



Baština Akademije nauka i umjetnosti Bosne i Hercegovine

Artificial Intelligence in Industry 4.0: The future that comes true: AI

Karabegović, Isak; editor

2024-09-17

<https://bastina.anubih.ba/handle/123456789/791>

Preuzeto s Baštine Akademije nauka i umjetnosti Bosne i Hercegovine

<https://bastina.anubih.ba/>



AKADEMIJA NAUKA I UMJETNOSTI BOSNE I HERCEGOVINE
АКАДЕМИЈА НАУКА И УМЈЕТНОСТИ БОСНЕ И ХЕРЦЕГОВИНЕ
ACADEMY OF SCIENCES AND ARTS OF BOSNIA AND HERZEGOVINA

Special Editions
Volume CCXV

Department of Technical Sciences
Volume 22

International Scientific Conference

**ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0:
THE FUTURE THAT COMES TRUE**

Editor
Isak Karabegović

SARAJEVO 2024



ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0: THE FUTURE THAT COMES TRUE



AKADEMIJA NAUKA I UMJETNOSTI BOSNE I HERCEGOVINE
АКАДЕМИЈА НАУКА И УМЈЕТНОСТИ БОСНЕ И ХЕРЦЕГОВИНЕ
ACADEMY OF SCIENCES AND ARTS OF BOSNIA AND HERZEGOVINA

Posebna izdanja
Knjiga CCXV

Odjeljenje tehničkih nauka
Knjiga 22

Međunarodna naučna konferencija
UMJETNA INTELIGENCIJA U INDUSTRIJI 4.0:
BUDUĆNOST KOJA SE OSTVARUJE

Sarajevo, 3 - 4. oktobra 2024.

Zbornik radova

Urednik
Isak Karabegović

SARAJEVO, 2024.



AKADEMIJA NAUKA I UMJETNOSTI BOSNE I HERCEGOVINE
АКАДЕМИЈА НАУКА И УМЈЕТНОСТИ БОСНЕ И ХЕРЦЕГОВИНЕ
ACADEMY OF SCIENCES AND ARTS OF BOSNIA AND HERZEGOVINA

Special Editions
Volume CCXV

Department of Technical Sciences
Volume 22

International Scientific Conference
**ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0:
THE FUTURE THAT COMES TRUE**

Sarajevo, 3rd - 4th October 2024

Proceedings

Editor
Isak Karabegović

SARAJEVO 2024

ARTIFICIAL INTELLIGENCE IN INDUSTRY 4.0:

The future that comes true

Sarajevo, 3rd – 4th October 2024

Publisher

Academy of Sciences and Arts of Bosnia and Herzegovina

For the publisher

Academician Muris Čičić

Organizing committee

Isak Karabegović, PhD

Ahmed Kovačević, PhD

Namik Hrle, PhD

Sead Pašić, PhD

Safet Isić, PhD

Samir Vojić, PhD

Mirha Bičo Čar, PhD

Mehmed Mahmić, PhD

Ermin Husak, PhD

Editor

Corresponding member Isak Karabegović

Reviewers:

Prof. Hazim Bašić, PhD, Prof. Domenico Guida, PhD, Prof. Safet Isić, PhD, Prof. Avdo Voloder, PhD, Academician Namik Hrle, Prof. Sead Pašić, PhD, Prof. Dejan Lukić, PhD, Prof. Jasmin Jahić, PhD, Prof. Maida Čohodar Husić, PhD, Prof. Marinko Maslar, PhD, Prof. Alempija Veljović, PhD, Prof. Živorad Kovačević, PhD, Prof. Vladimir Tankonogiyi, PhD, Prof. Žoran Miljković, PhD, Doc. Adnan Omerhodžić, PhD, Prof. Emir Kurtović, PhD, Prof. Aida Habul, PhD

Translation and lecturer for English

Mirela Midžić

DTP

Mehmed Mahmić

The press

Štamparija Fojnica d.d.

Fojnica

Printing

100

Sarajevo 2024

EBSCO

ISBN 978-9926-410-99-5

CIP zapis dostupan u COBISS sistemu Nacionalne i univerzitetske biblioteke BiH pod

ID brojem 61285126

CONTENTS

Preface	1
Jasmin Jahić INFLUENCE OF ARTIFICIAL INTELLIGENCE ON METHODOLOGIES AND PROCESSES FOR ENGINEERING SOFTWARE-ENABLED SYSTEMS IN INDUSTRY 4.0	3
Marija Mojsilović, Selver Pepić, Muzafer Saračević APPLICATION OF ADVANCED TECHNOLOGIES OF ARTIFICIAL INTELLIGENCE IN THE OPTIMIZATION OF PRODUCT QUALITY IN INDUSTRY	23
Zoran Miljković, Aleksandar Jokić, Milica Petrović MACHINE LEARNING WITHIN INDUSTRY 4.0: FROM DECISION TREES TO VISUAL TRANSFORMER ARCHITECTURE	41
Aleksandar Rodić BUILDING AI-SUPPORTED COLLABORATIVE AWARENESS IN INDUSTRIAL HUMANOID: CONCEPTUAL FRAMEWORK AND METHODOLOGICAL APPROACH	62
Lejla Banjanović-Mehmedović, Anel Husaković, Azra Gurdić Ribić, Naser Prljača, Isak Karabegović ADVANCEMENTS IN ROBOTIC INTELLIGENCE: THE ROLE OF COMPUTER VISION, DRL, TRANSFORMERS AND LLMs	94
Sadko Mandžuka, Edouard Ivanjko, Krešimir Vidović USE OF ARTIFICIAL INTELLIGENCE IN TRAFFIC TECHNOLOGY AND TRANSPORT	128
Vidosav D. Majstorović, Sonja Dimitrijević, Vladimir Simeunović, Dragan Stošić ARTIFICIAL INTELLIGENCE AND ITS APPLICATION IN MANUFACTURING	139
Samir Lemeš PROMPT ENGINEERING	159
Viktorio Malisa MACHINE SAFETY AND WORKPLACE IMPLICATIONS OF ARTIFICIAL INTELLIGENCE	171

Isak Karabegović, Ermin Husak, Samir Vojić, Edina Karabegović, Mehmed Mahmić TREND INNOVATION OF ARTIFICIAL INTELLIGENCE AND ROBOTIC TECHNOLOGY: IMPLEMENTATION IN ADVANCED ROBOTIC SYSTEMS	179
Savo Stupar, Elvir Šahić, Maida Cico, Mirha Bičo Ćar THE ROLE OF BLOCKCHAIN TECHNOLOGY IN THE TRANSFORMATION OF DIGITAL MARKETING	201
Amra Bratovcic IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE, SMART SENSORS, ROBOTS AND DIGITAL TRANSFORMATION IN FOOD AND AGRICULTURAL PRODUCTION	221
Edisa Drekoović, Isak Karabegović, Žaklina Teofilović A COMPREHENSIVE VIEW OF THE APPLICATION OF AI IN RECRUITMENT AND SELECTION	234
Osman Lindov AI SOLUTIONS FOR SUSTAINABLE, SAFETY AND RESILIENT TRANSPORT AND MOBILITY	251
Lemana Spahić, Almir Badnjević, Asim Kurjak, Lejla Gurbeta Pokvić INTEGRATION OF ARTIFICIAL INTELLIGENCE-BASED SYSTEMS IN DIAGNOSTIC PATHWAYS: TRUEAID CASE STUDY	270
Madžida Hundur Hiyari, Nejra Merdović, Faruk Bećirović, Emina Mrđanović, Adna Softić LIVER DISEASE CLASSIFICATION USING MACHINE LEARNING	280
Nemanja V. Majstorović, Sonja Dimitrijević, Aleksandar Todorović ARTIFICIAL INTELLIGENCE AND ITS APPLICATION IN DENTISTRY	288
Amina Radončić PREDICTING BEHAVIORAL PATTERNS IN BRAZILLIAN WHITEKNEE TARANTULAS – A MULTI-OUTPUT MACHINE LEARNING ANALYSIS OF DIETARY EFFECTS IN ACANTHOSCURRIA GENICULATA	309
Faruk Bećirovic, Nejra Merdovic, Madžida Hundur Hiyari, Merima Smajlhodžić-Deljo, Irma Salkic, Adna Softic ASSESSING THE ACCURACY OF LOGISTIC REGRESSION AND BERT IN SENTIMENT ANALYSIS AND MENTAL DISORDER CLASSIFICATION	317
LIST OF AUTHORS	327

PREFACE

The Academy of Sciences and Arts of Bosnia and Herzegovina (ANU BiH) continuously monitors and studies the development and implementation of new technologies in all industries. One of the latest contributions to these processes is the International Conference entitled "Artificial Intelligence in Industry 4.0: The Future that comes true", organized by the Department of Technical Sciences. Currently, the world is facing major changes. Research and development of innovation in new technologies, as well as rapid pace of implementation, especially digitalization and automation, have a major role in shaping the future world. The world has undergone a transformation with technological innovation that is constantly changing. It is necessary for all of us to adapt to the coming changes. A new era with all its challenges is ahead of us, a business climate limited by uncertainty, increased competitiveness, rapidly evolving technologies, infrastructure engineering, growing uncertainties, etc. Research seeks to understand the key challenges and opportunities that are accelerating in order to achieve strategic and competitive advantages. The development and implementation of new technologies is motivated by technical and economic reasons, such as: improving the quality of finished products (machining, etc.), reducing working hours (assembly processes), increasing the degree of homogeneity - constant quality (in all processes related to robotic application), increasing the speed of safety (in aggressive, flammable, explosive and other areas with a high degree of protection), reducing the required routine workforce and process repeatability, minimizing costs production and overall maintenance, meeting the requirements of the competition, as well as customer requirements. We are at the very beginning of the fourth industrial revolution. Industry 4.0 is a production concept in which everything is networked, so that in the production processes of machines, devices or their sensors are connected wirelessly, but also connected to a system that can make decisions driven by a large amount of data. Although Industry 4.0 concept is being implemented in our environment, the concept itself is not yet widespread. It is expected that the implementation of this concept will improve and enhance all aspects and segments of human life. Within the fourth industrial revolution, a new value chain was formed, which primarily relied on the cyber-physical system (CPS) and its associated service, most often realized in the cloud (Cloud Computing). It is necessary to create a framework that describes the key issues and emphasizes possible answers. Major changes in all industries, including new business methods, transformation of production systems, consumption, delivery and transport itself, are happening due to the implementation of new technological advances such as: robotics and automation, internet of things (IoT), big data, 3D printing, smart sensors, radio frequency identification (RFID), virtual and augmented reality, artificial intelligence, advanced security systems, etc. Industry 4.0 has an impact on companies by increasing automation and optimization, constant monitoring, personalization and customization, data mining, virtual reality, improving the work environment, increasing the use of robotics, etc. The implementation of Industry 4.0 in companies requires employees who are capable for the tasks of the future and who are certainly willing, able and educated to develop the future. The world is at the epicenter of artificial intelligence. Generative artificial intelligence unlocks the door to an infinite realm of possibilities, providing machines and devices with the power to create, innovate and mimic human behavior with high

precision. Artificial intelligence as the base technology of Industry 4.0 is applied in various sectors, such as production processes, medicine, logistics and security, and it can transform the work processes of the future, reduce costs and encourage innovation. When we talk about the development potential of artificial intelligence in the Western Balkans, as well as in Bosnia and Herzegovina, not enough attention has been paid to this sector, and the potential is extremely large. We are not only on the threshold of something extraordinary, we are already in it. The rapid development of Industry 4.0 technologies, and artificial intelligence itself, has led to increasing concern among experts. Leading artificial intelligence experts have signed a call to pause all massive artificial intelligence experiments until we can fully understand the potential risks and consequences. This means a recommendation to tread carefully and wisely, because the path we choose today will shape the very course of our technological evolution tomorrow.

We must note that the success or failure of the implementation of artificial intelligence as a technology of Industry 4.0 in all countries in the world, including the Western Balkans, lies in the hands of all participants in the chain, from producers to end users or customers. The Academy of Sciences and Arts of Bosnia and Herzegovina thanks the Society for Robotics of Bosnia and Herzegovina, and the Foreign Trade Chamber of Bosnia and Herzegovina, for co-organizing the conference. The organizers of the conference hope that the scientific and professional public will especially benefit from the vast knowledge and experience that experts in this field have summarized in their presentations.

This conference is an opportunity to connect the academic community with the real sector and achieve cooperation, and we hope that it will encourage inspiring discussions and contribute to the development of artificial intelligence and Industry 4.0 in Bosnia and Herzegovina, the Western Balkans and the world.

On behalf of all those who participated in organizing the conference and publishing the Proceedings, I would like to thank all the employees of the Academy of Sciences and Arts of Bosnia and Herzegovina for their great commitment to the success of this conference.

Sarajevo, September, 2024

*Corresponding member Isak Karabegović, editor
Academy of Sciences and Arts of Bosnia and Herzegovina*

Influence of Artificial Intelligence on Methodologies and Processes for Engineering Software-Enabled Systems in Industry 4.0

Jasmin Jahić*¹

Abstract: *Software currently presents is a corner stone of systems in Industry 4.0 (I4.0). To engineer software for these systems, engineers follow different methodologies and processes. These methodologies and processes aim to systemise production of high-quality software systems and make it possible to reproduce success in software engineering projects.*

With the introduction of AI in software engineering, actions that engineers perform are changing. Consequently, challenges and responsibilities of developers change. It becomes valid to ask: how will processes and methodologies in software engineering change with the introduction of AI? That means, what will be the new challenges that software engineering methodologies and processes need to solve, and which current challenges will simply disappear or become irrelevant. To answer these questions, in this paper, we abstract and summarise actions and aims of processes and methodologies in software engineering. We make predictions of what is it that humans bring to the table when it comes to software engineering, where can AI assist humans, and where AI has potential to fully replace humans. We discuss these predictions in the context of quality properties of I4.0 systems (e.g., security, safety), which must be taken into account when engineering I4.0 software-enabled systems.

Keywords: *Industry 4.0, AI, software engineering, software architecture, methodologies, processes*

1. Introduction

Industry 4.0 aims towards manufacturing flexibility, increased efficiency, and higher productivity[1]. To achieve these, it relies on communication, information, and intelligence technologies. The corner stone and the key enabler of all these technologies is software.

Engineering of software system is not an easy task. It involves many more actions than simply writing source code (Figure 1). To solve any problem using software, it is first necessary to solve it on a logical level. Programming languages, and software libraries serve to express logical solutions using source

*¹University of Cambridge, UK
E-mail: jj542@cam.ac.uk

code. Scaling the development of software solutions is a management and soft-skills challenge in software engineering. To produce solution faster and with adequate quality, it is necessary to organise development using development methodologies (e.g., as one would organise work in factories to increase throughput). These methodologies facilitate productivity, collaboration, and help to create high-quality software that fulfils its business goals. These methodologies are not of much use if they are not supported by proper processes and infrastructure. These days, source code version control, DevOps, and other infrastructure support development methodologies. But with time, software grows and it becomes hard to manage it, in particular when it comes to introducing new features. This is especially the case when software, besides providing desired functionality, needs to meet explicitly defined quality requirements (e.g., security, safety). To deal with these challenges, architects engage into system-level engineering activities. Software system architecture enhances communication, provides guidance, and serves to limit solution space for developers to ensure that software will meet its quality properties. Finally, all the activities that are performed in a software engineering project serve to fulfil business goals. These goals are not necessarily related to financial gains, but also can be centred around philanthropy or solving societal problems. As we can conclude from this overview of actions in software engineering, creating and extending software-enabled systems is a very complex undertaking. Adding AI into software engineering has potential to make it easier to deal with some of the challenges associated with presented activities and change the way how we create software systems. However, at this point, it is still not clear what the extent of those changes will be. In this paper, we focus on processes and methodologies in software engineering impacted by adoption of AI and try to make our predictions considering engineering of I4.0 systems.

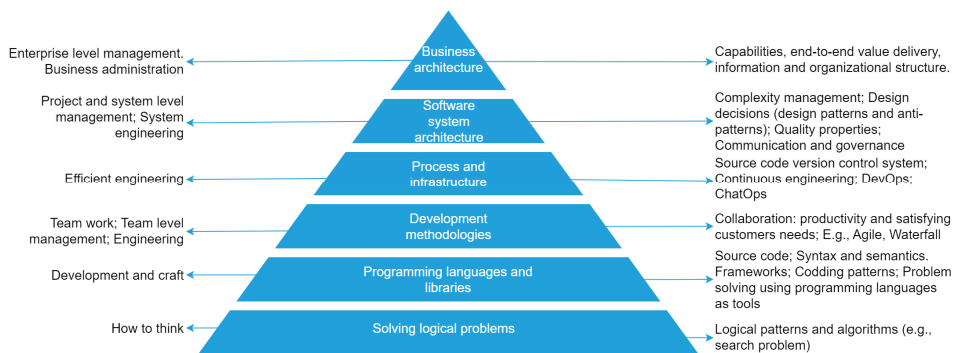


Figure 1: Actions involved in creating software-enabled systems

Besides guaranteeing that software meets its functional requirements, software engineers have to also ensure that their software-enabled systems meet quality

requirements. Depending on the system type, industry, number of users, context of deployment, and many other factors, different quality properties of software are of importance when designing such systems. Typical examples of system types in Industry 4.0 are systems focusing on smart manufacturing facilities, Internet of Things (IoT) devices, and AI-enabled cyber-physical systems (CPS²) such as (semi-)autonomous vehicles. Some of the common quality properties of these systems are real-time constraints, safety, security, power consumption, and certification constraints (compliance with industrial standards often requires following engineering methodologies and processes). While in general information systems these quality properties are often important, in I4.0 these quality properties are the main system driving factors. If some of these requirements is not met, then system can be considered a failure. In case of safety-critical systems, a failure to ensure safety levels can lead to loss of human life or huge financial damage. To ensure that system qualities will be met with solutions, engineering of these I4.0 systems often requires significant effort. This is especially the case due to heavy coupling between software and hardware (effectively characterising them as embedded systems). Introduction of AI into such engineering could potentially make the engineering effort less demanding and help engineers to create systems that will solve more complex problems, while at the same time ensuring that these systems optimally use available computing resources and minimize power consumption.

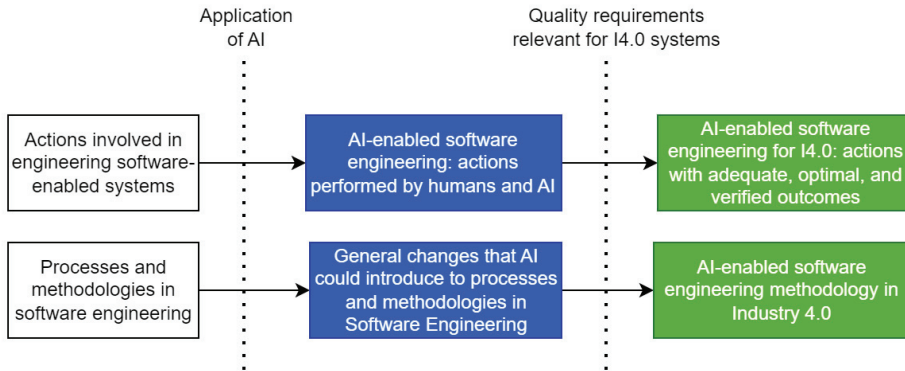


Figure 2: Research methodology

Therefore, in this paper we seek to answer the following question: *How will AI change actions in software engineering performed by humans, and in particular how will that change affect existing software engineering methodologies and processes for software-enabled systems in Industry 4.0?* Methodologies and processes are an important part of software engineering mainly responsible for

²https://ecssria.eu/2024_1.4

ensuring that software meets its quality properties. Changes in them will have a profound impact on the way we design, implement, test, deploy, monitor, fix, and extend software solutions. To make predictions on how introduction of AI in software engineering could change its processes and methodologies, we first discuss general changes that AI could introduce in these areas (Figure 2). Then, we gather quality requirements relevant for I4.0 systems and map them to processes and methodologies, as well as software engineering actions, responsible for creating solutions. The outcome of this research is one of the first attempts to predict how AI will change software engineering in I4.0. We also summarise a list of open questions that future research in this domain should answer. This paper introduces two main contributions:

1. Predictions on how AI will change actions, processes, and methodologies in general software engineering.
2. Predictions and discussion on how changes in general software engineering due to adoption of AI will change engineering of I4.0 systems.

This paper is organised as follows. In Section 2, we introduce related work in this domain. In Section 3, we present an abstract view of existing methodologies and processes in software engineering. In Section 4, we present our view on the changes that AI will potentially introduce to engineering of software systems in I4.0. We conclude this paper in Section 5, with a list of open questions for future research.

2. Related Work

Waterfall model [2] was one of the first methodologies for software development. It largely tried to imitate manufacturing processes from factories with focus on creating and following a plan. However, even when the Waterfall was introduced, its authored expressed scepticism regarding some of its parts. The biggest among the issues with Waterfall were long delivery cycles (time from requesting a change in software and delivering a new software version containing that change) and rigidity in following a plan instead the ability to respond to a change request. Agile³ manifesto tried to put the focus on fast delivery of incremental software updates.

While this has indeed ensured that the created software has desired functionality, lack of long-term plan (roadmap) in Agile often leads to naturally grown architectures. In such architectures, it is very hard to ensure system quality properties such as safety and security, as the focus is on short-term increments instead of long-term system behaviour. Continuous engineering (CE) [3] put the focus on creating a link between business strategy, development, and

³<https://agilemanifesto.org/>

engineering operations. The goal of CE is a constant deployment and delivery of new software versions that are well tested (some companies are capable of delivering and deploying several software versions in the same day). To add actions in this process related to system-level planning (software system architecture), researchers and practitioners introduced continuous architecture [4]. This approach combines continuous engineering with dedicated time and effort for system design and long-term planning, with these actions running in parallel.

Industry 4.0 systems often have strong quality requirements (e.g., safety, security, power consumption). There have been several attempts to adjust Agile to work for engineering I4.0 systems and support their quality properties[5][6][7]. Besides these, one approach to software development in I4.0 that has gained a lot of popularity in recent years is model driven development. Model driven development focuses on abstraction of component logical actions from implementation, and generation of source code from high level models[8]. It enables planning on a higher abstraction level and configuration of functionality through predefined sets of parameters.

When it comes to adoption of artificial intelligence in software engineering, certain tools are making great breakthroughs (e.g., CoPilot⁴). Scientific research also demonstrates benefits of AI in software engineering for writing source code and for performing associated actions (mostly by using large language models, such as ChatGPT)[9][10][11]. Other works focus on how AI can assist software architects in creating system design [12]. Despite benefits that the existing approaches demonstrate, they still suffer from low accuracy and low reliability [13][14][15]. There are a quite few attempts to discuss ethics around using AI solutions for engineering of software system. One of the most fundamental works in this domain is the Copenhagen manifesto [16]. It revolves around an idea that AI in software engineering must be human-centred and lists 12 principles to guide adoption of AI in software engineering.

To the best of our knowledge, this work is a pioneering work in both trying to predict i) how AI will change actions, processes, and methodologies in general software engineering, and ii) how changes in general software engineering due to adoption of AI will change engineering of I4.0 systems. While other approaches work on individual challenges, we take the wider software engineering scope into the account and aim to set a track for the whole discipline of software engineering.

⁴ <https://github.com/features/copilot>

3. Common Actions, Methodologies and Processis in Software Engineering

Engineering software systems comprises several actions (Figure 3). Different stakeholders (e.g., customers, business, marketing) have their wishes and concerns (a stakeholder is anyone affected by or concerned with an outcome of a project). Software architects consider wishes and concerns from stakeholders to understand the scope of the system and use them to define drivers that capture system’s problem domain.

In the solution domain, architects seek to make decisions that will address problem domain (system requirements and drivers) and create conceptual solutions that developers need to implement. To do so, architects rely on their own knowledge (experience) and their creativity. If we assume that there is a space that contains all adequate solutions for drivers of a certain system, then only a portion of those is known to architects or documented in general. The challenge for architects is how to reuse solutions unknown to them, but otherwise documented. Conceptual solutions created by architects serve as a blueprint for software implementation. It is necessary to support implementation effort with operations, such as building binaries, testing them, integrating contributions from different teams into a software solution, deployment of different software versions to target environment and their delivery to end users. As part of system execution, it might be also necessary to monitor the system and take adequate actions based on the monitoring results.

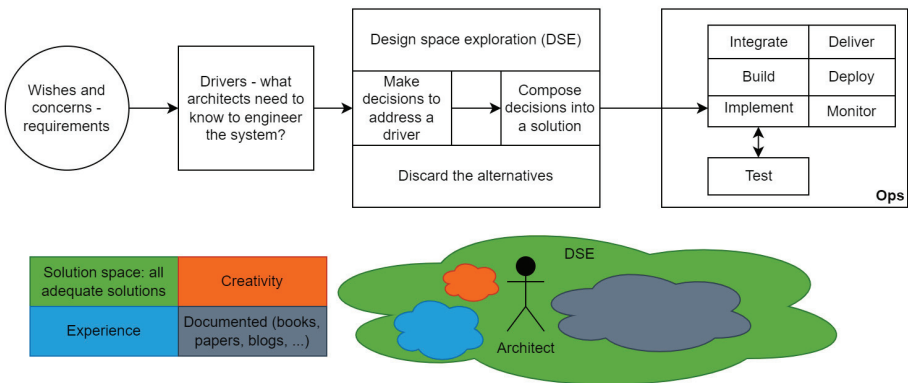


Figure 3: Actions involved in engineering software-enabled systems

Actions discussed above exist in every software project to a greater or a lesser degree. Engineers of software systems can organise their work (mentioned

actions) using different procedures and schedules (e.g., Waterfall[2], Agile⁵). Let us discuss one example of such organisation based on combination of Agile Scrum⁶(Figure 4) and continuous architecture [4].

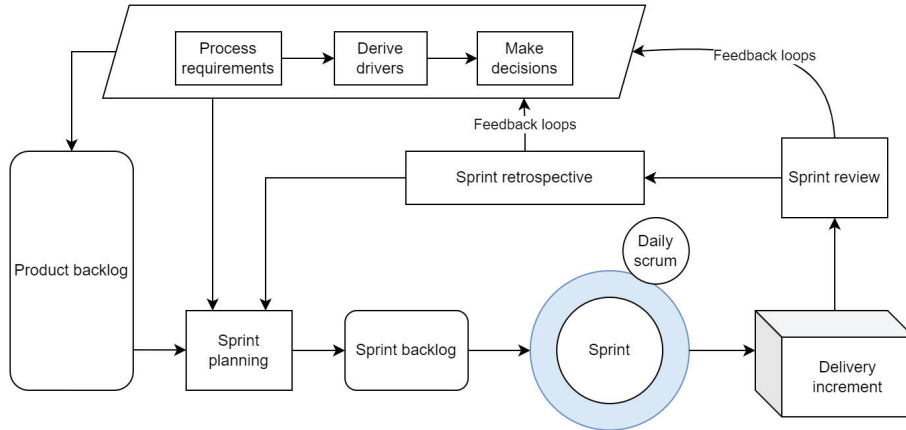


Figure 4: An example of organisation of continuous architecture using Scrum development methodology

In this example of the work organisation (processes and methodology), architects process requirements, derive system drivers, and based on these make architectural decisions about the system. The outcome of these activities is a (updated) project roadmap. Engineers of a system use the roadmap to create/update a prioritized list of work (**product backlog**). In Scrum, engineering activities are organised in sprints. A sprint usually takes two to four weeks. During **sprint planning**, engineers prioritise activities from the backlog that can reasonably be implemented within the duration of one sprint (**sprint backlog**). Each day, engineers have short standup meetings (**daily scrum**), during which they provide brief update on their work. At the end of the sprint, engineers deliver an increment to the previous product (**delivery increment**). During **sprint review**, the team presents to stakeholders the delivery increment and gathers feedback from them. Internally, the team working on the system performs **sprint retrospective** to determine what went well during the sprint and what could be improved. Outputs from the sprint review and sprint retrospective are fed to the system architects who then decide about the roadmap.

It is important to notice here that actions specified in Figure 3 are present in engineering every software-enabled system. They can be explicitly specified and

⁵<https://agilemanifesto.org/>

⁶<https://www.scrum.org/resources/what-scrum-module>

performed or handled implicitly. Engineers working on software systems find a way to organise their work. That organisation can conform to one of the existing development methodologies (e.g., Waterfall) or can be customised. Regardless, there is always either explicit or implicit organisation of work.

4. Influence of AI on Methodologies and Processes in Software Engineering

In the previous section (Section 3), we have discussed common actions, methodologies, and processes in software engineering. Let us first discuss how artificial intelligence is changing these (Section 4.1). To understand how is this relevant for I4.0 systems in particular, we discuss quality properties of these systems derived from different relevant industry standards (Section 4.2). Finally, we try to predict how AI will change actions in software engineering performed by humans and consequently processes around such software engineering, with the focus on systems in I4.0.

4.1. AI and Software Engineering

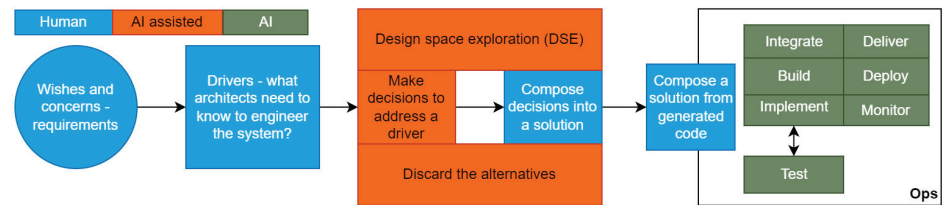


Figure 5: AI-enabled software engineering: coverage of work by humans and AI

In the last few years, we have witnessed significant developments in artificial intelligence and its application in software engineering (as discussed in Section 2). This development will have a significant impact on actions that humans perform in (future) software engineering. The key questions we are starting to ask is, which of the software engineering actions will be performed by humans, which of these actions will be supported by AI, which actions are suitable to be fully replaced by AI, and what are potentially new actions that engineers will perform. In other words, if AI is able to replace engineers doing some actions, what is it that humans bring to the table? To answer this question, we observe influence that introduction of AI has in the context of software engineering (Figure 5).

Understanding the context of a problem that a system should solve still remains in the domain of human intelligence. Even the most advance AI models today face the challenge of comprehending the problem domain. One example that

illustrates this is the case of stopping Waymo's robotaxis after they encountered a man with a t-shirt that contains STOP sign⁷. Another example is the AI chatbot provided by Google that advised people to eat one stone per day as stones might contain minerals⁸. Previous research also confirms these conclusions regarding AI and its struggle with the contextual understanding of problems it needs to solve. Experiments with application of AI in software system architecture have demonstrated that AI models are unable to fully comprehend context for which they are generating solutions[12]. These experiments show that AI cannot differentiate between different abstraction levels – effectively meaning that it cannot reason what is more important to architects who consume such AI generated solutions. If AI will ever be able to replace a human or be of an assistance to humans in this domain, then such AI would be (very close to) artificial general intelligence (AGI)[17] and would have expert domain knowledge about domain in which we try to apply it. Furthermore, as AI cannot grasp the context of stakeholders, their background, wishes and concerns, it also cannot differentiate between what is it that indeed drives software projects (we simply cannot capture all implicit requirements, it still takes a human to know). Such AI would need to have knowledge of advanced social skills and broad civic and domain knowledge. Consequently, it means that AI is unsuitable for deriving drivers. In conclusion, we can say that for defining a problem domain (motivation for creating a system, reasons for it to exist, domain in which it should exist with associated limitations and assumptions), AI at the moment (and in foreseeable future) is not of much value.

Architectural decisions are first artefacts in the solution space that engineers produce when developing software-enabled systems. Previous research [12] has demonstrated that AI can assist with interpreting knowledge that exists in software architecture domain and assist in this process. This research shows that AI is also suitable for deriving new solution patterns previously not considered by architects. Therefore, for design space exploration, already today AI can act as an assistance tool, as it can suggest solutions for individual (smaller scale) architecture problems. However, AI does not have a notion of right and wrong (correct and incorrect, true or false). Furthermore, it is still struggling to compose solutions that would be adequate for a combination of explicitly expressed and assumed systems' requirements composing a larger problem domain. Therefore, composing individual system architecture decisions into a system solution remains in the hands of human engineers and architects.

⁷<https://www.carexpert.com.au/car-news/how-a-t-shirt-stopped-this-autonomous-car-in-its-tracks>

⁸<https://www.cnet.com/tech/services-and-software/glue-in-pizza-eat-rocks-googles-ai-search-is-mocked-for-bizarre-answers/>

When it comes to code generation, test generation, and supporting other operations in software engineering, AI proved to perform quite well[18][11]. For now, AI still performs these activities on a small scale. But as it has been shown to be the case in other engineering areas, there is no doubt that AI is capable of generating solutions that outperform human engineers due to its ability to extract complex patterns from data or connect huge number of data points and data dependencies not obvious to humans when creating a solution⁹. Therefore, we can expect that in future AI can take over writing the code for individual problems, generating tests, interpreting the code to new engineers, writing scripts for deployment and delivery, and maintaining the whole operations infrastructure. However, what still remains to be a challenge is generation of full solutions for dedicated design. Much of the design that architects provide is abstracted. In such context, developers have implicit understanding with architects how to implement systems to conform to the prescribed architectural design. Eliminating humans from the loop would mean eliminating this implicit understanding. Therefore, it is realistic to expect that humans will still perform the role of an integrator of code components individually generated by AI.

4.2. Key Software Quality Properties in Industry 4.0

Industry 4.0 covers a large group of system types. These are buildaround artificial intelligence, cyber-physical systems, Internet of Things (IoT), and cloud computing. These underlying technologies enable the most important properties of I4.0, such as connectivity, distributed computing, and human-centric engineering. Let us observe key quality architectural properties of these systems and let us discuss standards that prescribe their engineering.

In terms of engineering, adopting artificial intelligence these days is mainly associated with two sets of actions: training and inference. While programming and using systems based on trained AI models is not a particularly challenging task, what remains as a big issue is the **power consumption** when running these models. The computing power required for AI is doubling every 100 days [19]. It is projected that this requirement will increase by more than a million times over the next 5 years [19]. Besides power consumption, another challenge is what does the adoption of AI mean for certification of **safety**-critical products (e.g., autonomous vehicles). Besides challenges with energy consumption, second greatest challenge with AI is the inability to properly test AI-based systems (especially significant constraint for safety-critical systems).

In cyber-physical systems (CPS), challenges are present around **safety** and **security** in the first place. Furthermore, there are also challenges related to management of such systems (**management of complexity, maintaining legacy**

⁹<https://spectrum.ieee.org/ai-3d-printing-better-ac>

systems, updating and extending the systems). When CPS are a bases for other systems, such is autonomous vehicles or avionics, there are strict standards that prescribe engineering processes and methodologies. Only by following these standards, it is possible to receive necessary **certification** for putting these systems on market. One of the most distinct quality properties in these systems are **real-time** constraints.

Systems composing Internet of Things (IoT) are dominated by low-power sensors and low-power communication and control devices. Therefore, in these systems **power consumption** and **secure communication** are the most important quality drivers. Besides these quality properties, **reliability** of such systems is important as well as the ability to **dynamically extend** them with adding new sensor and computing nodes.

Cloud-based systems are mostly concerned with **optimal use of available computing hardware** and with **power consumption**, where these two properties are tightly coupled. For example, finding an optimal scheduling of workloads for available computing nodes greatly saves power consumption, making the whole system more energy efficient.

Here is a summary of the most relevant standards that exist in these domains:

- IS26262¹⁰ is a safety standard for road vehicles. It prescribes development of safety-critical products on system, hardware, and software level.
- IEC 61508¹¹ is an international standard for Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems (ISO 26262 is an adaptation of this standard). It prescribes different safety integrity levels, and according to each of these levels, it prescribes types of required testing (e.g., functional coverage for the lowest safety level, MC/DC coverage for the highest safety level).
- ISO 21448¹² focuses on safety of the intended functionality. For that, it provides guidance on the applicable design, verification and validation measures.
- IEEE 2851¹³ is a standard for Functional Safety Data Format for Interoperability within the Dependability Lifecycle. It describes methods, description languages, data models, and database schema to enable the exchange/interoperability of data across all steps of the product's lifecycle.
- ISO/SAE 21434¹⁴ is a cybersecurity engineering standard for concept, product development, production, operation, maintenance and

¹⁰<https://www.iso.org/obp/ui/#iso:std:iso:26262:-1:ed-2:v1:en>

¹¹<https://www.tuvsud.com/en-us/services/functional-safety/iec-61508>

¹²<https://www.iso.org/standard/77490.html>

¹³<https://standards.ieee.org/ieee/2851/10780/>

¹⁴<https://www.iso.org/standard/70918.html>

decommissioning of electrical and electronic (E/E) systems in road vehicles, including their components and interfaces.

- ISO/DPAS 8800¹⁵ standard prescribes safety-related properties and risk factors impacting the insufficient performance and malfunctioning behaviour of Artificial Intelligence (AI) within a road vehicle context.
- ISO/IEC TR 5469¹⁶ standard describes the properties, related risk factors, available methods and processes related to i) use of AI inside a safety related function to realize its functionality, ii) use of non-AI safety related functions to ensure safety for an AI controlled equipment, and iii) use of AI systems to design and develop safety related functions.
- DO-178C¹⁷ standard focuses on software in airborne systems and their equipment. It focuses on tools, model-based development, verification, formal methods, and safety considerations.

From the summary of technologies commonly present in I4.0 and the summary of the relevant standards, we can conclude that the key quality properties relevant for I4.0 systems are:

- power consumption
- optimal use of available computing hardware
- safety
- security
- management of complexity, maintenance of legacy systems, updates and (dynamic) system extensions
- real-time
- reliability
- certification

At the same time, while meeting these requirements, systems also must be performant to exercise its functionalities, cost-efficient, and profitable.

4.3. AI-enabled Software Engineering in Industry 4.0

In Section 4.1, we have summarised actions, methodologies, and processes in software engineering that AI could take over, could assist with, and those that require exclusive human involvement. In the previous section (Section 4.2), we have summarised the most important quality properties of I4.0 systems. Let us merge these two and discuss how AI can potentially change software engineering for I4.0. Outcomes of actions in software engineering can be

¹⁵<https://www.iso.org/standard/83303.html>

¹⁶<https://www.iso.org/standard/81283.html>

¹⁷<https://www.rtca.org/training/do-178c-training/>

adequate (good enough), optimal, and formally verified. AI does not provide guarantees about formal verification of its generated solutions. However, could help to guide formal verification of solutions. Furthermore, considering large number of parameters that condition trade-offs between system quality parameters, AI can provide optimal solutions that humans cannot even comprehend.

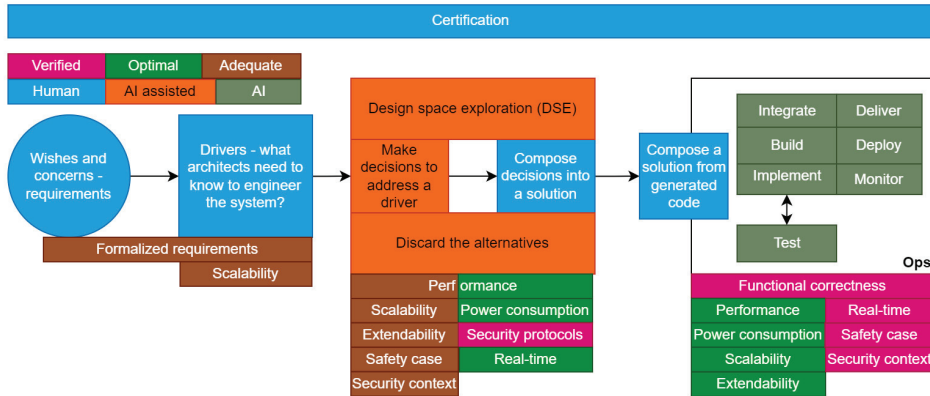


Figure 6: Different levels of quality of solutions: adequate, optimal, verified, and how AI can help to support them

As previously stated, understanding the problem domain is a task for humans. AI hardly can help in this domain.

On the other hand, creating architectural solutions is an action where AI can assist with architectural decisions. While providing optimally performant design for the whole system is hard to generate (especially considering that AI cannot even grasp the context in which the system should operate and architects cannot provide all possible details), providing optimal solutions for individual problems is possible. In that context, AI could help to generate optimal solutions such as scheduling and deployment configurations, considering different workloads, communication latencies, network configurations, and available computing nodes. Considering that architecture is an abstraction, it does not contain all the details that could facilitate optimal design performed by AI. In that context, it is also important to emphasize that systems evolve and do so in a way that is hard to predict. While AI can assist with design decisions, considering quality properties such as scalability, extendability, and building safety and security cases it is only possible to create adequate designs. On the other hand, considering concrete problems such as power consumption in a design context with enough information and real-time constraints (e.g., scheduling policies, priorities, available computing nodes), AI has potential to generate optimal design solutions. It is important here to emphasize once more the ability of AI to

discover dependencies between quality properties normally not obvious to humans. Therefore, besides other benefits, AI could be used to suggest optimal trade-offs between key quality properties. However, AI remains inadequate for verification of concrete design solutions, such is verification of security protocols. Verification of these still requires formal methods.

On the implementation level, however, AI has a lot more potential. We have created high-level programming languages to enable easier writing of code, management of large code bases, and management of complexity that stems from large code bases. However, there is no obstacle that prevents AI from generating solutions directly in assembly code, optimised for the target instruction set architectures (ISA). In that context, AI has a potential to generate optimised implementations in form of binaries (it still remains an open question how engineers would handle this, if AI is not able to generate full solutions but pieces of the solution that engineers have to integrate). In that context, AI could optimise binaries also for power consumption. Furthermore, generated solutions could take into the account target deployment platform. AI, with its ability to understand far more dependencies in software system solutions than humans, could optimally extend existing solutions with new features directly in the source code. Although AI cannot guarantee formal verification, we could expect that it generates functionally correct code, eliminating the need for testing on the lower levels (e.g., unit test level). This puts validation of the solution in the perspective, where AI testing could act as fuzzy testing[20] on higher abstraction levels (e.g., integration and system testing). Lifting the testing and validation into a more abstract domain definitely has benefits, as long as humans can keep some level of the validation and testing control. If AI is able to observe the functional rules of programming languages, we could claim that it (almost) achieves formally verified functionally correct source code. While AI cannot formally verify real-time constraints, safety case, and security in a broader context, it can help to generate solutions that are optimised towards these goals.

The major question, however, still remains. What does this mean for certification of such systems? At this point, we do not have a clear answer to this question, and it remains an open point for future work.

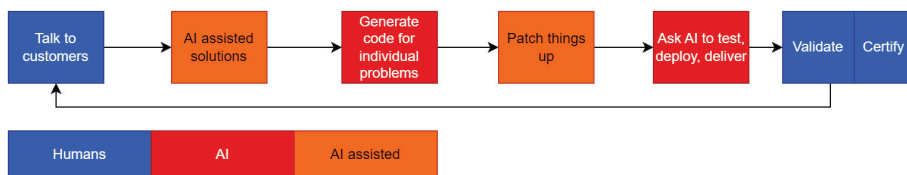


Figure 7: A prototype of an AI-enabled software engineering methodology

In Section 3 (Figure 4), we have introduced an example of methodology that supports continuous software architecture. Considering the potential benefits that AI could offer to software engineers and architects, the question is how that influences software engineering methodologies (Figure 7)? Does it deem certain actions unnecessary and creates a need for others? Interaction with those that define the problem domain still is in the domain of human work. In the solution domain, engineers can use AI tools to reach adequate (and in some cases optimal) solutions. If we are truly able to generate code snippets for well-defined problems, then the job of an engineer is to understand such solutions (to make sure that they do not do something different from the expressed requirements), and patch them into an integrated solution. Under the assumption that the generated code is functionally correct, engineers can use AI to test the solutions on the higher abstraction level and ask AI to generate scripts that will handle other operations (e.g., deploy, deliver). However, it is still the task of the humans to validate if the produced solutions satisfy requirements. If we cannot rely on AI to capture requirements and context around them, then we cannot rely on AI to validate solutions (impossible to validate if AI does not understand against what it needs to validate the solution).

Consequently, it is also the job of humans to deal with the certification of safety-critical products. There is no doubt that certification in future will change. It is quite likely that there will be a possibility to engage into a process of continuous certification through continuous compliance with standards [21], reducing the effort and time for verifying evolving safety-critical systems. However, it is also clear that humans have to remain the essential part of that process (because only humans understand the context in which solutions will perform).

One model that illustrates the discussion above in a quite nice way is (somewhat archaic but still relevant) V-model[22]. The V-model is a development process applied in many domains, including software system engineering. The main idea of the V-model is verification and validation of all actions in software projects. Validation is the assurance that a system meets acceptance and suitability criteria made by (external) stakeholders and their high-level requirements. Verification is the assurance that a system meets a defined set of design specifications, shows absence of bugs and meets quality specifications (for example, including standards).

The main elements of the V-model are:

- Concept of operations: describes stakeholder needs and the operating environment. This action considers writing the list of stakeholders and their operations in the system.
- Requirements and architecture: suggests collecting detailed requirements about how an ideal system should perform by analysing stakeholder needs

and suggesting techniques and approaches for meeting these needs. This action should also list requirements that are not feasible.

- Detailed design: the architectural design in this stage is detailed further so it is ready for implementation.

After implementation, the project undergoes testing and integration activities:

- System verification and validation. System tests are set up by a business team to test the overall functionality and quality of a system.
- Integration, test and verification. Integration tests verify if parts of the system that are developed independently can communicate, collaborate, and coexist among themselves. On the lowest abstraction level are unit tests. These tests verify if individual system parts can function correctly, isolated from the rest of the system.
- Operations and maintenance. Once tested, the project enters the operation and maintenance stage.

Although Waterfall-like structure behind V-model might not be suitable for the majority of project these days, it still captures essential actions and dependencies between them that each software engineering project should perform. When we apply the conclusions from this chapter on how AI is changing software engineering, we get division of work in V-model as illustrated in Figure 8.

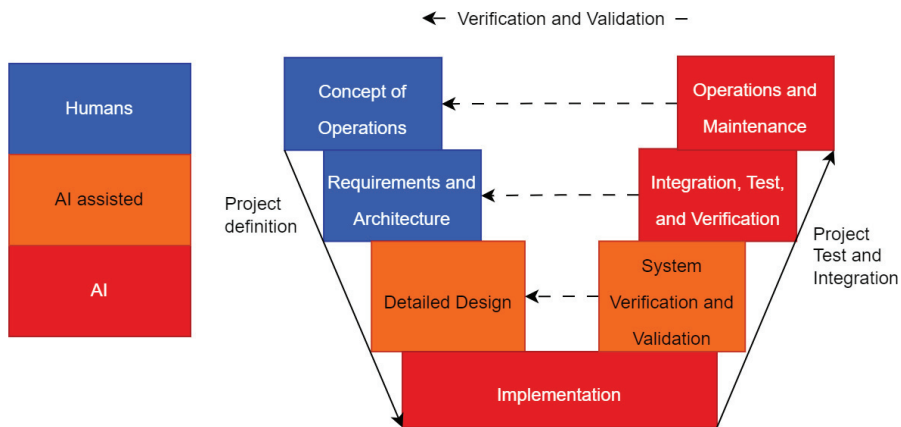


Figure 8: V-model complemented with AI-enabled software engineering: coverage of work by humans and AI

5. Conclusion

Adoption of AI in software engineering is current state of practice. This adoption will only increase in future. However, researchers and practitioners struggle to adopt AI in their engineering processes and methodologies in a way that would maximise their benefits. To deal with this problem, we have observed what kind of actions AI can take over from humans and where it can assist. Furthermore, we also discussed in this paper why some of the actions will most probably remain under the responsibility of humans. The conclusion is that adoption of AI in software engineering might completely re-shape this domain, in terms of how we organise work, responsibilities of developers, perceived value of delivery, and focus of invested effort in projects. For example, it might happen that writing code, tests, and scripts for all associated operations might represent only a small fraction of software costs, where bringing in other experts (such as experts on performance, distributed systems, UI, safety, security, etc.) might be the focus when engineering future systems.

In this paper, we have derived new sets of actions that engineers might be performing in AI-enabled software engineering for I4.0. We have suggested how to organise those actions into a new methodology. Finally, here we list a set of open questions that we consider that future research in this domain must try to answer:

- If AI increases productivity of software engineers to the point where a single engineer can deliver the same or more as previously a team of engineers, will we still need software development methodologies such as Agile? In that case, will the software engineering task be reduced in effort (to only connecting components generated by AI) and will we dedicate more resources to domain experts (performance architects, safety and security experts, UI experts, etc.)?
- If AI is able to generate correct code, will there still be a need for low level (e.g., unit level) testing?
- If AI can generate code for small problems, will the focus of engineers move to integrating small generated solutions to solve big problems?
- If AI can understand machine code, is there a need anymore for higher level programming languages? Can AI instead generate machine code from high level specifications?
- If AI can understand machine code, can it edit, update, and extend software on the machine code level directly?
- If AI can easily rewrite (machine code) solutions, does code have the same value anymore as (we think) it has today? Is there a need, in that case, for any documentation concerning source code itself?
- Can AI assist with guiding tools and processes for formal verification?

6. References

- [1] B. Chunguang, P. Dallasega, G. Orzes and J. Sarkis, "Industry 4.0 technologies assessment: A sustainability perspective," *International Journal of Production Economics*, pp. Volume 229,, 2020.
- [2] W. Royce, «Managing the development of large software systems: concepts and techniques,» chez *ICSE '87: Proceedings of the 9th international conference on Software Engineering*, 1987.
- [3] B. Fitzgerald et K.-J. Stol, «Continuous software engineering: A roadmap and agenda,» *Journal of Systems and Software*, vol. 123, pp. 176-189, 2017.
- [4] M. Erder et P. Pureur, *Continuous Architecture - Sustainable Architecture in an Agile and Cloud-Centric World*, Morgan Kaufmann, 2016.
- [5] R. Dovleac, A. Ionica et M. Leba, «Knowledge Management Embedded in Agile Methodology for Quality 4.0,» chez *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, Singapore, 2021.
- [6] H. B. Christensen, S. C. Jepsen et T. Worm, «Agile Architecting of Distributed Systems for Flexible Industry 4.0,» chez *16th Conference on Computer Science and Intelligence Systems (FedCSIS)*, Sofia, Bulgaria, 2021.
- [7] A. Marnewick et C. Marnewick, «The Ability of Project Managers to Implement Industry 4.0-Related Projects,» *IEEE Access*, vol. 8, pp. 314-324, 2020.
- [8] D. Schmidt, «Model-Driven Engineering,» *IEEE Computer*, vol. 39, n° %12, 2006.
- [9] J. Yang, H. Jin, R. Tang, X. Han, Q. Feng, H. Jiang, S. Zhong, B. Yin et X. Hu, «Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond,» *ACM Transactions on Knowledge Discovery from Data*, vol. 18, n° %16, 2024.
- [10] Z. Zhang, C. Chen, B. Liu, C. Liao, Z. .: Y. H. Gong, J. Li et R. Wang, «Unifying the Perspectives of NLP and Software Engineering: A Survey on Language Models for Code,» 10.13140/RG.2.2.32710.69440/1, 2024.

- [11] Z. Z. Chen, K. Ning, Y. Wang, J. Zhang, D. Zheng, M. Ye et Jiachi, «A Survey of Large Language Models for Code: Evolution, Benchmarking, and Future Trends,» <https://arxiv.org/abs/2311.10372>, 2024.
- [12] J. Jahic et A. Sami, «State of Practice: LLMs in Software Engineering and Software Architecture,» chez *21st IEEE International Conference on Software Architecture (ICSA 2024)*, Hyderabad, India, 2024.
- [13] M. Tufano, C. Watson, G. Bavota, M. D. Penta, M. White et D. Poshyvanyk, «An Empirical Study on Learning Bug-Fixing Patches in the Wild via Neural Machine Translation,» chez *2018 33rd IEEE/ACM International Conference on Automated Software Engineering (ASE)*, Montpellier, France, 2018.
- [14] M. Ciniselli, N. Cooper, L. Pascarella, D. Poshyvanyk, M. Di Penta et G. Bavota, «An Empirical Study on the Usage of {BERT} Models for Code Completion,» chez *18th IEEE/ACM International Conference on Mining Software Repositories, MSR 2021*, Madrid, Spain, 2021.
- [15] A. Mastropaolo, S. Scalabrino, N. Cooper, D. Nader-Palacio, D. Poshyvanyk, R. Oliveto et G. Bavota, «Studying the Usage of Text-To-Text Transfer Transformer to Support Code-Related Tasks,» chez *18th International Conference on Mining Software Repositories*, 2021.
- [16] D. Russo, S. Baltes, N. van Berkel, P. Avgeriou, F. Calefato, B. Cabrero-Daniel, G. Catolino, J. Cito, N. Ernst, T. Fritz, H. Hata, R. Holmes, M. Izadi et F. Khomh, «Journal of Systems and Software,» *Generative AI in Software Engineering Must Be Human-Centered: The Copenhagen Manifesto*, p. 10.1016/j.jss.2024.112115, May 2024.
- [17] S. Baum, «A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy,» *SSRN Electronic Journal*, vol. 10.2139/ssrn.3070741, 2017.
- [18] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang, W. Ye, Y. Zhang, Y. Chang, P. S. Yu, Q. Yang et X. Xie, «A Survey on Evaluation of Large Language Model,» *ACM Trans. Intell. Syst. Technol.*, vol. 15, n° %139, p. 45, June 2024.
- [19] S. Zhu, T. Yu, T. Xu, H. Chen, S. Dustdar, S. Gigan, D. Gunduz, E. Hossain, Y. Jin, F. Lin, B. Liu, Z. Wan, J. Zhang, Z. Zhao, W. Zhu, Z. Chen, T. S. Durrani, H. Wang, J. Wu et T. Zh, «Intelligent Computing: The Latest Advances, Challenges, and Future,» *Intelligent Computing*, vol. 2, n° %16, 2023.

- [20] A. Nappa et E. Blázquez, *Fuzzing Against the Machine: Automate Vulnerability Research with Emulated IoT Devices on Qemu*, Packt Publishing, Limited, 2023 .
- [21] T. Santilli, «Characterizing Software Architectural Metrics for Continuous Compliance in the Automotive Domain,» chez *21st IEEE International Conference on Software Architecture (ICSA 2024)*, Hyderabad, India, 2024.
- [22] K. Forsberg et H. Mooz, «The Relationship of System Engineering to the Project Cycle,» *Incose*, vol. 1, n° %11, pp. <https://doi.org/10.1002/j.2334-5837.1991.tb01484.x>, 1991.

Application of Advanced Technologies of Artificial Intelligence in the Optimization of Product Quality in Industry

Marija Mojsilović^{*1}, Selver Pepić¹, Muzafer Saračević²

Abstract: *This paper investigates the application of advanced technologies in the optimization of product quality in a dedicated industry. As the demands for high-quality products in this sector are increasing, so is the need for efficient optimization methods, which encourage the development and implementation of innovative technological solutions. Focusing on artificial intelligence, soft computing and related techniques enables improvement of the production process, identification of problems in early stages, prediction of potential defects and optimization of production parameters, with a deep consideration of how to improve quality control, reduce costs and achieve greater competitiveness in the market. The aim of the work is focused on the research of the input factors that influence the quality of the product. An analysis of literature and practical studies investigates how these technologies can improve processes in industry, exploring how they bring benefits and what challenges they can represent. Through this work, the door is opened for further development and implementation, to create more efficient processes and superior product quality.*

Keywords: *Advanced technologies, Quality, Artificial intelligence, ANFIS, Industry.*

1. Introduction

In the modern age, the dedicated industry faces increasing demands for high-quality products that must meet strict standards and specifications. To remain competitive in the market and meet the needs of their customers, companies in the dedicated industry are increasingly turning their attention to advanced technologies that allow them to optimize the quality of their products. Through the implementation of artificial intelligence, soft computing, and other innovative tools, these companies are creating new approaches to quality control, defect prediction, and manufacturing process improvement. This paper explores how advanced technologies are used in a dedicated industry to achieve superior product quality, reduce costs and increase market competitiveness. Through the analysis of the application of these technologies, we investigate how they are integrated into production processes, and identify key challenges and obstacles, as well as opportunities for further progress. This topic is essential

^{*1}Academy of Professional Studies Sumadija, College in Trstenik, Trstenik

²University of Novi Pazar, Department of Computer Sciences, Novi Pazar

in the context of modern industry, where technological progress is becoming a key factor in achieving success and long-term sustainability.

The importance of digital technologies for manufacturing services is often highlighted, but the current literature does not sufficiently address how companies can use digital methods to improve their service offerings. This article contributes to current theory by exploring how digitization can enable the development of servitization for manufacturers [1]. Digital technologies are increasingly becoming a key resource for the future competitiveness of modern organizations [1].

With the ever-increasing consumption of natural resources and growing market demands for high standards, the focus on sustainable development is becoming increasingly important in modern process industries, especially when it comes to product quality. Therefore, manufacturers are extremely interested in implementing new process monitoring techniques to improve product quality and process efficiency [2]. Product quality is the result of a combination of properties and characteristics that determine the extent to which the product can meet customer needs. The main goal of this study is the analysis of factors that influence purchase decisions, which includes the assessment of product quality and competitive prices [3].

2. Theoretical Settings of the Research

Software that uses the principles of soft computing, unlike traditional computing, is focused on using approximate models and providing solutions to complex real-world problems [4].

The application of soft computing explores the integration of artificial intelligence tools such as neural networks and fuzzy technologies into a comprehensive hybrid framework for solving real problems [5]. ANFIS training includes determining those parameters through the application of optimization algorithms. In the first ANFIS model, developed by Young, a hybrid learning approach was proposed for training [6]. Adaptive neuro-fuzzy inference system (ANFIS) is one of the most popular representatives of hybrid neuro-fuzzy systems [7][8]. Using neural networks, we can develop a fuzzy inference model that is used to estimate membership function parameters based on available input-output data. In the ANFIS network algorithm, the data is forwarded through four layers, applying the least squares method to determine the appropriate parameters [9].

In this research, the bell membership function was chosen because of its capacity to model non-linear data. The bell-shaped membership function is defined as follows:

$$\mu(x) = \text{bell}(x; a_i, b_i, c_i) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^{b_i}}$$

Training and evaluation of the ANFIS network was performed using the Matlab software package. Input-output data from experimental tests were collected to identify key parameters for a given output data using an ANFIS network, along with an estimation of the prediction error (RMSE). In this research, the accuracy of the prediction is evaluated through the least squares error, which can be expressed by the corresponding equation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - Q_i)^2}{n}}$$

Where P_i and Q_i are experimental and calculated data [9].

2.1. Methodology and data

The subject of research in this paper is the analysis of the impact of various input factors on the quality of products in the dedicated industry. Then identify the key factors that have the most significant impact on product quality and explore ways in which the application of advanced technologies can contribute to the optimization of these factors to achieve better product quality in a dedicated industry.

This research paper aims to investigate the influence of various input factors on product quality in a dedicated industry and to identify which of these factors contribute the most to improving or reducing product quality. To achieve an optimal solution and improve product quality, which is the main goal of this research, it is necessary to consider the influencing factors that will be key to the analysis. Various factors can significantly affect product quality, as shown in Table 1. Some of these factors include:

- Starting material - The casting as a starting material has a key influence on the quality of the product. Its purity, homogeneity and mechanical properties play a vital role in the final product. Any irregularities or impurities in the casting can result in a reduction in the quality and performance of the product, which can have negative consequences for the end user and the company's reputation. Therefore, it is important to carefully control and optimize the casting production process to ensure the high quality and competitiveness of the product on the market.

The range of 0 to 100 percent for starting material affecting product quality describes the spectrum of material quality from the lowest to the highest level. At 0% there is a material with numerous impurities and variations in structure, while at 100% it represents a material of high purity and homogeneity.

This range allows precise measurement of the quality level of the starting material and its influence on the quality of the final product.

- **Cutting tool** - The choice of cutting tool has a significant impact on the quality of the product. Parameters such as cutting tool material, blade geometry, cutting speed and cooling directly affect the accuracy, surface finish and durability of the product. Using the appropriate cutting tool with properly set parameters ensures efficient material removal without damage, resulting in high-quality products and reducing the need for additional processing. Therefore, careful selection and quality maintenance of cutting tools are key factors for achieving the desired performance and competitiveness in the market.

The range of 0 to 100 percent for a cutting tool that affects product quality covers the spectrum of tool performance and characteristics, starting from the lowest to the highest level. At 0% there may be poorly maintained or inappropriate tool material, leading to poor processing and lower product quality. On the other hand, at 100% there is a high-quality cutting tool with optimal geometry and material, which enables precise and efficient material processing and results in high-quality products. This range allows the evaluation of the performance level of the cutting tool and its impact on the quality of the final product.

- **Operator** - The operator's influence on product quality is undeniably crucial. Their experience, skills and attention during the production process play a crucial role in ensuring a high standard of quality. Through precise machine setup, proper execution of operations and constant monitoring of processes, operators ensure that products are manufactured according to specifications and standards. Their ability to recognize and respond to potential problems or defects during production can have a significant impact on preventing unwanted defects and improving product quality. Therefore, operator training, motivation and support are vital to maintain a high level of quality and competitiveness in the market. The range from 0 to 100 percent for the operator reflects different levels of skill, experience, and efficiency in the production process, starting from the lowest to the highest level. At 0% there may be novice or poorly trained operators who may cause errors or defects in production. On the other hand, they can be 100% expert operators with rich experience and skills.

- **Digitization** - Digitization has a profound and comprehensive impact on product quality. Through the application of digital technologies such as automation, artificial intelligence and data analytics, manufacturers can optimize production processes, identify potential defects and improve quality controls at every step of production. This enables more precise control of production parameters, reducing quality variations and improving product consistency. Also, digitization enables proactive prediction of potential problems and faster resolution of possible irregularities, which significantly contributes to increasing product quality and user satisfaction.

Digitization range: 0% - Low digitization. Absence or minimal use of digital technologies in the production process. 100% - High digitization. Full implementation of digital technologies such as automation, artificial intelligence, and data analytics in the production process.

- Machine - The impact of machines on product quality is extremely significant. The efficiency and accuracy of the machines directly affect the production process, which ultimately results in the quality of the final products. Accurate calibration and regular machine maintenance are key factors in achieving production consistency and minimizing quality variations. Modernization of machines with advanced technologies enables automatic process monitoring and control, reducing the possibility of human error and increasing precision. Also, innovative technologies such as 3D printing machines enable the production of complex components with high accuracy and finish. Therefore, investing in high-quality machines and their continuous improvement are key steps towards achieving superior product quality.

The range of 0 to 100 percent for a machine ranges from poorly maintained and inefficient machines to high quality and reliable. At the lower end of the range, machines can be imprecise and cause machining errors, while at the higher end, they provide accurate and efficient material processing. This range enables the evaluation of the efficiency of the machines in production and the impact on the quality of the final products.

- Manufacturing time - The manufacturing time of a product plays an important role in determining its quality. The speed of the production process can lead to too frequent omissions, insufficient control, or lack of attention to detail, which can result in low product quality. On the other hand, products made too quickly may not have enough time for complete processing or drying, which may also affect the final quality. Therefore, manufacturers need to determine the optimal manufacturing time for each product, taking into account the required steps, processes and time required to achieve the desired quality. The balance between the efficiency of the production process and ensuring a high standard of quality is key to successfully meeting the expectations of users and maintaining competitiveness in the market. The range of 0 to 100 percent for manufacturing time indicates the spectrum of time frames required to complete the manufacturing process, ranging from minimum to maximum time. At the lower end of the range, there is rapid production that can lead to reduced attention to detail, while at the higher end, it allows for thorough machining and precision, ensuring high-quality final products. This range makes it possible to estimate the optimal production time to achieve the desired level of product quality.

Table 1: Input and output parameters

Input and output	Parameter description	
	<i>The name of the input parameters</i>	<i>Range</i>
Input 1	Starting material	0% - 100%
Input 2	Cutting tool	0% - 100%
Input 3	Operator	0% - 100%
Input 4	Digitization	0% - 100%
Input 5	The machine	0% - 100%
Input 6	Build time	0% - 100%
Output	Product Quality	0% - 100%

3. Results and Discussion

Table 2 shows the results of applying the Anfis method, including the mean error value, the mean deviation, the root mean square error, and the root mean square error for the training and test data. Table 3 also shows the reliability coefficient of the obtained model for the input factor that has the greatest influence on the output size.

Table 2: The impact of one input on the output

INPUT 1
TRAINING – ERROR
SV = -0.000000 SD = 0.175803 MSE= 0.030718
RMSE= 0.175266
TEST – ERROR
SV = -0.013962 SD = 0.091634 MSE= 0.008302
RMSE= 0.091117
ALL DATA – ERROR
SV = -0.0020983 SD = 0.1658 MSE= 0.02735
RMSE= 0.16538

Table 3: Reliability of the model

Reliability of the model
TRAINING DATA
R=0.99998
TEST DATA
R=1
ALL DATA
R=0.99998

Input 1, which corresponds to the starting material or casting, shows the lowest root mean square error (RMSE) value compared to the other inputs. This indicates that the starting material has the most significant influence on the quality of the product.

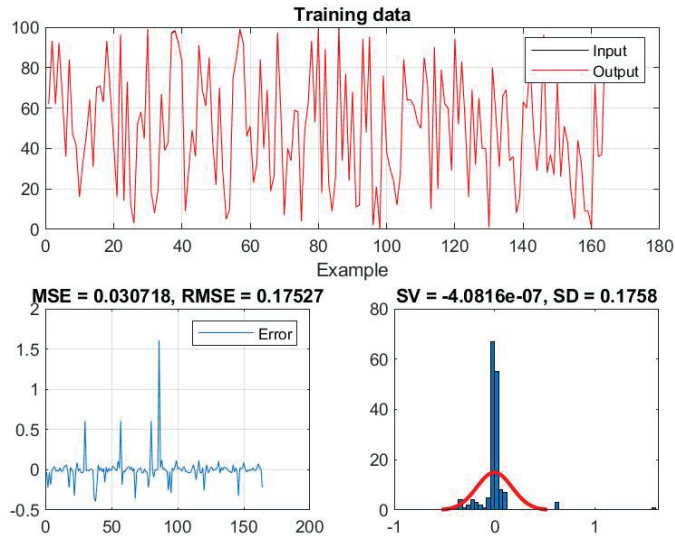


Figure 1. ANFIS network training - one input - achievements

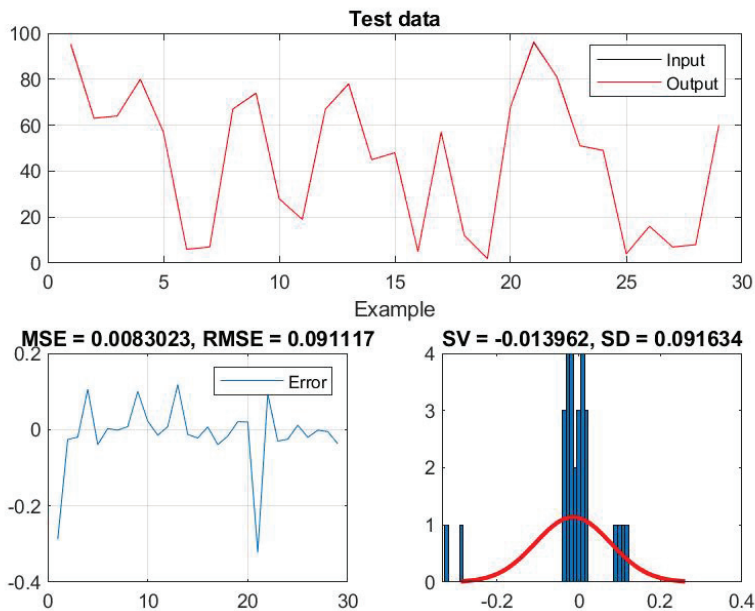


Figure 2. ANFIS network test - one input – achievements

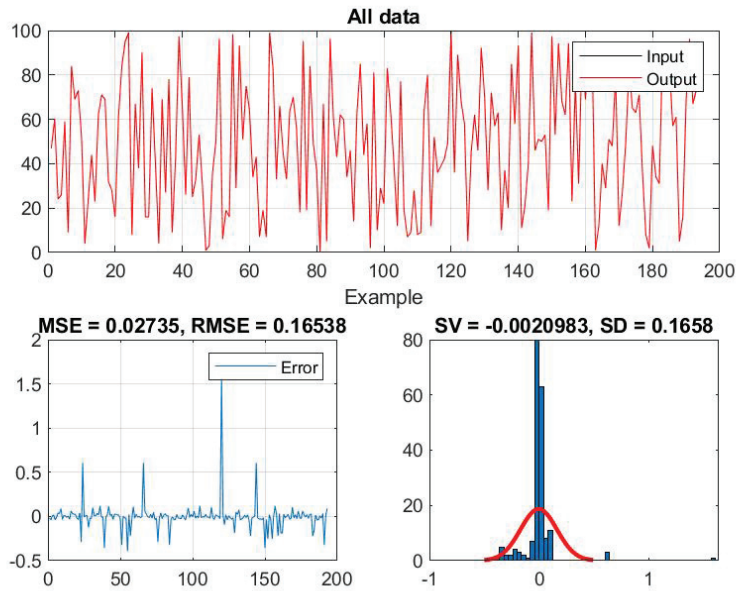


Figure 3. All ANFIS network data - one input

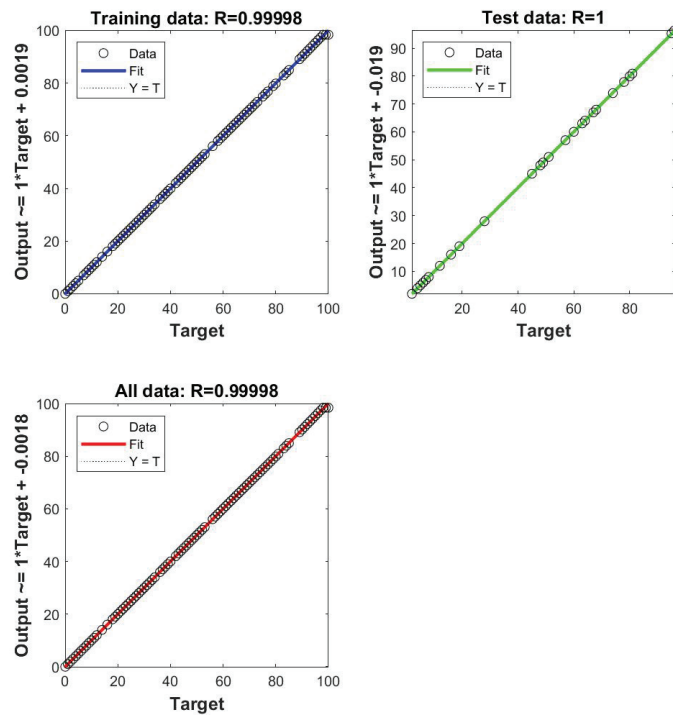


Figure 4. Regression of training, test and all data - one input - achievement

Figures 1, 2 and 3 show complete data on errors during training, test and all data obtained by applying the ANFIS methodology in the Matlab software package. Figure 4 shows the regression analysis and reliability of the model, while Figure 5 shows the graphical interpretation of the training data.

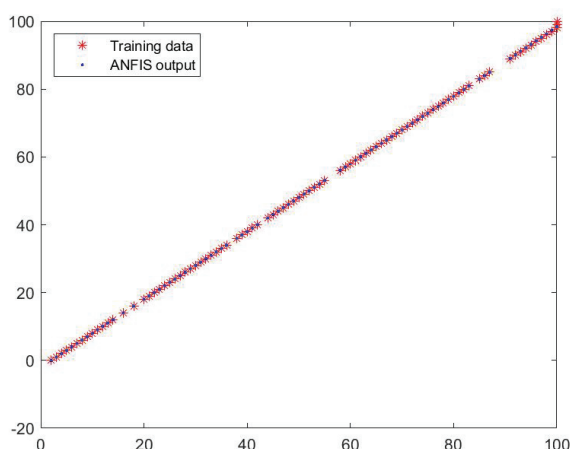


Figure 5. Graphical interpretation of training data - one input – achievements

In addition to the one input that affects the output the most, the two inputs that most affect the quality of the product were obtained, that is, the two inputs that in combination give the smallest RMSE error. Inputs 1 and 3, that is, the starting material and the operator, have the greatest influence on the quality of the product.

Table 4 shows all error values for training, testing and all data, taking into account the two most influential inputs - input 1 and input 3 - concerning the output size. Also, table 5 shows the reliability coefficient of the obtained model.

Table 4: The impact of two inputs on the output

INPUT 1 - 3

TRAINING – ERROR

SV = -0.000000 SD = 0.000082 MSE= 0.000000

RMSE= 0.000082

TEST – ERROR

SV = -0.000007 SD = 0.000070 MSE= 0.000000

RMSE= 0.000069

ALL DATA – ERROR

SV = -0.0000146 SD = 0.000080 MSE= 0.000000

RMSE= 0.00008

Table 5: Reliability of the model

Reliability of the model
TRAINING DATA
R=1
TEST DATA
R=1
ALL DATA
R=1

Figures 6, 7 and 8 show complete data on errors during training, test and all data for the influence of two combined factors. Figure 9 shows the regression analysis and reliability of the model, while Figure 10 shows the graphical interpretation of the training data.

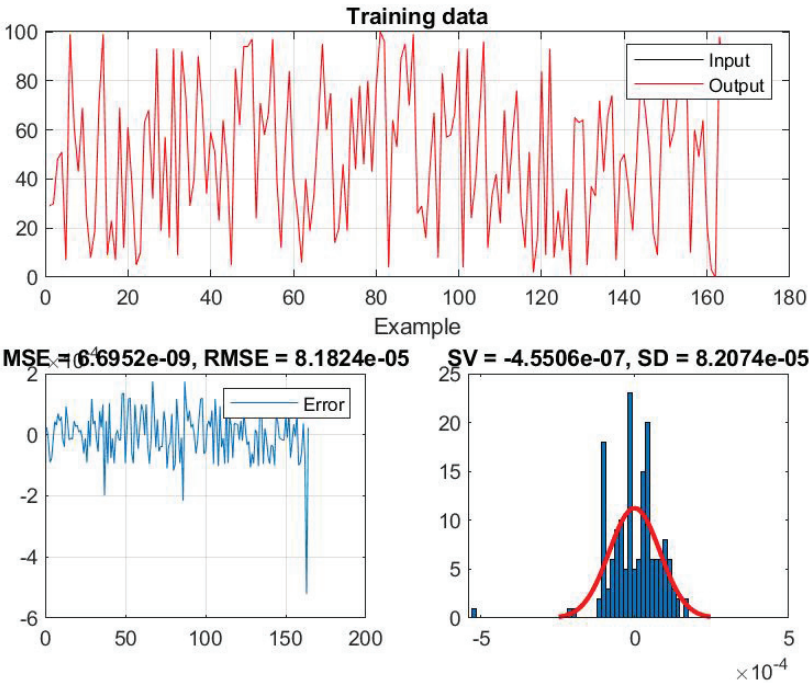


Figure 6. ANFIS network training - two inputs - achievements

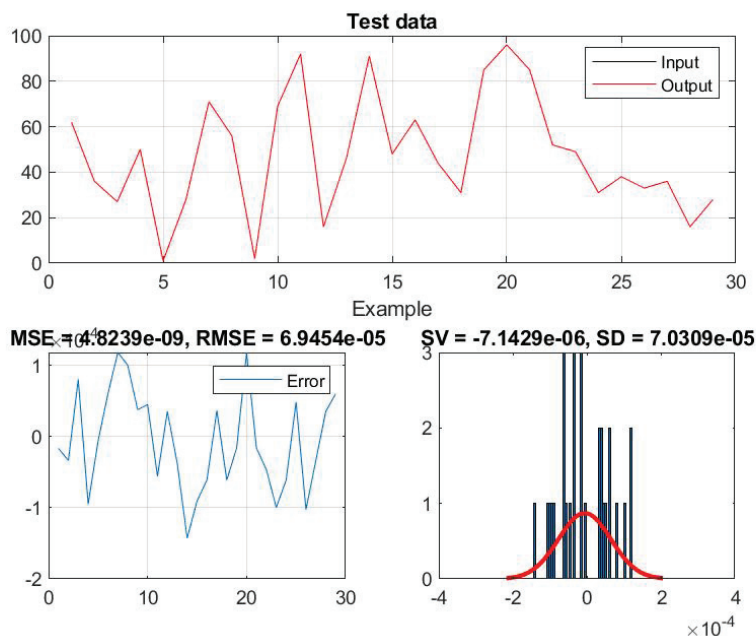


Figure 7. ANFIS network test - two inputs – outputs

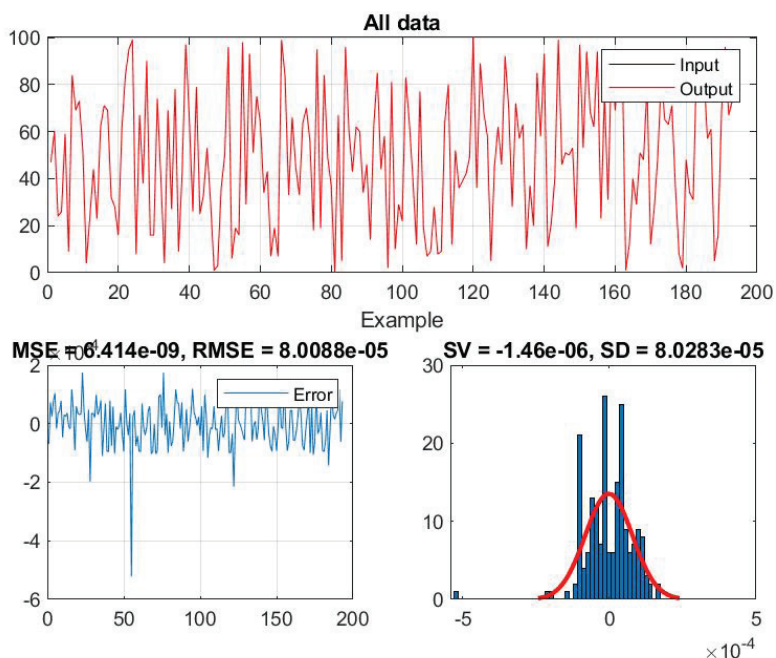


Figure 8. All ANFIS network data - one input

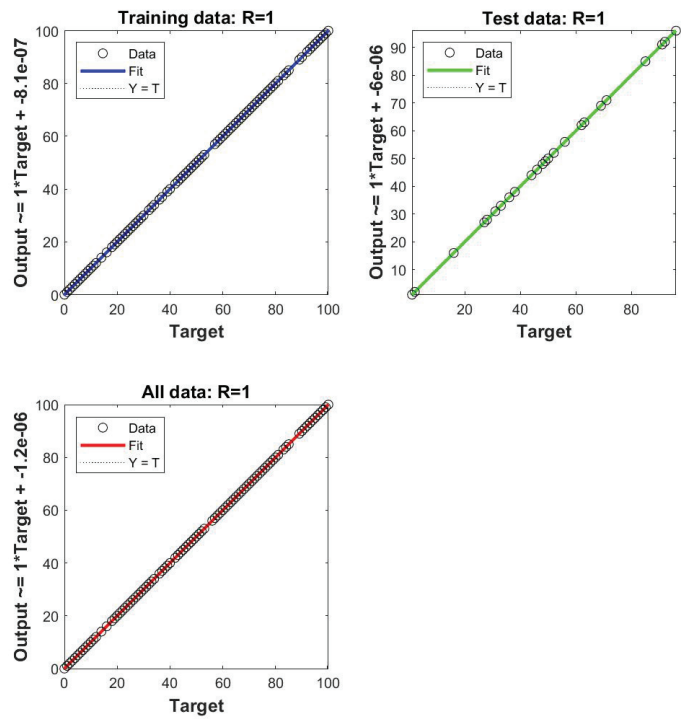


Figure 9. Regression of training, test and all data - two inputs - achievements

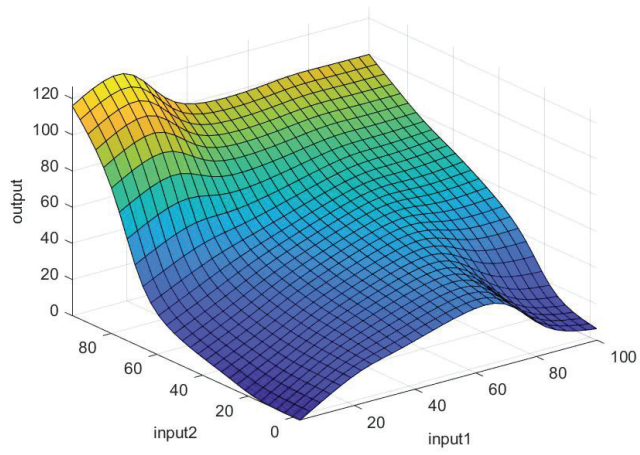


Figure 10. Graphical interpretation of training data - Matlab - two inputs – achievements

Table 6 shows the training, test, and all data for the three input factors that together have the greatest impact on product quality. In Table 7, the reliability of the model for the three combined factors is shown. The obtained research results highlight three key input factors - cutting tool, operator, and machine - as the main determinants of product quality.

The analysis showed that these factors are essential to achieve a high standard of production. The cutting tool is crucial because it directly affects the machining precision and product finish, while the skills and experience of the operator play a key role in the correct use of the tool and the optimization of the process. On the other hand, machines represent the basis of the production system, and their efficiency, precision, and reliability directly affect the quality of the final products. These three factors, when optimally coordinated and managed, provide the foundation for achieving high quality and competitiveness in the production process.

Table 6: The impact of three inputs on the output

INPUT 2 - 3 - 5
TRAINING – ERROR
SV = -0.000000 SD = 0.000029 MSE= 0.000000
RMSE= 0.000029
TEST – ERROR
SV = 0.000005 SD = 0.000024 MSE= 0.000000
RMSE= 0.000024
ALL DATA – ERROR
SV = -0.00000382 SD = 0.000283 MSE= 0.000000
RMSE= 0.0000282

Table 7: Reliability of the model

Reliability of the model
TRAINING DATA
R=1
TEST DATA
R=1
ALL DATA
R=1

Figures 11, 12, and 13 show complete data on errors during training, tests, and all data for the influence of three combined factors. Figure 14 shows the regression analysis and reliability of the model, while Figure 15 shows the graphical interpretation of the training data.

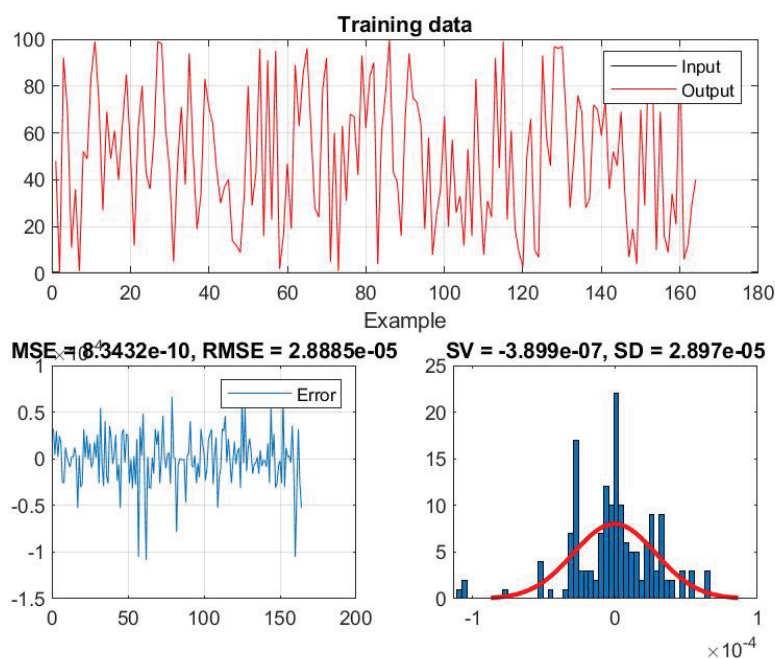


Figure 11. ANFIS network training - three inputs - achievements

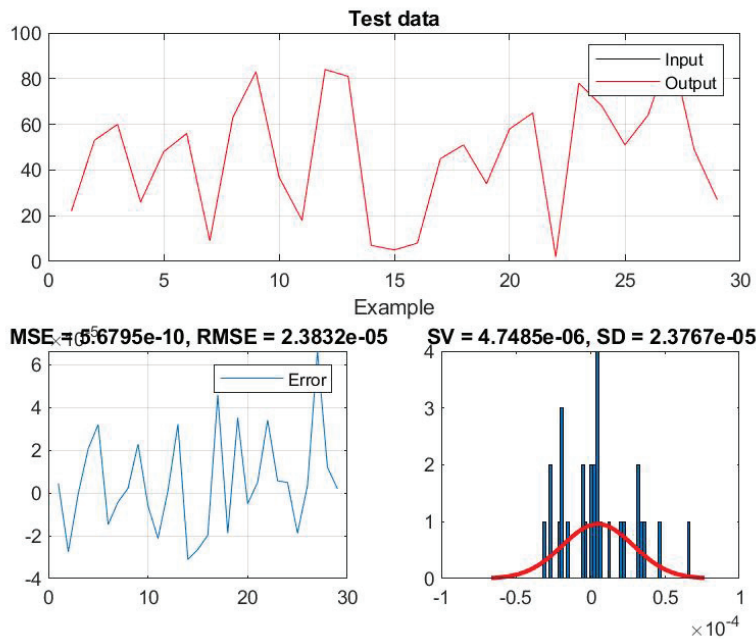


Figure 12. ANFIS network test - three inputs – achievements

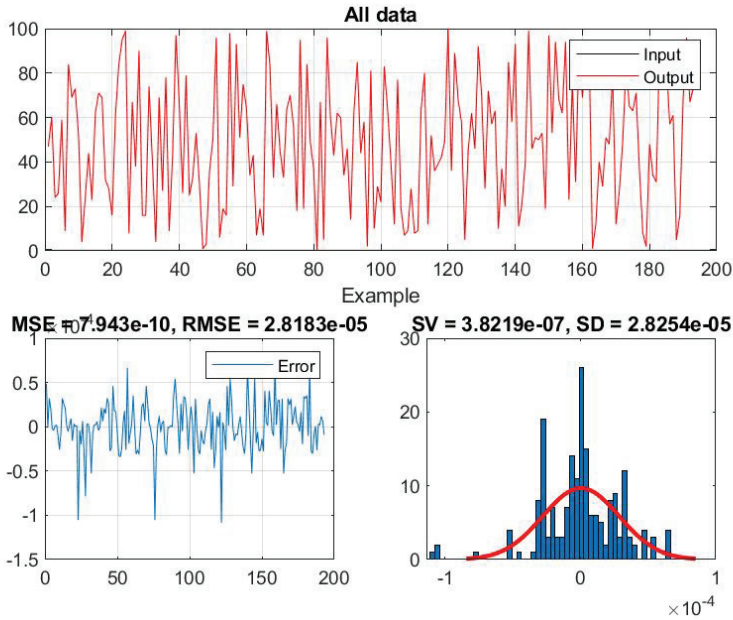


Figure 13. ANFIS network test - three inputs – achievements

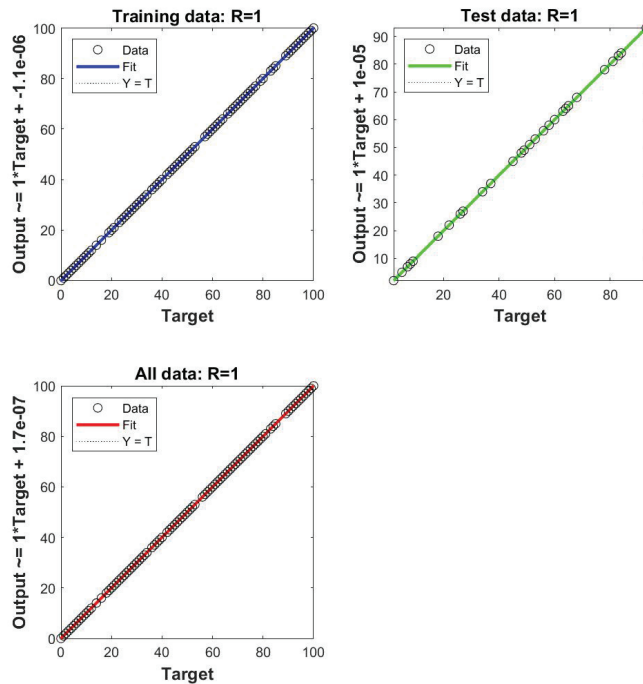


Figure 14. Regression of training, test, and all data - three inputs - achievements

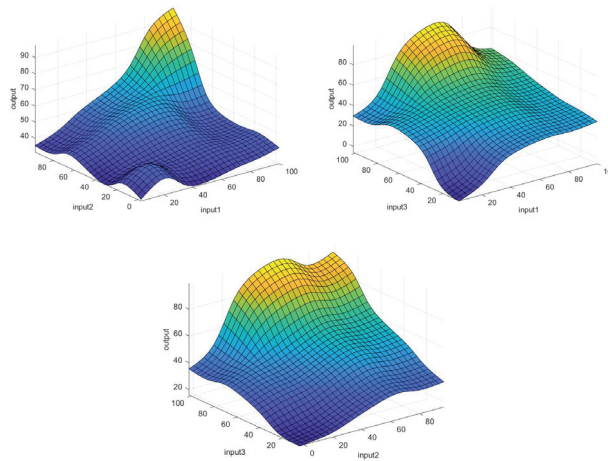


Figure 15. Graphical interpretation of training data - Matlab - three inputs – achievements

4. Conclusion

In the final consideration of the research of advanced technological methods in the optimization of product quality in the dedicated industry, the key influence of input factors on the quality of product manufacturing was highlighted. It has been shown that the starting material, the casting, has a key role in determining the quality of the product in the production process. Its chemical structure, mechanical characteristics, and surface treatment directly affect the performance and durability of the final product. In the dedicated industry, the quality and homogeneity of the metal alloys used as starting material can significantly affect the strength, corrosion resistance, and other key properties of the product.

Then, the two factors that together give the best result, i.e., the most impact on quality, were investigated, namely input 1 and input 3. In production, the combination of quality starting material and the skills of an experienced operator creates a top-quality product. That combination allows every element of the material to be used in the best possible way, providing the end product with exceptional value and impeccable quality.

After analyzing the previous factors, the focus was on the research of three basic factors whose joint influence is crucial for achieving high product quality. These factors are the cutting tool, operator skills, and machine characteristics. Establishing an optimal balance between these elements is an essential step toward improving the production process. The harmony and efficiency of this trio make it possible to achieve exceptional quality of the final product.

Potential improvements could include further refinement of techniques and technologies related to cutting tool processes to increase the accuracy and efficiency of material processing. Also, investing in operator training on the latest methods and tools could contribute to better utilization of the machines' potential. Also, the implementation of a system for monitoring and analyzing the performance of machines could enable the identification of potential problems or deficiencies in operation, which would enable faster problem-solving and optimization of the production process.

An AI-related advancement may include the application of advanced data analysis algorithms to monitor the performance of production systems. These algorithms would enable the automatic detection of patterns and trends in production data, enabling faster and more accurate decisions in process optimization. Also, the introduction of a self-learning artificial intelligence system would enable the automatic adjustment of production parameters to achieve optimal product quality with minimal human supervision.

5. References

- [1]Coreynen, W., Matthyssens, P., & Van Bockhaven, W. (2017). Boosting servitization through digitization: Pathways and dynamic resource configurations for manufacturers. *Industrial marketing management*, 60, 42-53.
- [2]Perera, Y. S., Ratnaweera, D. A. A. C., Dasanayaka, C. H., & Abeykoon, C. (2023). The role of artificial intelligence-driven soft sensors in advanced sustainable process industries: A critical review. *Engineering Applications of Artificial Intelligence*, 121, 105988.
- [3]Rajasa, E. Z., Manap, A., Ardana, P. D. H., Yusuf, M., & Harizahayu, H. (2023). Literature Review: Analysis Of Factors Influencing Purchasing Decisions, Product Quality And Competitive Pricing. *Jurnal Ekonomi*, 12(01), 451-455.
- [4]Ibrahim, D. (2016). An overview of soft computing. *Procedia Computer Science*, 102, 34-38.
- [5]Denai, M. A., Palis, F., & Zeghib, A. (2004, October). ANFIS based modelling and control of non-linear systems: a tutorial. In 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583) (Vol. 4, pp. 3433-3438). IEEE.
- [6]Karaboga, D., & Kaya, E. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*, 52, 2263-2293.

- [7] Anđelković, B., Istraživanje i razvoj novih metoda za proračun steznih sklopova primenom neuronskih mreža i fazi logike, doktorska disertacije, Mašinski fakultet Niš, 2006.
- [8] Sugeno, M.; Kang, G.T.: Structure identification of fuzzy model, *Fuzzy Sets Syst.* 28 (1988) 15–33
- [9] Mojsilović, M., Cvejić, R., Pepić, S., Karabašević, D., Saračević, M., & Stanujkić, D. (2023). Statistical evaluation of the achievements of professional students by combination of the random forest algorithm and the ANFIS method. *Heliyon*, 9(11).

Machine Learning Within Industry 4.0: From Decision Trees to Visual Transformer Architecture

Zoran Miljković^{*1}, Aleksandar Jokić¹, Milica Petrović¹

Abstract: Machine learning models have been revolutionizing the manufacturing industry for a decade now. Many powerful models have been developed to solve previously thought not solvable problems. In the era of rapid development of the machine learning field, the selection of optimal models for the considered problem within the Industry 4.0 setting is still a hard question. In the first part of this work, we provide an overview of the use cases of machine learning for Industry 4.0. Afterwards, we provide details of the most commonly utilized machine learning models. Finally, the experimental evaluation is performed to compare the analyzed models for two different use cases, namely visual inspection and predictive maintenance. Two machine learning models are compared for each task in order to highlight the main differences and experimental results.

Keywords: Machine learning, Industry 4.0, visual transformer, artificial neural networks, convolutional neural networks, decision trees.

1. Introduction

Artificial Intelligence (AI) is a crucial driver in the fourth industrial revolution, offering substantial advantages to the industrial sector through improved operational efficiency, cost reduction, and data-driven decision-making. Advanced machine learning algorithms based on AI excel in data analysis and process optimization[1], outperforming human-created algorithms in enhancing manufacturing productivity and efficiency[2]. These algorithms enable machines to operate continuously and undertake tasks that are either too laborious or dull for human workers. Moreover, AI-based algorithms can perform complex tasks that traditional algorithms struggle with, such as recognizing patterns in large datasets, making predictions from unstructured data, and adapting to new information without explicit programming. Furthermore, by providing comprehensive insights and predictive analytics, AI algorithms aid in more informed decision-making at all levels of an organization[3]. They can be used to identify trends, forecast outcomes, and provide recommendations, enabling management to make strategic decisions based on data. Another significant benefit of the implementation of AI regards the integration of various systems

^{*1}University of Belgrade – Faculty of Mechanical Engineering Production engineering department
E-mail: zmiljkovic@mas.bg.ac.rs, ajokic@mas.bg.ac.rs, mmpetrovic@mas.bg.ac.rs

and processes within the industrial setup, ensuring seamless communication and interaction across different platforms and machines, leading to more cooperative and coordinated operations. Finally, these integrated processes respond more quickly to market changes, and optimize their operations in ways that ensures competitive advantage, sustainability, and waste reduction[4].

Particularly, one of the production subsystems that can benefit extensively from AI is predictive maintenance. Machine learning-enabled predictive maintenance allows companies to anticipate equipment failures before they occur. This predictive capability helps reduce downtime and maintenance costs, as maintenance can be scheduled before a machine breaks down, preventing disruptive and expensive emergency repairs.

Vibration, acoustic, electrical, and temperature data from machinery can be used to detect anomalies or patterns that indicate impending failures[5]. Machine learning models such as Support Vector Machines [6] or Artificial Neural Networks [7] can classify vibration signals into regular or potential failure states, allowing for timely maintenance actions before a breakdown occurs. Machine learning algorithms, such as Random Forests [8] or Gradient Boosting [9], can integrate various data points (temperature, vibration, and operational hours) to predict when a piece of equipment might fail, which enables proactive maintenance scheduling. Learning algorithms can analyze sounds from equipment to detect changes that may indicate mechanical issues.

Deep learning models, particularly Convolutional Neural Networks (CNNs), are adept at processing audio data and identifying subtle changes in acoustic patterns [10] that might signify potential failures. Transformer architecture based on large language models can be utilized to use maintenance logs made by operators to predict future failures[11]. Data-driven digital twins[12] (virtual replicas of physical systems) can simulate and analyze the behavior of machinery under different conditions. By utilizing AI algorithms within these simulations, companies can predict failures and understand potential maintenance needs without interrupting actual operations. Deep learning models, particularly Long Short-Term Memory networks[13] are effective in analyzing time-series data from sensors to detect patterns that precede equipment failures. These models can learn from historical data and make accurate predictions about the future machinery state. Having that in mind, in this chapter, the analysis of decision trees and artificial neural networks for machine failure prediction will be made, based on the predictive maintenance dataset[14].

Moreover, AI-based models can be used to improve quality control processes in manufacturing. By analyzing images or sensor data, these models can identify product defects or anomalies, ensuring that only items meeting the highest quality standards reach customers. Particularly, traditional inspection processes are being turned into more efficient, accurate, and automated

operations. CNNs are particularly effective for image processing tasks, enabling the identification of various defect types, such as cracks, misalignments, or incorrect dimensions, with high accuracy. Deep learning models can be trained to recognize and classify different types of surface flaws on materials or parts. For example, a neural network can differentiate between acceptable variations and defects on a product surface. Unsupervised learning algorithms, e.g., Autoencoders[15], are used to learn the normal patterns of data and detect deviations that signify defects or quality deviations. Finally, Large Language Models (LLMs) can analyze customer feedback, reviews, or reports to identify common quality issues or trends[16]. This information can be used to improve customer feedback and, therefore product design and manufacturing processes. In some inspection applications, CNN models are integrated with robotic systems to automate the inspection process. Robots equipped with real-time vision systems [17]can conduct consistent and detailed inspections and relieve human workers from dull tasks[18]or environmentally hazardous surroundings. Therefore, in this chapter, the utilization of CNNs for the visual inspection of casting parts will be performed and analyzed.

Another aspect of Industry 4.0 is the integration of machine learning algorithms and robotics, which transforms manufacturing environments and creates smarter, more flexible, and more highly efficient processes[19]. This integration allows robots to perform tasks with greater autonomy, adaptability, and precision. Autonomous mobile robots enhanced with AI-based control algorithms[20] navigate complex industrial environments independently, avoiding obstacles, optimizing routes[21], and transporting materials[22]. They use sensor data processed by machine learning algorithms to interpret their surroundings and make real-time decisions. Moreover, AI-driven vision systems enable robots to interpret visual data [23]and carry out complex tasks like quality control, inspection, and object recognition. These systems use advanced machine learning-based image processing algorithms, such as CNNs, to analyze images or video streams in realtime.

Furthermore, LLMs can be used to enhance the interaction between humans and robots, enabling more intuitive and effective collaboration. Natural language processing (NLP) and machine learning based on LLMs allow robots to understand and respond to human commands, while predictive algorithms help them anticipate human actions and adjust robot behavior accordingly. Collaborative robots can work alongside humans without the need for physical barriers. All these capabilities are enabled by machine learning models, all while ensuring safety and efficiency. Machine learning models help collaborative robots to improve their performance over time, learning from each human interaction.

In summary, artificial intelligence, especially machine learning, represents one of the cornerstones of Industry 4.0, driving significant

transformations across industries by enabling smarter, more efficient, and more adaptive manufacturing systems.

2. Machine Learning Models

A machine learning model is a computational framework of a system or process whose parameters are learned by utilizing data and a specific training algorithm. After the training, the model can identify patterns or relationships within the data, allowing it to make a prediction without being explicitly programmed for a specific task. More formally, the machine learning model (1) is a function f with the set of tunable parameters θ that can map the input data X to the output prediction Y . The input/output pairs of data define the dataset (D). The training process consists of the optimization of parameters θ based on the defined loss function.

$$Y = f(X; \theta) \quad (1)$$

A well-trained machine learning model needs to generalize well to new and unseen data that are usually part of the validation and test datasets. Generalization is achieved through techniques such as cross-validation and regularization. The following sections will provide the definition of machine learning models such as decision trees, artificial neural networks, and convolution neural networks.

2.1. Decision trees

Decision trees are one of the most commonly utilized machine learning algorithms for both classification and regression tasks (Fig. 1). They work by incrementally developing a decision tree by breaking down a dataset (feature space) into smaller subsets based on conditions regarding the input features. The final result is a tree with root node, decision nodes, and leaf nodes that can be used to make predictions based on the features of a new data point.



Figure 1. Abstract illustration of a decision tree

The root node represents the entire dataset (D), which is then split into two or more sets based on the most dominant threshold of feature. The goal is to find the combination of features and thresholds (θ_t) that best separates the dataset at node t (2) and (3).

$$D_t^{left} = \{(X, Y) \in D \mid X_i \leq \theta_t\}, \quad (2)$$

$$D_t^{right} = \{(X, Y) \in D \mid X_i > \theta_t\}, \quad (3)$$

To determine the best split, metrics such as Gini impurity (4) or entropy (5) are used:

$$G(t) = 1 - \sum_{k=1}^K p_k^2, \quad (4)$$

$$H(t) = - \sum_{k=1}^K p_k \log p_k, \quad (5)$$

where K is the number of classes, and p_k represents the probability that the sample belongs to the k -th class. Entropy measures the randomness or uncertainty in the dataset. Gini impurity measures the probability of a randomly chosen element being incorrectly classified according to the distribution of labels in the subset. Different algorithms are used to determine how to make a split and, therefore, create a decision tree Classification and Regression Trees[24],

Iterative Dichotomiser 3[25], and C4.5[26]. These algorithms partition the data into subsets that have the highest purity or lowest impurity.

$$\theta_t^* = \arg \min_{X_t, \theta_t} \left(\frac{n_t^{left}}{n_t} \cdot G(D_t^{left}) + \frac{n_t^{right}}{n_t} \cdot G(D_t^{right}) \right) \quad (6)$$

where θ_t^* represents the optimal threshold value for the optimal feature X_t at node t , and n represents the number of samples. Decision nodes are utilized to split each subset of the feature space further. These nodes can be split further into additional decision nodes or become a final (leaf) node. Leaf nodes represent the outputs of the decision tree where no further splitting is possible. In classification, it represents the class of the data points in this node. In regression, it typically represents the mean of the target values. Once the tree is built, it can be used to make predictions. For a new data point, the decision tree is used by testing the relevant conditions (decision and root nodes) until a leaf node is reached.

2.2. Artificial Neural Networks

Artificial Neural Networks (ANNs) represent machine learning models that are inspired by human neural networks. They are utilized for approximation of complex functions, classification, and time series analysis[27]. The structure of the ANN consists of layers of neurons that are all interconnected with tunable weights. The ANN has a minimum of three layers, the input layer, where the input vectors are passed; the hidden layers, which represent the hyperparameter of the model; and the output layer used to generate output or result of a network. The number of neurons in the input and output layer is defined by the number of features in input and output vectors, while the number of neurons in each hidden layer is also a hyperparameter. The learning process of an ANN is performed by adjusting the weight parameters until the difference (defined by a loss function) between desired and predicted output vectors is minimal.

The input vector is defined as $\mathbf{X} = (x_1, x_2, \dots, x_n)$, where n represents the number of features. The output of all neurons in the hidden and output layers are defined by (7)

$$h^l = f(W^l h^{l-1} + b^l) \quad (7)$$

where h represents the output of each neuron in layer l , f is the activation function, W is a matrix of weights between $l-1$ and l layer, and b represents the bias value. The type of activation function used is another hyperparameter of the ANN; some of the most utilized are:

- Sigmoid activation function $f_s(x) = (1 + e^{-x})^{-1}$
- Hyperbolic tangent activation function $f_t(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
- Rectified Linear unit (ReLU) activation function $f_r(x) = \max(0, x)$
- Gaussian error Linear Unit (GeLU) activation function

$$f_g(x) = \frac{x}{2} \left(1 + \tanh \left(\sqrt{\frac{2}{\pi}} (x + 0.445x^3) \right) \right)$$

The formal definition of the general training algorithm is given in equation (8):

$$W^{new} = W^{old} + \eta \frac{\partial L}{\partial W}, \quad (8)$$

where η represents the learning rate, and L is the loss function.

2.3. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) represent a group of artificial neural network architectures that utilize convolution as a primary processing operation. Unlike the standard ANNs defined in Section 2.2, CNNs do not have all the neurons connected in subsequent layers, and the architecture is generated by utilizing different layers. Another distinguishing factor is that CNNs are typically utilized in the Deep Learning (DL) domain, while ANNs are used in traditional Machine Learning (ML). Even though DL is a subdomain of ML, the main difference is that DL is used when there are large and complex datasets, while for ML, some sort of feature engineering is required first. Moreover, DL models are characterized by utilizing an enormous number of parameters and many subsequent layers.

One of the most commonly used layers in CNNs is the Convolutional Layer, where the main parameters are the filter size, stride, padding, and number of filters. After the convolutional layer, activation layers are usually implemented with a ReLU activation function. To minimize the resolution of the input feature maps and, therefore, parameters in the network, the average or max pooling layers are utilized. Regularization is achieved by using dropout or batch norm layers. A detailed explanation of the individual layers used in CNNs can be found in [28]. Numerous CNNs have been proposed throughout the years for different problems, mainly in the computer vision domain. Some of the most commonly utilized architectures include alexnet[29], VGG[30], resnet[31], mobilenet[32], and densenet[33].

2.4. Visual Transformer Architecture

A Visual Transformer (ViT)[34] is a type of artificial neural network architecture developed for computer vision tasks, leveraging the power of transformer architectures that have shown remarkable success in natural language processing. The entire ViT architecture is given in Fig. 2.

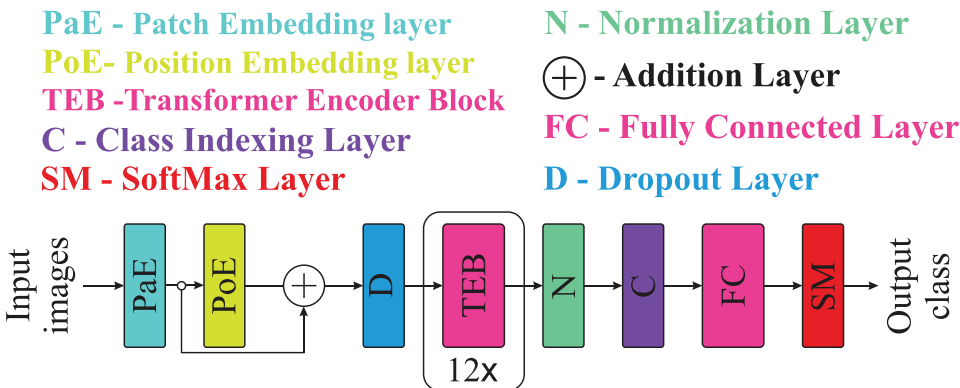


Figure 2. Visual Transformer (ViT) architecture.

As can be seen, the input images are processed through a sequence of layers, with the main block of layers defined within the Transformer Encoder Block (TEB), which is repeated 12 times. The output of ViT architecture is a vector that represents the class of the input image.

The ViT utilizes the process of dividing input images into non-overlapping patches of fixed size. Each patch (shown in green – Fig. 3) is reshaped into a vector, and a fully connected layer is applied to each vector to map it to an embedding dimension by using the Patch Embedding layer (PaE). This process is shown in Fig. 3. Each patch embedding layer output value is represented with a green rectangular cuboid. The concatenation layer is utilized to add an additional vector of weights at the start of the output (shown with a black rectangular cuboid), which is later used to make a prediction of the class output.

Afterwards, the whole output (all green and black vectors) is multiplied with learnable weights from the Position Embedding layer (PoE) and added with the values before multiplication by using skip connection and addition layer (Fig. 2.).

Unlike Recurrent Neural Networks (RNNs) or CNNs, which have an inherent notion of sequence and locality, transformers process the entire image simultaneously.

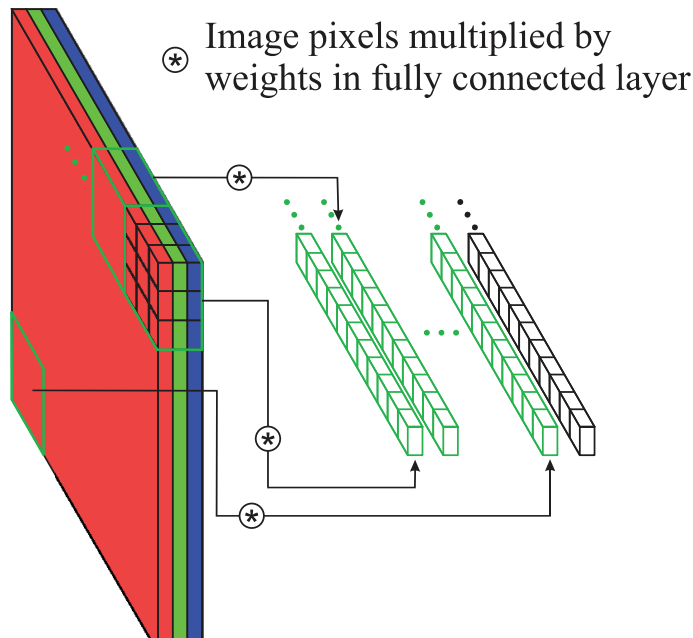


Figure 3. Patch Embedding layer (PaE) layer of ViT architecture

Therefore, there is a lack of a built-in mechanism to understand the sequential order of the input image patches, and the position embedding layer plays a crucial role in enabling the model to understand the order of image patches (and pixels) in the input sequence. The rest of the ViT architecture consists of several blocks of transformer encoder structure, which is shown in Fig. 4.

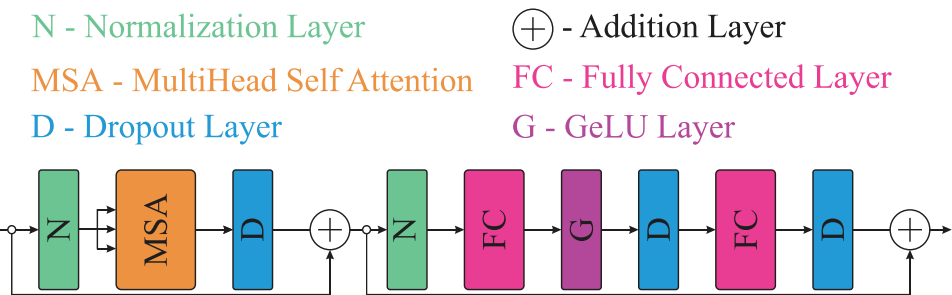


Figure 4. Transformer Encoder Block (TEB) structure

The Transformer Encoder Block (TEB) consists of the two parts separated by different skip connections and addition layer. The first part includes the Normalization layer that performs channel-wise normalization, the MutliHead Self Attention block, which will be covered in detail later, and dropout layers that are used for regularization purposes. In the second part, after the

normalization layer, the fully connected layer (explained in Section 2.2) is activated with the GeLU activation function. Afterwards, the dropout layer, fully connected layer, and another dropout layer are connected, and their final output is added to the input of the second part of TEB. Finally, the core of the transformer architecture is a MultiHead Self Attention (MSA) block (Fig. 5). Within MSA, the input values are copied three times, and they represent Query, Key, and Value matrices. These matrices are connected to three different fully connected layers, and the output of the key and query is multiplied with the matrix multiplication layer.

Afterwards, the output matrix is divided with the square root of the dimensions of the matrix within the scale layer, and the probabilities are generated with SoftMax. The output of those layers represents the attention filter that maps the importance between image patches. When the value output is multiplied with the attention filter, the result represents the filtered value matrix (attention map) that focuses on individual objects within the image. Twelve of these Self-attention heads (all with different weight values, focusing on other objects/parts of the image) are concatenated within the concatenation layer and connected with another fully connected layer to provide the final output of MSA.

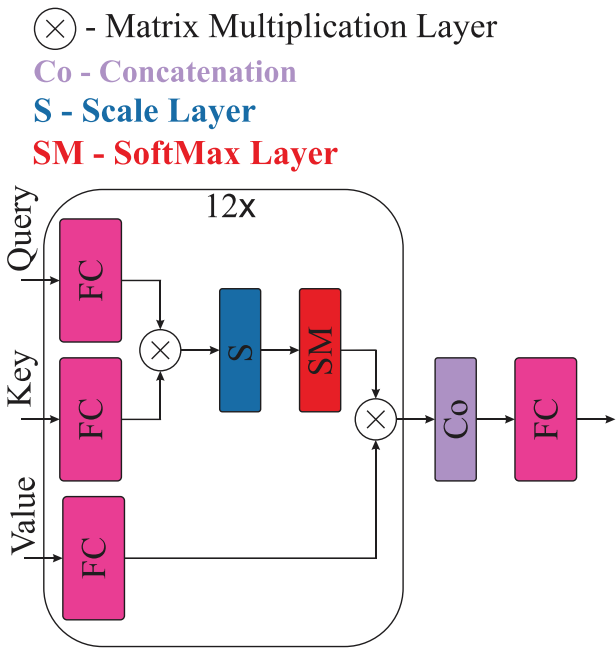


Figure 5. MultiHead Self Attention (MSA) block structure

3. Machine Learning Models within Industry 4.0 Settings: Experimental Results

The experimental evaluation is performed on the publicly available datasets [35][14]. The first dataset incorporates synthetic data regarding the predictive maintenance of milling machine based on seven features. The dataset contains 10000 data points; a few data samples are shown in Table 1. *Type* represents the product quality requirements (H – high, M – medium, L – low), while *Target* represents the output classes, 1 – failure (machine), 0 – non-failure. Even though the dataset has seven features, only six have been used for machine learning since product ID is just a randomly assigned number to each part and provides no additional information regarding the machine itself. The machine learning models are used to predict when the analyzed milling machine will fail in order to plan ahead for the maintenance process. The distributions of individual features within the dataset are given in Fig. 6.

Table 1. Samples of the predictive maintenance dataset

No.	Product ID	Type	Air temperature [K]	Process temperature [K]	Rotational speed [RPM]	Torque [Nm]	Tool wear [min]	Target
1	29462	H	298.8	309.2	1425	53.9	135	0
2	14909	M	298.9	309.2	1412	44.1	140	0
3	47230	L	298.9	309.1	2861	4.6	143	1
□								
998	57153	L	298.5	308.2	1444	40.5	170	0
999	57154	L	298.6	308.2	1361	68.2	172	1
1000	24835	M	298.6	308.3	1589	34.1	174	0

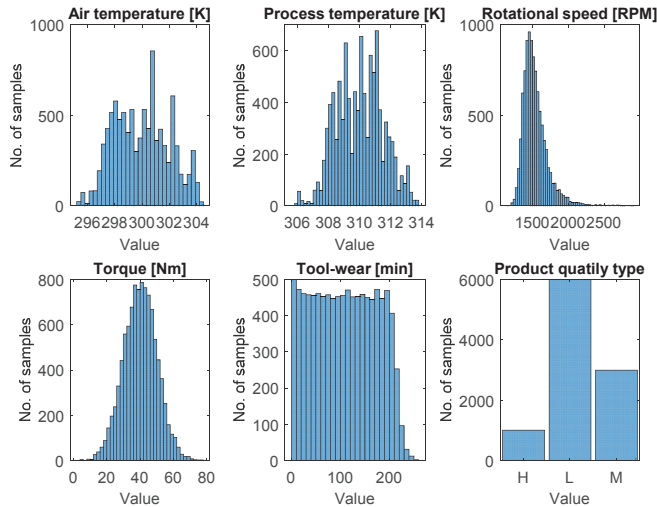


Figure 6. Distributions of features from the dataset.

As it can be seen, there is a wide variety in feature space with normal, uniform, and skewed distributions. After the data analysis is performed, machine learning can begin. As the input vector is relatively small (having six values), the utilization of CNNs or transformer architecture would probably result in overfitting.

Therefore, the problem of predictive maintenance is performed by utilizing artificial neural networks (fully connected networks) and decision trees. All the algorithms have been implemented using the Deep learning toolbox within MATLAB R2023b. The architecture of the utilized ANN is shown in Fig. 7. As it can be seen, the architecture of ANN is 6 [50-30-10]₃ 1. We utilized three hidden layers with 50, 30, and 10 neurons receptivity, whereas the first hidden layer has ReLU activation function, and two other hidden layers use tanh. Since the problem is binary classification, the output layer utilizes a standard sigmoid activation function.

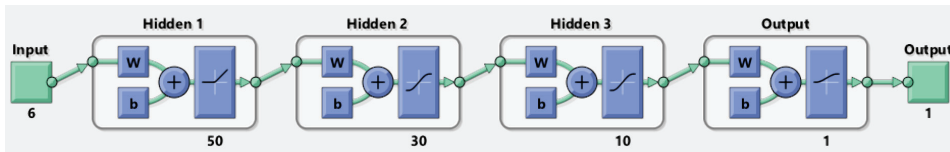


Figure 7. Architecture of the ANN

The training process of the ANN is shown in Fig. 8. The utilized training algorithm is Scaled conjugate gradient descent [36] with cross-entropy loss function. The dataset is randomly divided into training (80%), validation (10%), and test (10%) subsets.

The other machine learning model utilized for the problem of predictive maintenance is decision trees, also implemented within MATLAB Deep learning toolbox. The data fed into the model is the same as in the case of ANN. The utilized splitting criterion is Gini impurity; the maximum number of splits is set to 100 with no surrogates. The K-fold cross-validation is performed to ensure the generalization properties of the model. After the training is completed, the output tree with classes for each leaf is given in Fig. 9. The leaf nodes are represented with blue dots, while the decision nodes are shown with blue triangles.

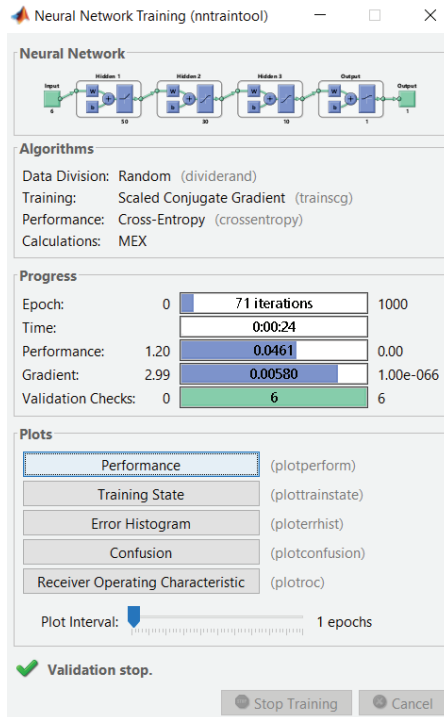


Figure 8. Training process of the ANN

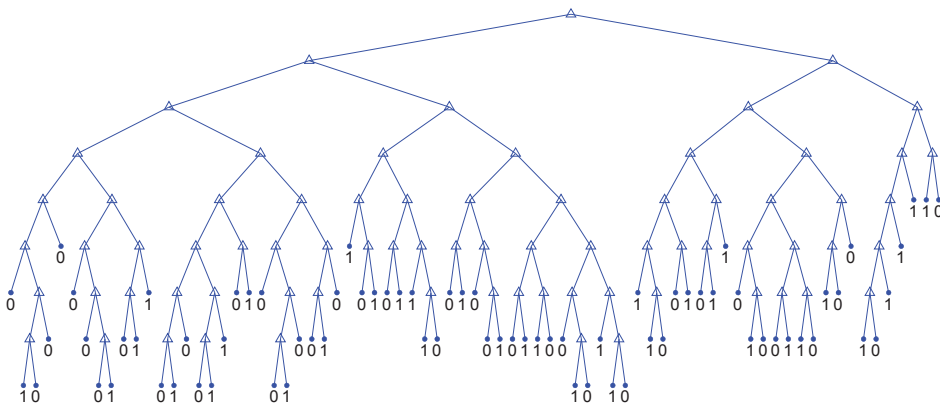


Figure 9. Decision tree

Moreover, as the entire decision tree would be hard to visualize, we present a part of the decision tree with decision values, decision node conditions and values in Fig. 10. As can be seen, the primary decision is based on the value of torque with a splitting value of 65.

The results are compared after the ANN and decision tree training are completed. The results are shown with two confusion matrices in Fig. 11.

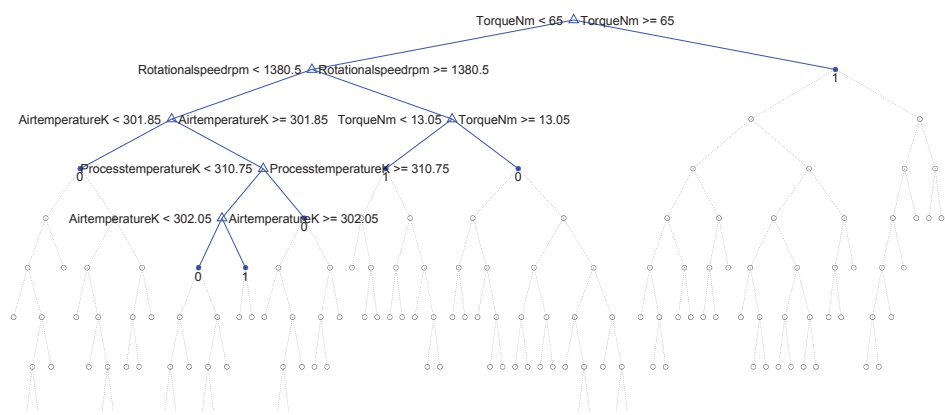


Figure 10. Decision tree with decision node conditions.

When the main parts of the confusion matrices are compared, the results seem similar. The results of the decision tree algorithm do not add to 100% due to the rounding of numbers to the first decimal point. Both ML models achieve 98.2% accuracy in detecting that non-failure will occur. Decision trees predict 0.01% more occurrences of failure compared to ANN. However, the decision trees also predict more non-failures when there is one, which means ANNs do better on this metric. On the other hand, ANNs also predict 0.01% more failure when there is none, which is a better occurrence compared to the last one. The experimental evaluation suggests that decision trees for the considered problem achieve better overall results. Another benefit of using a decision tree compared to ANN is the interpretability of the model. By using decision trees, it is always clear why some input generated predicted class based on the tree and the value of the decision variables. However, for ANN, it is not so easy to interpret why some weights in the network have considered values, and it is much harder to interpret why some input generated its paired output. Therefore, in the real-world Industry 4.0 setting, it is significantly easier to implement decision trees due to their ease of interpretation for humans.

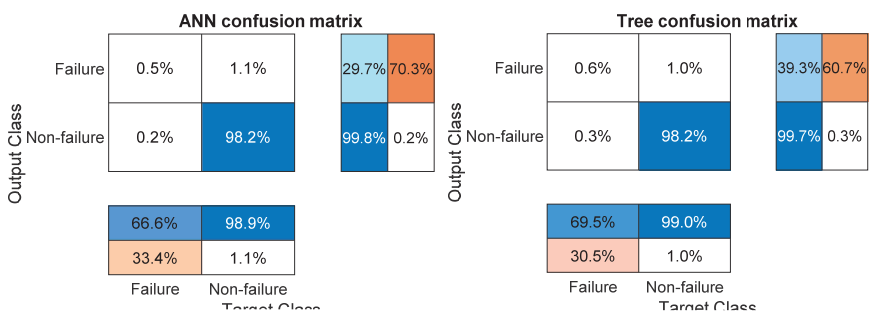


Figure 11. Confusion matrices of the analyzed models

The second experimental evaluation is performed on the visual inspection task. The considered dataset contains images for visual inspection of casting parts. Images are classified into two categories ok parts and parts with defects. The considered images are grayscale with a resolution of 512×512 . The dataset is divided into training and testing subsets (130 images), whereas we further divide the training dataset into validation (130 images) and training (1040 images). Three data augmentation techniques are considered: rotating the image ± 90 degrees, random scaling of the images, and mirroring the image along the Y axis. The samples of the correctly cast parts are shown in the first row of Fig. 12, while parts with defects are given in the second row.

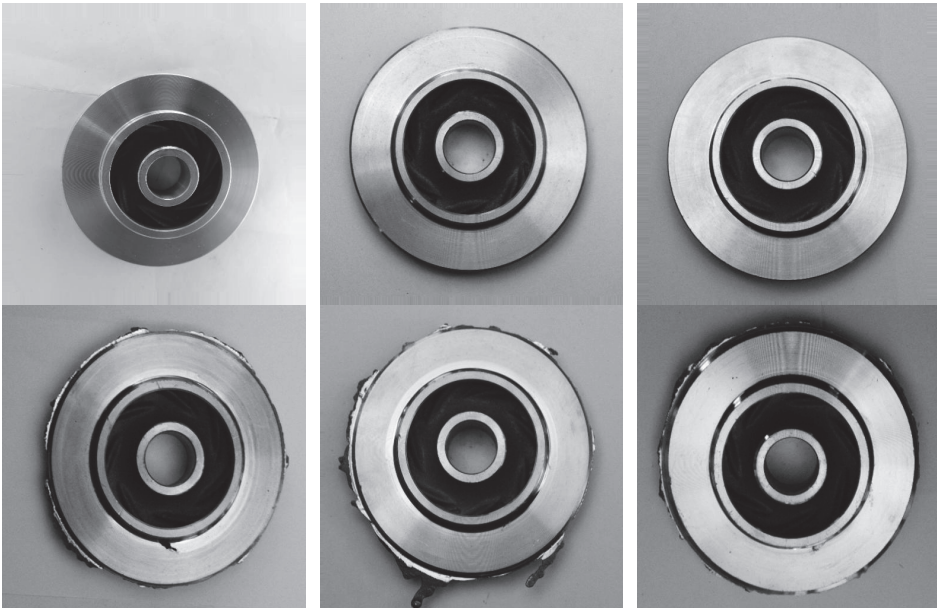


Figure 12. Sample images from the visual quality inspection dataset

A deep learning toolbox within MATLAB R2023b was utilized for the training of deep neural networks. For the considered problem, two deep neural networks are utilized: Visual Transformer (ViT) [34] and a densely connected convolutional network [33]. Both networks are pretrained on imagenet dataset and fine-tuned for the considered problem. The utilized dense network has an overall of 707 layers, with more than 800 connections between them. On the other hand, the utilized ViT has 143 layers with 167 connections. Therefore, since the size of networks differs significantly, to make a fair comparison, the ViT was trained for eight epochs, while the dense network was trained for four. Other training parameters were the same: mini-batch size of 12, Adam optimizer[37], and initial learning rate of 0.0001. The training process for dense network and ViT can be seen in Figs. 13 and 14, respectively.

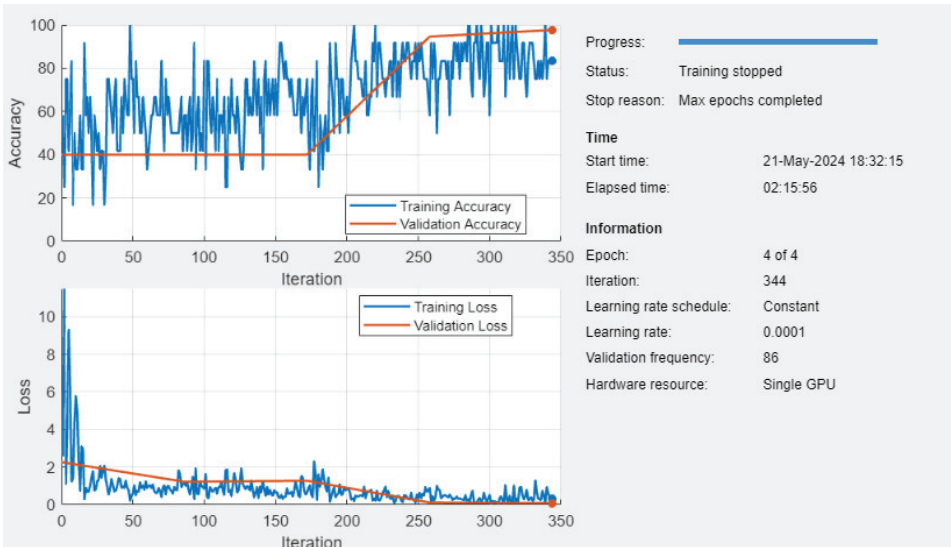


Figure 13. Learning process of the dense net

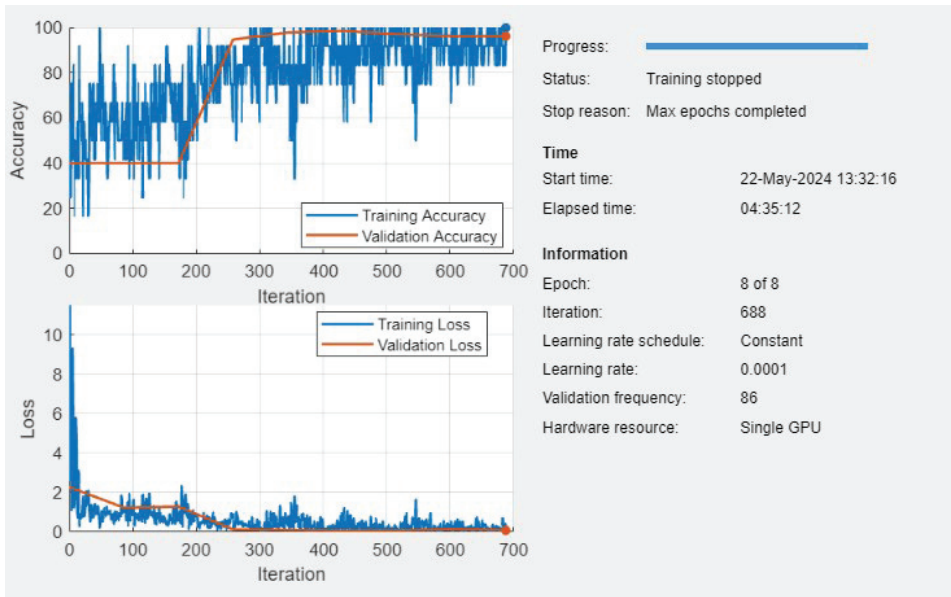


Figure 14. Learning process of ViT

As it can be seen, the loss function on both training and validation datasets gradually decreases during the training process for both networks. Also, training and validation accuracy increases and oscillates around 90%, indicating that the training process is progressing well. The final results of both networks are shown with test set confusion matrices in Fig. 15.

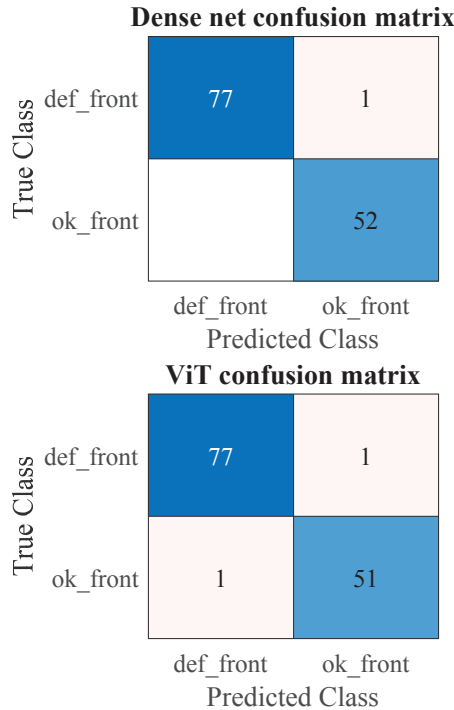


Figure 15. Confusion matrices of the dense net and ViTmodels

As shown in Fig. 15, both machine learning neural networks perform exceptionally well, achieving around 99% accuracy on the test set. The dense network performed slightly better by correctly classifying one sample more than the ViT transformer.

4. Conclusion

In this chapter, the authors analyzed the usage of different machine learning models within the Industry 4.0 domain. The main focus of the experimental evaluation was on the tasks of predictive maintenance and visual inspection. For predictive maintenance, machine learning models, namely, fully connected artificial neural networks, are compared to the decision trees on the publicly available dataset. The experimental evaluation shows a similar level of accuracy in detecting machine failure, with a slight advantage shown by decision trees. Another highlighted factor regarding the decision between these two machine learning models is the interpretability of the system, whereas decision trees have a significant advantage. The second experimental evaluation included the comparison of deep learning models on visual inspection task. Visual transformer architecture is compared to the dense convolution neural network. Both networks were pretrained on the ImageNet dataset and fine-tuned for visual inspection. The experimental results show high levels of accuracy (around 99% on the test set) of both analyzed networks, with a small advantage achieved by the densenet.

5. Acknowledgment

This research has been financially supported by the Ministry of Science, Technological Development, and Innovation of the Serbian Government, through the project "Integrated research in macro, micro, and nano mechanical engineering – Deep learning and cybersecurity of cyber-physical systems within Industry 4.0" (contract No. 451-03-65/2024-03/200105, 05/02/2024).

6. References

- [1] Petrović, M., Mitić, M., Vuković, N., Miljković, Z.: Chaotic particle swarm optimization algorithm for flexible process planning, *The International Journal of Advanced Manufacturing Technology*, Vol. 85, No. 9–12, pp. 2535–2555, 2016.
- [2] Miljković, Z., Petrović, M.: Application of modified multi-objective particle swarm optimisation algorithm for flexible process planning problem, *International Journal of Computer Integrated Manufacturing*, Vol. 30, No. 2–3, pp. 271–291, 2017.
- [3] Jokić, A., Petrović, M., Miljković, Z.: Mobile robot decision-making system based on deep machine learning, in *9th International Conference on Electrical, Electronics and Computing Engineering (IcETRAN 2022)*, 2022, pp. 653–656.
- [4] Petrović, M., Vuković, N., Mitić, M., Miljković, Z.: Integration of process planning and scheduling using chaotic particle swarm optimization algorithm, *Expert Systems with Applications*, Vol. 64, pp. 569–588, 2016.

- [5] Vorkapić, N., Kokotović, B., Živanović, S.: Wavelet transform packet for signal analysis using standard deviation and signal spectrum, *Innovative Mechanical Engineering*, Vol. 3, No. 1, pp. 28–38, 2024.
- [6] Goyal, D., Choudhary, A., Pabla, B. S., Dhama, S. S.: Support vector machines based non-contact fault diagnosis system for bearings, *Journal of Intelligent Manufacturing*, Vol. 31, pp. 1275–1289, 2020.
- [7] Zhang, W., Peng, G., Li, C.: Bearings fault diagnosis based on convolutional neural networks with 2-D representation of vibration signals as input, in *MATEC web of conferences*, 2017, Vol. 95, p. 13001.
- [8] Behera, S., Choubey, A., Kanani, C. S., Patel, Y. S., Misra, R., Sillitti, A.: Ensemble trees learning based improved predictive maintenance using IIoT for turbofan engines, in *Proceedings of the 34th ACM/SIGAPP symposium on applied computing*, 2019, pp. 842–850.
- [9] Aziz, N., Akhir, E. A. P., Aziz, I. A., Jaafar, J., Hasan, M. H., Abas, A. N. C.: A study on gradient boosting algorithms for development of AI monitoring and prediction systems, in *2020 International Conference on Computational Intelligence (ICCI)*, 2020, pp. 11–16.
- [10] Kumar, A., Gandhi, C. P., Zhou, Y., Kumar, R., Xiang, J.: Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images, *Applied Acoustics*, Vol. 167, p. 107399, 2020.
- [11] Usuga-Cadavid, J. P., Lamouri, S., Grabot, B., Fortin, A.: Using deep learning to value free-form text data for predictive maintenance, *International Journal of Production Research*, Vol. 60, No. 14, pp. 4548–4575, 2022.
- [12] Rathore, M. M., Shah, S. A., Shukla, D., Bentafat, E., Bakiras, S.: The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities, *IEEE Access*, Vol. 9, pp. 32030–32052, 2021.
- [13] Jiang, Y., Dai, P., Fang, P., Zhong, R. Y., Zhao, X., Cao, X.: A2-LSTM for predictive maintenance of industrial equipment based on machine learning, *Computers & Industrial Engineering*, Vol. 172, p. 108560, 2022.
- [14] Matzka, S.: Explainable artificial intelligence for predictive maintenance applications, in *2020 third international conference on artificial intelligence for industries (ai4i)*, 2020, pp. 69–74.
- [15] Tang, T.-W., Kuo, W.-H., Lan, J.-H., Ding, C.-F., Hsu, H., Young, H.-T.: Anomaly detection neural network with dual auto-encoders GAN and its industrial inspection applications, *Sensors*, Vol. 20, No. 12, p. 3336, 2020.
- [16] Shi, J., Li, J., Ma, Q., Yang, Z., Ma, H., Li, L.: CHOPS: CHat with custOmer Profile Systems for Customer Service with LLMs, *arXiv preprint arXiv:2404.01343*, 2024.
- [17] Jokić, A., Petrović, M., Miljković, Z.: Semantic segmentation based stereo visual servoing of nonholonomic mobile robot in intelligent manufacturing environment, *Expert Systems with Applications*, Vol. 190, p. 116203, 2022.

- [18] Miljković, Z., Mitić, M., Lazarević, M., Babić, B.: Neural network Reinforcement Learning for visual control of robot manipulators, *Expert Systems with Applications*, Vol. 40, No. 5, pp. 1721–1736, 2013.
- [19] Petrović, M., Miljković, Z., Jokić, A.: A novel methodology for optimal single mobile robot scheduling using whale optimization algorithm, *Applied Soft Computing*, Vol. 81, p. 105520, 2019.
- [20] Jokić, A., Petrović, M., Miljković, Z.: Methods for visual servoing of robotic systems: A state of the art survey (in Serbian), *Tehnika*, Vol. 73, No. 6, pp. 801–816, 2018.
- [21] Vuković, N., Mitić, M., Miljković, Z.: Trajectory learning and reproduction for differential drive mobile robots based on GMM/HMM and dynamic time warping using learning from demonstration framework, *Engineering Applications of Artificial Intelligence*, Vol. 45, pp. 388–404, 2015.
- [22] Petrović, M., Jokić, A., Miljković, Z., Kulesza, Z.: Multi-objective scheduling of a single mobile robot based on the grey wolf optimization algorithm, *Applied Soft Computing*, Vol. 131, p. 109784, 2022.
- [23] Miljković, Z., Vuković, N., Mitić, M., Babić, B.: New hybrid vision-based control approach for automated guided vehicles, *The International Journal of Advanced Manufacturing Technology*, Vol. 66, No. 1–4, pp. 231–249, 2013.
- [24] Breiman, L., Jerome, F., R.A., O., Stone, C. J.: *Classification and regression trees*. Chapman and Hall/CRC, 2017.
- [25] Quinlan, J. R.: Induction of decision trees, *Machine learning*, pp. 81–106, 1986.
- [26] Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., Data, M.: Practical machine learning tools and techniques, in *Data mining*, 2005, Vol. 2, No. 4, pp. 403–413.
- [27] Vuković, N., Petrović, M., Miljković, Z.: A comprehensive experimental evaluation of orthogonal polynomial expanded random vector functional link neural networks for regression, *Applied Soft Computing Journal*, Vol. 70, pp. 1083–1096, 2018.
- [28] Petrović, M., Jokić, A., Kulesza, Z., Miljković, Z.: Deep learning of mobile service robots, in *Book Service robots – Advances in Research and Applications*, Nova Science Publishers, New York, 2021, pp. 77–97.
- [29] Krizhevsky, A., Sutskever, I., Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, in *Advances in neural information processing systems (NIPS)*, 2012, pp. 1097–1105.
- [30] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition, in *3rd International Conference on Learning Representations (ICLR)*, 2015, pp. 1–14.
- [31] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [32] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: MobileNets: Efficient convolutional neural networks for mobile vision applications, *arXiv preprint arXiv:1704.04861*, 2017.

- [33] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q.: Densely connected convolutional networks, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
- [34] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S.: An image is worth 16x16 words: Transformers for image recognition at scale, *arXiv preprint arXiv:2010.11929*, 2020.
- [35] Nevil, K., Priyank, V., Jenish, H., Ronak, B.: Public Datasets, Casting Product Image Data for Quality Inspection, <https://www.kaggle.com/datasets/ravirajsinh45/reallife-industrial-dataset-of-casting-product>, 2020.
- [36] Møller, M. F.: A scaled conjugate gradient algorithm for fast supervised learning, *Neural networks*, Vol. 6, No. 4, pp. 525–533, 1993.
- [37] Kingma, D. P., Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.

Building AI-Supported Collaborative Awareness in Industrial Humanoids: Conceptual Framework and Methodological Approach

Aleksandar Rodić *¹

Abstract: *With advancements in artificial intelligence, web-oriented applications, cloud technologies, computer-sensor networks, and wireless communications, there is significant potential for the rapid development of robotic systems designed for tasks in both mass and customized industrial production. Collaborative robots, a new type of intelligent robotic system, leverage AI and digital systems to work alongside humans, facilitating the execution of complex technological tasks with greater flexibility and adaptability to infrastructural changes and specific production requirements. This paper presents the conceptual framework and methodological approach for developing artificial consciousness in collaborative robotic systems intended for industrial production. Artificial consciousness enables robots to comprehend the physical and social context of the workplace more quickly and reliably in environments where there is technologically driven cooperation between biological systems (human workers) and technological systems (intelligent service robots). Enhanced perception and mutual understanding, improved predictability, intuitive communication between partners, and developed awareness (self-awareness, awareness of surroundings, and behavior control) are crucial features for collaborative systems in industry, whether it involves human-robot or robot-robot cooperation. This paper will detail the structure of a collaborative robotic system based on the application of AI, cloud architecture, and web-oriented applications.*

Keywords: *intelligent systems, industrial humanoids, collaborative robots, AI-driven robots, context-aware robots.*

1. Introduction

Collaborative work refers to the practice of individuals or groups working together towards a common goal or shared outcome. It emphasizes cooperation, mutual support, and collective effort among participants. The key aspects of a collaborative work can be pointed out as given in the text to follow: a) *Shared Goals* assumes that a collaborative work involves individuals or teams aligning their efforts towards achieving a common objective or solving a shared problem. b) *Mutual Accountability* means that participants in collaborative work

^{*1}Mihajlo Pupin Institute, Univeristy of Belgrade
E-mail: aleksandar.rodic@pupin.rs

are accountable not only to themselves but also to each other. Each person's contribution is essential to the overall success. c) *Open Communication* assumes that effective collaboration requires open and transparent communication among team members. This includes sharing information, ideas, and feedback freely. d) *Division of Labor* means that tasks and responsibilities are often divided among team members based on their strengths, skills, and expertise. This division helps optimize efficiency and leverage individual strengths. e) *Interdependence* means a collaborative work often involves interdependence among team members. Each person's work or contribution impacts and is influenced by the work of others. f) *Respect and Trust* means that successful collaboration is built on a foundation of respect for each other's contributions, skills, and perspectives. Trust among team members fosters a positive working environment. g) *Conflict Resolution* is due a collaboration involves different perspectives and approaches, effective teams develop mechanisms for resolving conflicts constructively. This ensures that disagreements do not hinder progress towards goals. h) *Shared Resources* assumes that collaborative teams often share resources such as tools, information, and expertise to maximize efficiency and effectiveness.

Collaborative work is essential in many contexts, including business, academia, healthcare, and community initiatives. It promotes innovation, problem-solving, and collective achievement by harnessing the strengths and expertise of multiple individuals towards a common purpose. In industrial manufacturing, collaborative work refers to the coordinated efforts of individuals, teams, and machines working together to manufacture products or deliver services efficiently and effectively. Here are several aspects that define collaborative work in industrial production: (i) *Cross-functional Teams*: Collaborative work often involves teams comprised of individuals with diverse skills and expertise (e.g., engineers, technicians, operators, quality assurance specialists). These teams work together to design, develop, and optimize production processes. (ii) *Integrated Workflows*: Collaboration in industrial production emphasizes seamless integration of workflows across different departments or stages of production. For example, design teams collaborate with manufacturing teams to ensure products are manufacturable and meet design specifications. (iii) *Shared Goals and Objectives*: Teams in industrial production share common goals such as meeting production targets, improving quality, reducing costs, or enhancing safety. These goals guide collaborative efforts and align the actions of team members. (iv) *Lean Manufacturing Principles*: Collaborative work often incorporates lean manufacturing principles, such as continuous improvement (kaizen), waste reduction, and just-in-time production. Cross-functional teams work together to identify inefficiencies and implement improvements. (v) *Use of Technology and Automation*: Collaborative work in modern industrial production leverages advanced technologies and automation. For instance, collaborative robots (cobots) work alongside human operators in

tasks such as assembly, welding, or packaging, enhancing productivity and flexibility. (vi) *Real-time Communication and Data Sharing*: Effective collaboration relies on real-time communication and data sharing among team members and across departments. Technologies like Industrial Internet of Things (IIoT) facilitate data collection and analysis, enabling informed decision-making and proactive problem-solving. (vii) *Quality Assurance and Continuous Monitoring*: Collaborative work includes processes for quality assurance and continuous monitoring of production outputs. Teams collaborate to identify defects, analyze root causes, and implement corrective actions to maintain product quality. (viii) *Safety and Ergonomics*: Collaboration in industrial production also focuses on ensuring workplace safety and ergonomic considerations. Teams work together to design workstations, tools, and processes that prioritize worker safety and comfort. (ix) *Training and Skill Development*: Collaborative work includes initiatives for training and skill development to ensure that team members have the necessary knowledge and competencies to perform their roles effectively. This may involve cross-training across different functions or roles within the production process.

A collaborative work in industrial manufacturing fosters a culture of teamwork, innovation, and continuous improvement. It emphasizes the importance of coordination, communication, and shared responsibility in achieving operational excellence and meeting customer demands efficiently.

1.1. Collaborative Robots

Collaborative robots, often referred to as cobots, are a specialized type of robot designed to work alongside humans in a shared workspace or environment. Unlike traditional industrial robots that typically operate separately from human workers behind safety barriers, cobots are intended to collaborate directly with humans without the need for extensive safety measures (though safety considerations are still crucial). The collaborative robots represent a shift towards more flexible, interactive, and human-friendly automation solutions in industrial and other settings. Their ability to enhance productivity, safety, and operational efficiency while supporting human workers makes them increasingly popular in modern manufacturing and beyond. The key characteristics of collaborative robots are: a) *Safety Features*: Cobots are equipped with advanced safety features to ensure safe interaction with human workers. This includes sensors (such as force, torque, and proximity sensors) that enable the robot to detect and respond to the presence of humans nearby, thereby reducing the risk of collisions and injuries. b) *Ease of Use*: Collaborative robots are designed to be user-friendly and easy to program. They often feature intuitive interfaces that allow non-experts to teach or program the robot for various tasks quickly. This flexibility enables rapid deployment and adaptation in dynamic production

environments. c) *Flexible and Adaptive*: Cobots are typically lightweight and mobile, making them easy to move and integrate into different areas of a production line or workspace. They can perform a wide range of tasks, from simple pick-and-place operations to more complex assembly or quality control tasks. d) *Interaction with Humans*: Unlike traditional robots that operate autonomously or in isolated environments, cobots are designed to interact physically and socially with human workers. They can hand tools or parts to humans, collaborate on assembly tasks, or assist in physically demanding or repetitive tasks. e) *Application Diversity*: Collaborative robots find applications across various industries, including manufacturing, healthcare, logistics, and agriculture. They are used in tasks such as assembly, packaging, inspection, material handling, and even in customer service roles. f) *Cost-Effectiveness*: While cobots initially may have higher acquisition costs compared to traditional robots, they can offer cost-effectiveness in terms of faster setup, reconfiguration, and the ability to work alongside existing human workforces without major infrastructure changes. g) *Adaptive Force and Power Limitations*: Cobots are designed with limited force and power capabilities to ensure safety. This means they are programmed to exert forces that are safe for humans, preventing accidental injuries in case of contact.

1.2. Key Topics

Researching intelligent aware collaborative robots with advanced artificial intelligence covers a wide range of different topics from technical to social. The most important topics can be pointed out as to follow: a) *Problems and opportunities of implementation artificial intelligence (AI) in Robotics*, the concept of AI artificial intelligence (AI) in robotics and types of AI to be integrated into robotic controllers (e.g., machine learning, deep learning, fuzzy inference, etc.). b) *Integration of AI into collaborative robots*, how AI enhances the capabilities of collaborative robots and explore specific AI techniques to be used in cobot controllers (e.g., sensor fusion, decision-making algorithms, cloud-driven operation, etc.). c) *Application of AI-enhanced collaborative robots* finding out which industries and sectors benefiting from the AI-powered cobots (e.g., manufacturing, logistics, delivery, etc.) and also case studies or examples showcasing successful implementations. d) *Challenges and limitations* analysing ethical considerations of AI in collaborative robots (e.g., job displacement, safety concerns, protection of personal information, etc.) and technical challenges in integrating advanced AI into cobot controllers. e) *Future trends and innovations* like predictions for the future of AI in collaborative robotics and emerging technologies and research directions (e.g., human-robot interaction, explainable AI). f) *Comparison with traditional robots* analysing advantages and disadvantages of AI-enhanced cobots compared to traditional industrial robots

and economic implications and return on investment. g) *Regulatory and safety standards* including current regulations governing the use of collaborative robots and safety protocols and standards for AI-powered cobots.

2. State-Of-The-Art In Collaborative Robotics

The evolution from traditional industrial robots to intelligent, aware collaborative robots reflects significant advancements in AI, sensor technology, computing power, and connectivity. Researching intelligent, aware collaborative robots with advanced artificial intelligence encompasses a broad spectrum of topics. One key area of interest is their evolution from traditional industrial robots to the sophisticated collaborative robots we see today. Collaborative robots are not only transforming industrial production by improving flexibility, safety, and efficiency but are also paving the way for new forms of human-robot interaction and cooperation. As technology continues to evolve, we can expect collaborative robots to become even more capable and integrated into various aspects of manufacturing and beyond.

Traditional industrial robots have been a staple in manufacturing since the mid-20th century. These robots are designed to perform repetitive, precise tasks such as welding, painting, assembly, and material handling. Key characteristics of traditional industrial robots include: a) *High Precision and Speed*: These robots are optimized for accuracy and efficiency in performing repetitive tasks. b) *Fixed Automation*: They are typically programmed to follow a set sequence of operations in a controlled environment. c) *Isolation from Humans*: Traditional industrial robots are often placed in cages or behind safety barriers to prevent human workers from coming into contact with them. This is because these robots lack the ability to detect and respond to humans in their vicinity, posing safety risks.

Several technological advancements have paved the way for the development of collaborative robots: a) *Artificial Intelligence (AI)*: AI algorithms have enabled robots to learn from their environment, make decisions, and adapt to new tasks. b) *Sensors and Machine Vision*: Advanced sensors and machine vision systems have given robots the ability to perceive their surroundings, detect objects, and recognize human presence. c) *Improved Computing Power*: Enhanced computing power has allowed for real-time data processing and complex decision-making. d) *Wireless Communication and IoT*: The integration of wireless communication and the Internet of Things (IoT) has enabled seamless connectivity between robots, machines, and human operators.

Emergence of Collaborative Robots, or cobots, represents a significant shift in the design and application of industrial robots. Key characteristics of collaborative robots include: a) *Human-Robot Interaction*: Cobots are designed to work alongside humans, without the need for safety barriers. They are

equipped with sensors and safety features to detect and respond to human presence, ensuring safe interaction. b) *Flexibility and Adaptability*: Unlike traditional robots, cobots can be easily reprogrammed and reconfigured for different tasks. This makes them ideal for small-batch production and tasks that require frequent changeovers. c) *Ease of Use*: Collaborative robots are designed to be user-friendly, with intuitive programming interfaces that allow even non-experts to set up and operate them. d) *Enhanced Safety Features*: Cobots are equipped with force and torque sensors that enable them to detect unexpected obstacles and stop operations to prevent accidents. This makes them inherently safer for working in close proximity to humans.

Intelligent, Aware Collaborative Robots is the latest generation of collaborative robots incorporates advanced AI and awareness capabilities, making them even more versatile and effective. These robots exhibit: a) *Artificial Consciousness*: Advanced AI enables robots to understand and interpret the physical and social context of their environment, improving their ability to cooperate with human workers. b) *Higher Degrees of Perception*: With improved sensory systems, these robots can gather and process detailed information about their surroundings, leading to better decision-making and task execution. c) *Intuitive Communication*: Intelligent cobots can engage in more natural and intuitive communication with human partners, using gestures, speech recognition, and other human-like interaction methods. d) *Predictive Capabilities*: AI allows these robots to predict human actions and adjust their behavior accordingly, enhancing collaboration and efficiency.

The evolution from traditional industrial robots to intelligent, aware collaborative robots reflects significant advancements in AI, sensor technology, computing power, and connectivity. Collaborative robots are not only transforming industrial production by improving flexibility, safety, and efficiency but are also paving the way for new forms of human-robot interaction and cooperation. As technology continues to evolve, we can expect collaborative robots to become even more capable and integrated into various aspects of manufacturing and beyond.

2.1. Examples of Collaborative Robots for Industry

There are several well-known examples of collaborative robots (cobots) that have made significant impacts across various industries. Here are some notable examples: (i) *Universal Robots (UR)*: Universal Robots [1] is one of the pioneers in collaborative robotics. Their robots are known for their flexibility, ease of use, and safety features. UR cobots are widely used in industries such as automotive, electronics, and food processing for tasks like assembly, packaging, and machine tending. (ii) *ABB YuMi*: ABB's YuMi[2] (short for 'You and Me') is a dual-arm collaborative robot designed for small parts assembly. It features

precise motion control and advanced vision capabilities, allowing it to handle delicate tasks alongside human operators. YuMi is used in industries like electronics, consumer goods, and pharmaceuticals. (iii) *FANUC CR series*: FANUC offers a range of collaborative robots under their CR series [3]. These cobots are designed for applications such as assembly, inspection, and machine tending. FANUC's cobots are known for their reliability, speed, and integration capabilities with existing FANUC robotic systems. (iv) *KUKA LBR iiwa*: KUKA's LBR iiwa [4] (Intelligent Industrial Work Assistant) is a sensitive lightweight robot designed to work safely alongside humans without the need for safety fencing. It is used in industries such as automotive, aerospace, and medical technology for tasks requiring precision and sensitivity. (v) *Rethink Robotics Sawyer and Baxter*: Rethink Robotics, acquired by HAHN Group, developed the Sawyer and Baxter robots [5]. These cobots are known for their intuitive programming interface and adaptive capabilities. They are used in manufacturing environments for tasks like machine tending, quality inspection, and assembly. (vi) *Cobots from Techman Robot*: Techman Robot [6] (TM Robot) produces collaborative robots known for their user-friendly interface and integrated vision systems. Their cobots are used in industries such as electronics, plastics, and food processing for applications ranging from pick-and-place to complex assembly tasks. (vii) *Doosan Cobots*: Doosan Robotics offers a range of cobots designed for industrial automation tasks [7]. Their cobots are known for their safety features, accuracy, and ease of integration. They are used in sectors such as automotive, logistics, and metalworking. (viii) *Optimus cobot, 2nd Gen*: The second generation of Elon Musk's Optimus robot [8] represents a significant advancement in collaborative robotics, particularly in car manufacturing and industry. With its advanced AI, robust sensory systems, human-robot interaction capabilities, and integration with modern manufacturing systems, the Optimus robot is poised to enhance productivity, safety, and flexibility in the automotive industry. As these technologies continue to evolve, the role of collaborative robots like Optimus will become increasingly pivotal in driving innovation and efficiency in manufacturing.

The Optimus co-bot, developed by Tesla automotive industry, is an example of the increasing interest of industry for developing high-tech industrial service robots dedicated to specific applications in car manufacturing like: a) *Assembly Line Operations*: The robot can perform precise and repetitive tasks such as installing components, tightening bolts, and assembling parts, reducing the physical strain on human workers and increasing overall productivity. b) *Quality Control and Inspection*: With its advanced vision systems, the Optimus robot can conduct detailed inspections of car components and finished vehicles, identifying defects and ensuring high-quality standards are maintained. c) *Material Handling and Logistics*: The robot can handle and transport materials and components across the manufacturing floor, optimizing logistics and

reducing downtime. d) *Welding and Painting*: Equipped with specialized tools, the Optimus robot can perform welding and painting tasks with high precision and consistency, enhancing the quality and appearance of the final product. e) *Customization and Flexibility*: The robot's ability to quickly adapt to different tasks and configurations makes it ideal for handling customized production runs and accommodating changes in manufacturing processes.

These examples highlight the diversity of collaborative robots in terms of capabilities, applications, and industries served. Each of these cobots has unique features and strengths, but all share the common goal of enhancing productivity, flexibility, and safety in industrial environments by working in collaboration with human operators.

2.3. Where Use Cobots

Collaborative robots are utilized in a wide range of industrial tasks and technological processes across various sectors due to their ability to work safely alongside humans and perform diverse functions. Here are examples of industrial tasks and processes where cobots are commonly employed: (i) *Assembly and Disassembly*: Cobots are used in assembly lines to handle small parts assembly and disassembly tasks. They can work alongside human operators to improve efficiency and reduce repetitive strain injuries. Universal Robots' UR cobots are widely used for assembly tasks in industries such as automotive and electronics. (ii) *Pick and Place Operations*: Collaborative robots excel in pick and place operations where they pick items from one location and place them in another. This is common in logistics, warehousing, and assembly operations. Techman Robot's TM cobots are used for pick and place tasks in industries like e-commerce and consumer goods. (iii) *Machine Tending*: Cobots are employed in machine tending tasks where they load and unload parts from machines such as CNC machines, injection molding machines, and 3D printers. FANUC CR cobots are used for machine tending in automotive and aerospace manufacturing. (iv) *Quality Inspection*: Collaborative robots equipped with vision systems and sensors are used for quality inspection tasks. They can inspect parts for defects, measure dimensions, and perform quality control checks. ABB YuMi cobots are used for quality inspection in industries such as electronics and pharmaceuticals. (v) *Packaging and Palletizing*: Cobots are employed in packaging lines to handle packaging materials, pack products into boxes, and palletize finished goods. They improve efficiency and throughput in packaging operations. KUKA LBR iiwa cobots are used for packaging and palletizing tasks in food and beverage industries. (vi) *Material Handling and Logistics*: Collaborative robots play a crucial role in material handling tasks, including sorting, transporting, and organizing materials in warehouses and distribution centers. Sawyer and Baxter cobots from Rethink Robotics are used in

logistics for material handling and order fulfillment. (vii) *Testing and Inspection*: Cobots are used for testing and inspection of products and components, performing tasks such as non-destructive testing, surface inspection, and functional testing. Cobots from companies like Doosan Robotics are utilized in automotive and electronics industries for testing and inspection tasks. (viii) *Collaborative Welding and Cutting*: Advanced cobots with welding and cutting capabilities are used in metalworking industries for tasks such as welding, soldering, and cutting operations. Universal Robots' UR cobots are adapted with welding tools for collaborative welding tasks in automotive and manufacturing sectors.

Cobots are designed to work alongside humans in various industrial tasks. However, there are certain tasks and industries where their application may be limited or not feasible due to technical, safety, or efficiency reasons. These are for example:

High-Speed Manufacturing. Tasks that require extremely high speeds, such as in some automotive assembly lines or high-speed packaging lines, may not be suitable for collaborative robots due to their typically slower operational speeds compared to traditional industrial robots.

High-Precision Assembly. Tasks that demand very high precision and accuracy beyond what current cobot technology can reliably achieve might not be suitable. This includes certain microelectronics assembly or intricate watchmaking processes.

High-Load Handling. While cobots can handle a range of loads, tasks involving very heavy or bulky items that exceed their lifting capacities may require traditional industrial robots with higher payload capabilities.

Extreme Environments. Industries dealing with extreme temperatures, hazardous chemicals, or other harsh environmental conditions may find it challenging to use standard cobots, which are often designed for typical factory environments.

Specialized Manufacturing Processes. Some specialized manufacturing processes, such as those requiring unique tooling or highly customized operations that are not easily adaptable to cobot programming, may still rely on traditional automation solutions. But, with higher-level of AI and learning even these tasks can be accomplished.

Complex Decision-Making. Tasks that require complex decision-making based on dynamic and unpredictable environments may be challenging for current cobots, which typically operate based on predefined algorithms and sensor inputs.

Long-Term Continuous Operation. Cobots are generally designed for intermittent or short-duration tasks. Applications requiring long-term continuous operation without frequent human intervention may still favor traditional automation systems designed for sustained uptime.

While collaborative robots continue to evolve with advancements in AI, machine learning, and sensor technology, addressing some of these limitations, these are some of the current constraints in which their application may be less feasible compared to traditional industrial robots or other automation solutions.

3. Features of Collaborative Robots

Collaborative robots, to be effectively utilized in industry, need advanced physical/manipulative, cognitive and social traits to meet demanding technological requirements.

3.1. Physical and Manipulative Traits

Collaborative service robots in industry require specific physical and manipulative attributes (traits) to effectively perform tasks alongside human workers while ensuring safety, efficiency, and adaptability. Here are key attributes that collaborative service robots should possess: (i) safety features, (ii) mobility and navigation, (iii) manipulation and dexterity, (iv) interaction and communication, (v) Sensory Perception, (vi) Ease of Programming and Operation, (vii) Adaptability and Learning, and (viii) Reliability and Maintenance.

Safety features assumes: a) Collision detection and avoidance: Robots should have sensors to detect the presence of humans and other objects in their vicinity. They should be able to slow down or stop to prevent collisions. b) Soft and rounded design: Physical attributes such as soft exteriors and rounded edges minimize the risk of injury in case of accidental contact with humans.

Mobility and navigation assumes: a) Omni-directional Movement: Robots should be capable of moving in multiple directions (e.g., forwards, backwards, sideways) to navigate complex environments. b) Obstacle Avoidance: Advanced navigation systems and sensors allow robots to navigate around obstacles autonomously.

Manipulation and Dexterity assumes: a) End effector variety: Robots should have interchangeable end effectors (tools or grippers) suitable for different tasks, such as handling objects of various shapes and sizes. b) Precision and Accuracy: High precision in movement and manipulation enables robots to perform delicate tasks with accuracy, such as assembly, inspection, and placing objects.

Interaction and Communication assumes: a) Speech recognition and synthesis: Robots should be able to understand and respond to voice commands from human operators or interact verbally in natural language. b) Gesture recognition: Capabilities to interpret human gestures and respond appropriately enhance collaboration and ease of communication.

Sensory Perception assumes: a) Vision systems: Advanced cameras and vision systems enable robots to perceive their environment, identify objects, and

navigate autonomously. b) Tactile sensors: Sensors that provide feedback on contact forces enable robots to handle fragile objects delicately and safely.

Ease of Programming and Operation assumes: a) Intuitive Programming Interfaces: User-friendly interfaces for programming and controlling robots enable non-experts to teach or reconfigure robots quickly. b) Remote Monitoring and Control: Capability for remote monitoring and control allows supervisors to oversee robot operations and intervene if necessary.

Adaptability and Learning assumes: a) Machine Learning and AI: Robots should have the ability to learn from experience, adapt to changing tasks or environments, and continuously improve their performance. b) Task Flexibility: Capability to perform a wide range of tasks and adapt to new tasks or workflows as needed in dynamic industrial settings.

Reliability and Maintenance assumes: a) Durability: Robust construction and components ensure reliability in industrial environments with frequent use. B) Predictive Maintenance: Systems for monitoring robot health and predicting maintenance needs reduce downtime and optimize operational efficiency.

These attributes collectively enable collaborative service robots to work effectively alongside human workers in industries such as manufacturing, logistics, healthcare, and hospitality. By possessing these traits, collaborative robots contribute to increased productivity, improved safety, and enhanced operational flexibility in diverse industrial applications.

3.2. Cognitive and Social Traits

Cobots to be implemented in industry should possess cognitive and social traits that enable effective interaction and cooperation with human workers. These traits are crucial for fostering a collaborative and productive work environment. Here are key cognitive and social traits that collaborative robots should have.

Cognitive traits assumes that a robot possesses or is able to practically demonstrate: (i) context awareness, (ii) task understanding, (iii) learning and adaptation, (iv) decision-making, and (v) problem-solving.

Context Awareness. Cobots should understand their operational context, including their surroundings, tasks being performed, and the presence of human operators. This awareness helps them adapt their behavior accordingly.

Task Understanding. Ability to comprehend task instructions and objectives given by human supervisors or through programming. This includes understanding task sequences, priorities, and dependencies.

Learning and Adaptation. Capability to learn from interactions and experiences in the workplace. Machine learning algorithms allow cobots to improve their performance over time, adjust to changes in tasks or environments, and optimize their behavior.

Decision-Making. Cobots should possess decision-making capabilities to autonomously select actions and responses based on sensory inputs, task requirements, and predefined rules or algorithms. This includes making decisions to ensure safety, efficiency, and task completion.

Problem-Solving. Ability to identify and respond to unexpected situations or challenges that may arise during task execution. Cobots should be equipped with algorithms or logic to troubleshoot issues and seek assistance if needed.

Social traits assumes that robot possesses or is able to present: (i) communication skills, (ii) collaboration and teamwork, (iii) social intelligence, (iv) adaptability to human preferences, (v) respect for safety and boundaries, (vi) transparency and explainability, (vii) empathy and assistance.

Communication Skills. Effective communication with human operators through verbal cues, visual signals, or digital displays. This includes conveying status updates, requesting assistance, or providing feedback on task progress.

Collaboration and Teamwork. Ability to collaborate seamlessly with human workers as part of a team. Cobots should understand their role within the team, coordinate actions with other robots or humans, and adapt to collaborative workflows.

Social Intelligence. Cobots should demonstrate social intelligence by interpreting human emotions, intentions, and non-verbal cues (e.g., gestures, facial expressions). This enhances their ability to engage with humans in a socially acceptable manner.

Adaptability to Human Preferences. Cobots should be able to adapt their behavior and interaction style based on individual human preferences or cultural norms. This fosters a more personalized and comfortable working relationship.

Respect for Safety and Boundaries. Cobots should prioritize safety by respecting human safety zones, following established safety protocols, and adjusting their speed or behavior to avoid collisions or accidents.

Transparency and Explainability. When making decisions or taking actions, cobots should provide transparency by explaining their reasoning or rationale to human operators. This enhances trust and facilitates effective collaboration.

Empathy and Assistance. While not necessarily emotional empathy, cobots should demonstrate a form of practical empathy by anticipating human needs, offering assistance when appropriate, and adjusting their behavior to support human operators.

These cognitive and social traits are essential for collaborative robots to integrate seamlessly into industrial environments and work effectively alongside human counterparts. By embodying these traits, cobots contribute to improved productivity, safety, and overall operational efficiency in diverse industrial applications.

3.3. Cobots and Personality Traits

Collaborative robots intended for industrial applications typically do not need personality traits or affective behaviors in the same way humans do. Whether such features are necessary and their importance for collaborative robots in industry are discussed in the text to follow.

Personality traits with industrial collaborative robots are not of primary interest while they are crucial for other types of human-centric robots like for example social robots (elderly care, nursery, etc.).

Utility vs. Personality. Primary focus is on utility. Industrial collaborative robots are primarily designed for functional purposes, such as performing tasks efficiently, safely, and reliably. Personality traits, which are more characteristic of human interaction, are not essential for their core functionality.

Consistency and Predictability. Primary focus is on consistent behavior. Robots in industrial settings benefit from consistent and predictable behavior. This ensures that they perform tasks reliably and can be easily integrated into existing workflows without introducing variability that might hinder operational efficiency.

Task-Oriented Interaction. Primary focus is on task efficiency. Cobots are programmed to optimize task performance and operational outcomes. Personality traits that deviate from this task-oriented focus could potentially distract from their primary purpose or introduce unnecessary complexity.

Affective behavior of collaborative robot may have an influence on their capabilities.

Safety and Interaction. While affective behavior like displaying emotions or responses might enhance human-robot interaction in some contexts, in industrial settings, clear communication and predictable responses are more critical for ensuring safety and efficiency.

Perception of Machines. Humans tend to anthropomorphize machines to some extent, but in industrial settings, clear distinctions between human and robotic roles help maintain operational clarity and safety protocols.

When considering the importance of personality traits and affective behaviour of collaborative robots the following items should be taken into account: a) efficiency and productivity, b) user experience and acceptance, and c) ethical considerations.

Efficiency and Productivity. The primary focus of collaborative robots in industry is on enhancing efficiency, productivity, and safety. Features like personality traits and affective behavior, while potentially interesting from a human interaction standpoint, do not significantly contribute to these primary goals.

Clear Communication. Effective communication, clear task execution, and reliable performance are crucial for user acceptance and integration of

cobots into industrial workflows. Predictable behavior and task-focused interaction facilitate smoother collaboration with human workers.

Ethical Alignment. In contexts where robots are closely integrated with human workers, ethical considerations arise regarding how robots are perceived and how their behavior affects human psychology and work dynamics. Clarity and transparency in robot behavior can help mitigate potential ethical concerns.

While personality traits and affective behaviors are not necessary for collaborative robots in industrial applications, maintaining clear, predictable, and task-focused interaction is paramount. These features ensure that cobots effectively fulfill their intended roles, enhance operational efficiency, and support safe collaboration with human counterparts in industrial environments.

4. Building Context Aware Humanoid

Building of intelligent, aware industrial humanoid service robot for accomplishing different collaborative tasks in manufacturing, is part of a long-term R&D project conducted in the Mihajlo Pupin Institute. In the first phase of development, the mechanical prototype of the robot was developed [9]. In the current phase of development, the intelligent robot controller is designed and the appropriate methodology of building is defined.

The 21 degrees of freedom (DOFs) bi-manual robot mechanism for customized manufacturing consists of several functional modules (Figs. 1 and 2): (i) 3 DOFs motorized cart, (ii) 4 DOFs robot torso, (iii) 2 DOFs robot head (pan-tilt camera), (iv) two robot arms of 6 DOFs every, and (v) two single DOF robot grippers (two-finger or optionally three-finger). The mechanical structure of the robot's torso was designed so that the robot can change its geometry from a fully extended position (Fig. 1, right) to a fully bent position (Fig. 1, left).

Control of the bi-manual collaborative robot (Fig. 2) is a complex task because the system consists of several mechanical subsystems (modules) that are kinematic and dynamically coupled to each other. The robot prototype is fully equipped by heterogeneous sensors measuring position, speed and forces/torques in the system. Robot also possesses necessary communication facility to connect to the network (Internet, Cloud). The hardware architecture of the control system has a distributed hierarchical structure. It is subordinated to the higher levels of control as shown in the block diagram presented in Fig. 3 as presented in [10].

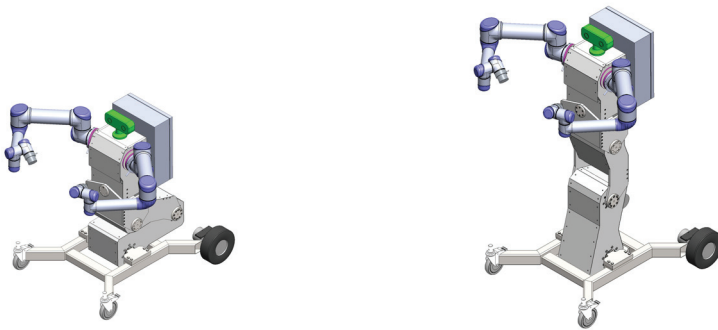


Figure 1. Wheel-based bi-manual collaborative industrial humanoid developed in Mihajlo Pupin Institute – Belgrade: contracted (left) and extended (right), [9]



Figure 2. Laboratory prototype of the 21 DOFs bi-manual collaborative industrial humanoid developed in Mihajlo Pupin Institute – Belgrade [9]

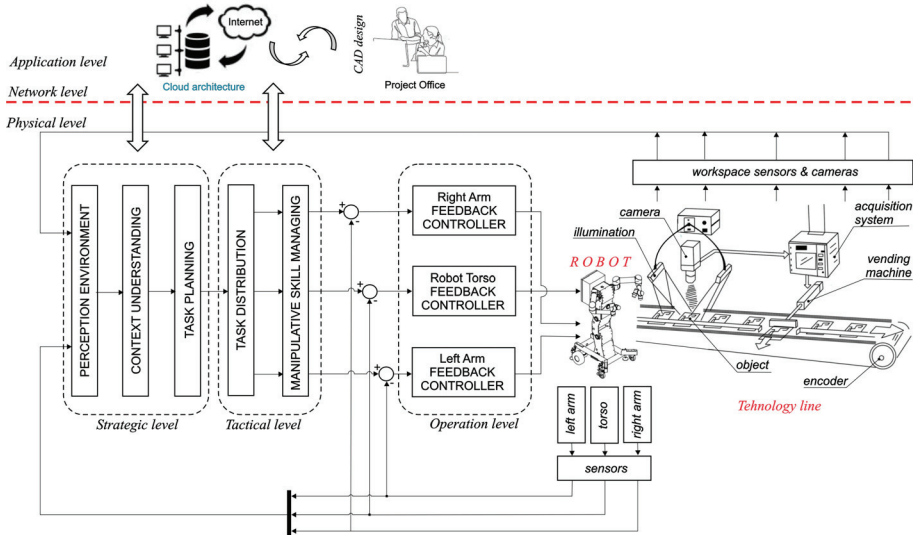


Figure 3. High-level system description of the cloud-enabled intelligent control of the mobile bi-manual service robot with variable geometry [10]

The hardware architecture, presented in Fig. 3, supports building of an intelligent, robot-aware controller. The conceptual framework and methodological insights are presented in the following paragraphs. The concept of building intelligent robot controller is based on implementation of advanced AI algorithms, cloud-driven robot behavior, web-oriented applications, use of digital twins, etc.

4.1. Artificial Intelligence and Cobots

Artificial intelligence (AI) plays a crucial role in enhancing the capabilities and effectiveness of collaborative robots (cobots) in industry [11]–[17]. The AI contributes to cobots by enabling advanced features such as better: (i) *Adaptability and Flexibility*: AI enables cobots to adapt to changing environments and tasks. Machine learning algorithms allow cobots to learn from experience and adjust their behavior accordingly. This adaptability is essential in dynamic manufacturing environments where tasks may vary. (ii) *Autonomous Decision-Making*: AI-powered cobots can make autonomous decisions based on real-time data and sensory inputs. They can analyze information from sensors and adjust their actions to optimize performance, improve efficiency, and ensure safety. (iii) *Enhanced Sensory Perception*: AI enables cobots to have advanced sensory perception capabilities. This includes computer vision for object recognition, depth sensing for precise manipulation, and tactile sensing for handling delicate objects—all of which are critical for safe collaboration with humans. (iv)

Predictive Maintenance: AI algorithms can analyze data collected from cobots to predict maintenance needs and prevent breakdowns. This proactive approach reduces downtime and improves overall equipment effectiveness (OEE). (v) *Collaboration with Humans:* AI facilitates natural and intuitive human-robot collaboration. Cobots can interpret human gestures, understand commands, and adjust their behavior to work effectively alongside human operators without compromising safety. (vi) *Continuous Improvement:* AI enables continuous learning and improvement in cobots' performance over time. Through data analytics and feedback loops, cobots can identify inefficiencies, refine processes, and optimize workflows for greater productivity. (vii) *Complex Task Execution:* AI allows cobots to handle more complex tasks that require sophisticated decision-making and problem-solving abilities. This includes tasks such as adaptive welding, autonomous navigation in crowded environments, and collaborative problem-solving in assembly processes. (viii) *Integration with IoT and Industry 4.0:* AI-driven cobots play a key role in the integration of Industrial Internet of Things (IIoT) and Industry 4.0 initiatives. They contribute to the digital transformation of manufacturing by enabling data-driven decision-making and connectivity across the production ecosystem.

Artificial intelligence is essential to maximizing the potential of collaborative robots in industry. It enables cobots to be smarter, more adaptive, safer, and more productive collaborators with human workers. As AI technologies continue to evolve, they will further enhance the capabilities of cobots, making them indispensable tools for modern manufacturing and beyond.

Bearing in mind the importance of AI in developing intelligent robots for industry, the OpenAI's ChatGPT is considered as a tool that has a great potential to support development of new cobot capabilities. Cobots are able to potentially connect online to the ChatGPT to ask for advice or interpretation about phenomena in its physical or social environment [18]-[24]. How this could be achieved and the potential effects on collaborative activities between robots and humans are elaborated in the following paragraph.

Hardware requirements

Computing Power. The robot would need sufficient onboard computing power to process natural language and interact with the AI model in real-time. This typically involves a capable processor and memory to handle AI inference tasks.

Internet Connectivity. Stable and high-speed internet connectivity is essential for the robot to communicate with the AI model in the cloud or a remote server where the AI is hosted.

Sensors and Perception Systems. To provide context for its queries, the robot may use various sensors such as cameras, microphones, environmental sensors,

and perhaps even specialized sensors for social interaction cues (e.g., facial recognition).

Protocols (Procedures)

API Integration. The robot would interact with the AI model through an API (Application Programming Interface). The API allows the robot to send queries and receive responses from the AI model. Protocols such as HTTP or WebSocket can be used for communication between the robot's software and the AI server.

Natural Language Processing (NLP). The robot's software needs to include NLP capabilities to formulate queries in natural language and understand responses from the AI model. This involves parsing text, understanding context, and extracting relevant information.

Error Handling and Recovery. Implementing error handling mechanisms is crucial to manage scenarios where the connection is disrupted or the AI model returns unexpected responses. This ensures robustness in the interaction between the robot and the AI.

Effects on Collaborative Activities

Enhanced Problem-Solving and Decision-Making. By tapping into AI models like ChatGPT, robots can access a vast repository of knowledge and expertise. This can aid in solving complex problems or interpreting ambiguous situations in the physical or social environment.

Improved Human-Robot Collaboration. Robots equipped with AI-driven advice capabilities can collaborate more effectively with human counterparts. They can seek guidance on tasks, clarify instructions, or even mediate in social interactions by interpreting human intentions or emotional states.

Adaptability and Learning. Continuous interaction with AI models allows robots to learn from previous interactions and improve their own decision-making processes over time. This adaptive learning capability enhances their ability to perform diverse tasks in dynamic environments.

Efficiency and Productivity. Quick access to information and advice from AI models can streamline workflows and reduce the time required for decision-making. This efficiency contributes to overall productivity in collaborative activities involving robots and humans.

Ethical and Social Considerations. Integrating AI advice capabilities should consider ethical guidelines, such as privacy protection, data security, and transparency in decision-making processes. Clear communication of the robot's reliance on AI advice ensures trust and collaboration with human partners.

Connecting collaborative robots to the OpenAI's ChatGPT models, looking for a quick advice or interpretation, can significantly enhance their capabilities in understanding and responding to the physical and social

environment. With appropriate hardware, protocols, and integration of AI capabilities, collaborative robots can contribute more effectively to collaborative activities, improving efficiency, decision-making, and human-robot interaction quality in various industrial and social settings.

4.2. Cloud-driven Cobot Capacities

Integrating cloud architecture with industrial collaborative robots can enhance their operational capacities in several dimensions [25]-[30]. There is an analysis of how cloud architecture could potentially address the current limitations of modern robots.

Memory and Processing Power

Scalability. Cloud architecture provides virtually unlimited storage and computational resources. Robots can offload intensive tasks such as complex computations, large-scale data processing, or advanced AI algorithms to the cloud. This allows them to overcome physical limitations of onboard memory and processing power.

Speed. Cloud computing offers high-speed parallel processing capabilities, enabling robots to handle real-time data analysis and decision-making more efficiently than relying solely on onboard processors.

Data Accessibility and Collaboration

Data Storage. Robots can store and access vast amounts of data in the cloud, including historical performance data, machine learning models, and environment maps. This enables them to continuously improve their performance and adapt to changing conditions.

Collaboration. Cloud-connected robots can easily share information with each other and with human operators in real-time. This facilitates collaborative tasks where multiple robots coordinate their actions based on shared data and goals.

Software Updates and Maintenance

Remote Updates. Cloud architecture allows for centralized management of software updates and patches. This ensures that robots always have access to the latest algorithms, security enhancements, and bug fixes without needing physical intervention.

Monitoring and Diagnostics. Continuous monitoring of robot performance and diagnostics can be conducted remotely through the cloud. This proactive approach to maintenance reduces downtime and improves overall reliability.

AI and Machine Learning Capabilities

Training and Deployment. Cloud-based AI platforms enable robots to leverage advanced machine learning models and algorithms that require extensive training data and computational resources. This includes natural language processing, image recognition, and predictive analytics.

Adaptability. Robots can adapt their behavior and decision-making based on real-time insights from cloud-based AI systems, enhancing their ability to handle complex and dynamic environments.

Security and Privacy Considerations

Data Security. Cloud providers typically offer robust security measures, including encryption, access controls, and regular audits, which can be leveraged to protect sensitive data transmitted and stored by robots.

Privacy. Efforts must be made to ensure that sensitive data, such as proprietary information or personal data, is handled securely and in compliance with relevant regulations.

By integrating cloud architecture with industrial collaborative robots can significantly enhance their operational capacities by providing scalable memory and processing power, facilitating data accessibility and collaboration, streamlining software updates and maintenance, empowering advanced AI capabilities, and addressing security and privacy concerns effectively. This approach not only expands the technical capabilities of robots but also supports their integration into increasingly complex and interconnected industrial environments.

4.3. Web-oriented Cobot Applications

Enabling a collaborative robot to connect to web browsers for quick access to various databases and information can significantly enhance its performance and capabilities in several key areas [31]-[37].

Context Awareness

Real-Time Information Retrieval. By accessing web databases and information sources, the robot can gather up-to-date contextual information relevant to its tasks. This includes environmental conditions, operational status of machinery, inventory levels, and any dynamic changes in the workspace.

Adaptive Decision-Making. With access to current data, the robot can make more informed decisions in real-time, adapting its actions based on the latest information available. This improves efficiency and responsiveness in collaborative tasks.

Social (Collaborative) Awareness

Human-Robot Collaboration. Access to web-based resources can enhance the robot's ability to understand and adapt to human behaviors and intentions. For instance, it can retrieve information about human preferences, work habits, and task priorities, thereby improving its collaboration with human partners.

Communication Enhancement. Web connectivity enables the robot to communicate effectively with humans by retrieving contextual information that facilitates clearer and more contextually relevant interactions.

Awareness of Human Health and Affective State

Health Monitoring. By connecting to health databases or wearable devices, the robot can monitor human health indicators such as heart rate, body temperature, and stress levels. This information can be crucial for adjusting its interactions and workload to ensure human well-being and productivity.

Affective Computing. Accessing affective computing databases or sentiment analysis tools allows the robot to recognize and respond to human emotions more accurately. This capability supports empathetic interactions and improves overall user satisfaction and comfort.

Business Partner Integration

Access to Business Data. Connecting to web browsers allows the robot to retrieve business-related data such as inventory levels, production schedules, customer orders, and market trends. This information helps the robot optimize its tasks and contribute more effectively to business operations.

Integration with Enterprise Systems. The robot can integrate with enterprise resource planning (ERP) systems or customer relationship management (CRM) platforms through web connectivity, enabling seamless data exchange and enhancing overall operational efficiency.

Security and Privacy Considerations

Secure Access. Ensure that web connections are secure, utilizing encryption and authentication protocols to protect sensitive data and prevent unauthorized access.

Data Privacy. Respect privacy regulations and guidelines when accessing personal or sensitive information, ensuring that data handling practices comply with legal requirements and ethical standards.

By leveraging web connectivity for accessing databases and information, collaborative robots can achieve greater context awareness, enhance human-robot collaboration, monitor human health and affective states, integrate with business processes, and ultimately improve their overall performance and utility in industrial and collaborative settings. This approach contributes to more

adaptive, efficient, and user-friendly robotic systems in diverse operational environments.

Achieving real-time or semi real-time decision-making supported by web-search tools for collaborative robots is nowadays feasible with the right approach and infrastructure. In order to achieve high-performance decision-making in this context it is required the following:

Fast and Reliable Internet Connectivity

Bandwidth. A robust internet connection with sufficient bandwidth is crucial for rapid data retrieval and communication between the robot and web-based resources.

Latency. Minimizing latency (delay in data transmission) ensures that the robot can access information quickly, which is essential for real-time decision-making.

Efficient Data Processing and Integration

Data Processing Speed. The robot's onboard processing capabilities should be optimized for rapid data parsing, analysis, and integration with real-time information retrieved from web-search tools.

Integration Framework. Implementing an efficient integration framework that allows seamless communication between the robot's software systems and external web APIs (Application Programming Interfaces) is essential. This framework should handle data formats, protocols, and security considerations effectively.

AI and Machine Learning Algorithms

Fast Algorithms. Utilize algorithms that are optimized for speed and efficiency in processing and making decisions based on dynamic and large datasets obtained from web sources.

Adaptive Learning. Incorporate machine learning techniques that can adapt to new information and refine decision-making processes over time based on feedback and changing environmental conditions.

Real-Time Data Analytics

Streaming Analytics. Implement real-time data analytics capabilities that can process incoming data streams from web-search tools immediately and derive actionable insights without significant delay.

Predictive Analytics. Use predictive models to anticipate future trends or events based on historical data and current information retrieved from web sources.

User Interface and Interaction Design

Intuitive Interface. Design a user interface that allows human operators to easily interact with the robot and understand the decision-making process facilitated by web-search tools.

Visualization. Use data visualization techniques to present complex information in a clear and actionable format, aiding both human operators and the robot itself in making informed decisions.

Security and Reliability

Secure Communication. Implement robust security measures to protect data exchanged between the robot and web-search tools, ensuring confidentiality, integrity, and availability.

Redundancy and Failover. Have contingency plans in place, such as redundant internet connections or fallback mechanisms, to maintain operation in case of network disruptions or failures.

Achieving high-performance decision-making supported by web-search tools involves a combination of advanced technology infrastructure, efficient data handling capabilities, optimized algorithms, and secure communication protocols. With these elements in place, collaborative robots can effectively leverage external web resources to enhance their real-time or semi real-time operational capabilities in dynamic industrial environments.

4.4. Cobots and Digital Twins

Digital twins refer to virtual representations or digital models of physical objects, processes, or systems. They are essentially digital counterparts that mimic real-world entities and can simulate their behavior, characteristics, and interactions in a virtual environment [38]-[44]. Here are key aspects and applications of digital twins:

Key Characteristics

Virtual Representation. Digital twins are virtual or digital replicas of physical entities, such as machines, products, systems, or processes.

Real-time Data Integration. They integrate real-time data from sensors, IoT devices, and other sources to simulate and mirror the real-world counterpart's behavior and performance.

Simulation and Predictive Capabilities. Digital twins can simulate various scenarios and predict outcomes based on historical data and predictive analytics.

Feedback Loop. They enable a continuous feedback loop where data collected from the physical entity updates the digital twin, and insights gained from the twin can inform decisions about the physical entity.

Lifecycle Management. Digital twins can represent the entire lifecycle of a product, process, or system—from design and development to operation, maintenance, and eventual disposal.

Applications in Manufacturing and Industrial Processes

Predictive Maintenance. Digital twins of machines predict maintenance needs and optimize schedules to reduce downtime.

Process Optimization. Simulate manufacturing processes to optimize efficiency, quality, and resource usage.

Product Development

Design Verification. Use digital twins to test and validate product designs virtually before physical prototyping.

Performance Monitoring. Monitor product performance and behavior in real-world conditions to improve design iterations.

Digital twins are powerful tools that bridge the physical and digital worlds, offering numerous benefits across industries by enhancing understanding, decision-making, and operational efficiency based on real-time data and simulations. Digital twins can significantly enhance the properties and capabilities of collaborative industrial robots by providing simulation, monitoring, and optimization capabilities throughout their lifecycle. Digital twins can be used to improve various aspects of collaborative robots.

Virtual Prototyping

Purpose. Create digital twins of robot designs to simulate and validate mechanical structures, kinematics, and ergonomics before physical prototyping.

Benefits. Reduces design iterations, accelerates development timelines, and ensures that robots are optimized for performance, safety, and human-robot interaction.

Performance Optimization

Simulation of Tasks. Use digital twins to simulate robot tasks in virtual environments, such as assembly, welding, or handling operations.

Optimization of Motion Planning. Evaluate and optimize motion trajectories, speed profiles, and path planning algorithms to maximize efficiency and minimize cycle times.

Real-Time Monitoring and Control

Sensor Integration. Connect physical robots to their digital twins, integrating real-time data from sensors (e.g., position, force, temperature).

Condition Monitoring. Monitor robot performance metrics (e.g., speed, accuracy) and health indicators (e.g., motor vibrations, temperature) to detect anomalies and predict maintenance needs.

Predictive Maintenance

Failure Prediction. Analyze data collected from digital twins to predict potential failures or maintenance issues before they occur.

Optimized Maintenance Scheduling. Optimize maintenance schedules based on actual usage patterns and predicted wear and tear, minimizing unplanned downtime.

Task Optimization and Adaptability

Dynamic Task Allocation. Use digital twins to dynamically allocate tasks among robots based on workload, skillsets, and environmental conditions.

Adaptive Behavior. Adjust robot behavior and responses in real-time to optimize collaboration with human workers and other robots in shared workspaces.

Training and Simulation:

Skill Development. Provide virtual training environments using digital twins to simulate complex tasks and scenarios for robot operators.

Scenario Testing. Test new operational procedures or task sequences in a risk-free virtual environment before implementation in the real-world.

Data-Driven Insights

Performance Analysis. Analyze historical data from digital twins to identify trends, performance bottlenecks, and opportunities for optimization.

Iterative Refinement. Use insights gained to iteratively improve robot design, operation strategies, and human-robot collaboration protocols.

Integration with Industry 4.0 Initiatives

Digital Thread. Integrate digital twins into broader Industry 4.0 frameworks to create a digital thread across design, manufacturing, and operational phases.

Smart Factory Integration. Enable seamless integration with other smart factory technologies and systems for enhanced automation, data exchange, and decision-making.

Case Example

In an automotive manufacturing setting, digital twins of collaborative robots are used to simulate assembly tasks. The digital twin allows engineers to optimize robot placement, refine assembly sequences, and validate the robot's

ability to work alongside human workers safely and efficiently. Real-time data from the physical robots are fed back to the digital twin, enabling continuous monitoring, predictive maintenance, and performance optimization.

Digital twins offer a comprehensive framework to enhance the properties of collaborative industrial robots by enabling virtual prototyping, real-time monitoring, predictive maintenance, operational optimization, and continuous improvement. They play a pivotal role in maximizing the efficiency, safety, and adaptability of robots in dynamic industrial environments.

5. Methodological Approach

Building collaborative traits into industrial service robots involves leveraging various methods and technologies to enhance their interaction, communication, and performance alongside human workers. Here are key building methods for implementing collaborative traits in industrial service robots:

Sensor Integration

Purpose. Sensors play a crucial role in enabling robots to perceive their environment, interact safely with humans, and adapt to changing conditions.

Types of Sensors. Include cameras for vision-based tasks, depth sensors for object detection and navigation, proximity sensors for detecting human presence, and force/torque sensors for handling tasks requiring delicate touch.

Implementation. Integrate sensor data with control algorithms (e.g., localization, object recognition) to enhance robot awareness and responsiveness.

Artificial Intelligence (AI) and Machine Learning.

Purpose. AI enables robots to learn, reason, and make decisions autonomously, enhancing their cognitive capabilities and adaptive behavior.

Techniques. Utilize machine learning algorithms for tasks such as object recognition, motion planning, task scheduling, and predictive maintenance.

Implementation. Train AI models on data collected from sensors and operational feedback to improve task performance, optimize workflows, and adapt to dynamic environments.

Natural Language Processing (NLP)

Purpose. Enable robots to understand and respond to human commands and inquiries through speech recognition and synthesis.

Techniques. Implement NLP algorithms for parsing and understanding natural language instructions, generating appropriate responses, and interacting verbally with human operators.

Implementation. Integrate NLP capabilities into robot control systems to facilitate intuitive communication and enhance user interaction.

Human-Robot Interfaces (HRI)

Purpose. Design intuitive interfaces that facilitate communication, collaboration, and control between human operators and robots.

Types of Interfaces. Include graphical user interfaces (GUIs), touchscreens, mobile applications, and wearable devices for remote monitoring and control.

Implementation. Develop user-centered designs that prioritize ease of use, clarity of information, and ergonomic considerations to support effective human-robot collaboration.

Collaborative Planning and Control

Purpose. Enable robots to collaborate seamlessly with human workers in performing complex tasks and shared activities.

Techniques. Implement algorithms for task allocation, motion coordination, and adaptive control strategies that consider human intentions, safety constraints, and task priorities.

Implementation. Integrate collaborative planning frameworks (e.g., shared autonomy, task-level programming) to enable robots to work alongside humans in a coordinated and efficient manner.

Behavioral and Social Skills

Purpose. Develop robots with social intelligence to enhance interaction quality, teamwork, and user acceptance in industrial environments.

Techniques. Incorporate algorithms for recognizing human emotions, responding appropriately to social cues, and adapting behavior to foster positive human-robot relationships.

Implementation. Implement behavioral models that govern robot actions, gestures, and expressions to convey intentions, maintain engagement, and ensure safety in collaborative tasks.

Safety and Ethical Considerations

Purpose. Ensure that robots operate safely alongside humans, respecting ethical guidelines and promoting trust and acceptance in the workplace.

Techniques. Integrate safety mechanisms such as collision detection, emergency stop systems, and safe operation zones (SOZs) to prevent accidents and protect human operators.

Implementation. Design robots with built-in safety features, conduct risk assessments, and adhere to industry standards and regulations to ensure safe deployment and operation in industrial settings.

By leveraging these building methods, industrial service robots can embody collaborative traits that enhance their functionality, interaction capabilities, and overall performance in diverse industrial applications. These methods enable robots to work effectively alongside human workers, contributing to improved productivity, safety, and operational efficiency in modern industrial environments.

3. Conclusion

With the launch of the Industry 4.0 initiative, collaborative industrial robots are gaining increasing importance. The demanding global market compels production companies to cater more closely to customer preferences. As a result, mass industrial production is gradually evolving into customized production, which is more flexible, faster, and better able to meet specific requirements without requiring significant infrastructural changes to technological production lines.

In manufacturing, human workers will increasingly share their workplaces with collaborative robots that assist with various technological tasks. Currently, humans are still more dexterous and creative than robots, but robots can work for extended periods without fatigue or loss of concentration. The goal is for service robots in the industry to increasingly approach human capabilities and take over previously irreplaceable roles. Achieving this requires the introduction of new digital tools and advanced artificial intelligence algorithms. These will enable future generations of service robots to not only perform with speed, precision, and repeatability but also possess a higher degree of understanding of their physical and social environments. This will allow robots to take on more substantial roles in manufacturing processes.

Moreover, the integration of remote control, predictive capabilities, and new analytical models will create a novel production environment for the so-called Factories of the Future. This paper presents the concept of controller development for an industrial service humanoid and the methodology that, at the current level of development, can lead to the desired goal: an intelligent, collaborative, and aware service robot.

4. References

- [1] Universal Robots: <https://www.universal-robots.com>, [Accessed: 30.06.2024]
- [2] ABB Robotics: <https://new.abb.com/products/robotics>, [Accessed: 30.06.2024]
- [3] FANUC Robotics: <https://www.fanucamerica.com>, [Accessed: 30.06.2024]
- [4] KUKA Robotics: <https://www.kuka.com/en-gb/products/robotics-systems/industrial-robots/lbr-iiwa>, [Accessed: 30.06.2024]
- [5] *Rethink Robotics Sawyer and Baxter*, <https://www.rethinkrobotics.com>, [Accessed: 30.06.2024]
- [6] Techman Robot: <https://www.tm-robot.com>, [Accessed: 30.06.2024]
- [7] DOSAN Robotics: <https://www.doosanrobotics.com/en/Index>, [Accessed: 30.06.2024]
- [8] Tesla OPTIMUS Humanoid: https://www.teslarati.com/tesla-optimus-hands-22-degrees-of-freedom-upgrade-2024/#google_vignette, [Accessed: 30.06.2024]
- [9] Rodić, A., Stevanović, I., Jovanović, M., *Smart Cyber-Physical System to Enhance Flexibility of Production and Improve Collaborative Robot Capabilities – Mechanical Design and Control Concept*, in Proceedings of the 27th International Conference on Robotics in Alpe-Adria Danube Region (RAAD 2018), Springer, In book: *Advances in Service and Industrial Robotics*, Eds. Nikos A. Aspragathos, Panagiotis N. Koustoumpardis, Vassilis C. Moulianitis, Part of the Mechanisms and Machine Science book series, Vol. 67, pp. 627-639, ISBN: 978-3-030-00231-2, online ISBN: 978-3-030-00232-9, DOI: 10.1007/978-3-030-00232-9_66, 2018.
- [10] Rodić, A., Šumarac, J., Stevanović, I., Jovanović, M., *Cloud-enabled bi-manual collaborative robot with enhanced versatility for customized production*, Springer, Book Chapter in *Advances in Service and Industrial Robotics*, Eds. Said Zeghloul, Med Amine Laribi, Juan Sandoval. Book Series: *Mechanisms and Machine Science*, ISSN: 2211-0984, ISSN 2211-0992 (electronic), ISBN: 978-3-030-75258-3, ISBN: 978-3-030-75259-0 (eBook), <https://doi.org/10.1007/978-3-030-75259-0>, 102, pp. 240–249, 2021
- [11] Diehl, F., Utz, H., & Jeschke, S., *Artificial Intelligence for Robotics and Autonomous Systems: Theory and Applications*, Springer, ISBN: 3031287150, 9783031287152, 2020.
- [12] Zhu, Y., Mottaghi, R., & Gupta, S., *Deep Reinforcement Learning for Industrial Robot Control*, IEEE Robotics and Automation Letters, Volume: 4, Issue: 2, pp. 1300-1307, ISSN: 2377-3766, 2019.

- [13] Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D., *Artificial Intelligence in Industry 4.0: State of the Art and Future Trends*, Journal of Manufacturing Systems, Volume: 45, pp. 134-146, ISSN: 0278-6125, 2016.
- [14] Garcia, E., Jimenez, M. A., De Santos, P. G., & Armada, M., *Application of AI Techniques in Industrial Robotics: A Review*, IEEE International Conference on Robotics and Automation (ICRA), pp. 3653-3658, ISSN: 1050-4729, 2007.
- [15] Mohan, S., & Likhachev, M., *Towards Autonomous Industrial Robots: Integrating AI Planning and Execution*, IEEE International Conference on Intelligent Robots and Systems (IROS), pp. 754-759, ISSN: 2153-0858, 2018.
- [16] Zhang, X., Wang, X., & Du, Y., *AI-Driven Predictive Maintenance for Industrial Robots*, IEEE International Conference on Automation Science and Engineering (CASE), pp. 355-360, ISSN: 2161-8089, 2019.
- [17] Rozo, L., Jimenez, P., & Torras, C., *Real-time AI-based Human-Robot Collaboration in Manufacturing*, IEEE International Conference on Robotics and Automation (ICRA), pp. 402-407, ISSN: 1050-4729, 2013.
- [18] S. Johnson, M. Patel, A. Kumar, *Enhancing Industrial Robotics with Natural Language Processing: A Study Using OpenAI's ChatGPT*, International Journal of Advanced Robotic Systems, 2022.
- [19] T. Smith, E. Wang, J. Liu, *Leveraging Conversational AI for Real-time Monitoring and Control of Industrial Robots*, IEEE Transactions on Industrial Informatics, 2023.
- [20] L. Brown, Y. Chen, M. Davis, *Integrating ChatGPT with Industrial Robotics for Enhanced Human-Robot Collaboration*, IEEE International Conference on Robotics and Automation (ICRA), 2022.
- [21] P. Gonzalez, K. Lee, R. Martin, *AI-driven Conversational Interfaces for Industrial Robot Programming*, IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 2021.
- [22] J. Wilson, F. Rodriguez, A. Singh, *Utilizing OpenAI's ChatGPT for Predictive Maintenance in Industrial Robotics*, IEEE International Conference on Automation Science and Engineering (CASE), 2023.
- [23] H. Kim, D. Park, S. Lee, *Natural Language Interfaces for Industrial Robots: A ChatGPT-based Approach*, ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2022.
- [24] F. Chen, A. Wang, R. Gupta, *Conversational AI for Robotics: Techniques and Applications*, Springer, 2023.
- [25] Mohanan, P., *Cloud Robotics: Integrating Cloud Computing with Robotic Systems*, CRC Press, ISBN: 9780367187352, 2020.
- [26] Kehoe, B., Patil, S., Abbeel, P., & Goldberg, K., *A Survey of Cloud Robotics: Techniques, Applications, and Challenges*, IEEE Transactions on

- Automation Science and Engineering, Vol.: 12, Issue: 2, Pages: 398-409, ISSN: 1545-5955, 2015.
- [27] Chen, J., Shi, W., & Zhao, Q., *Cloud Robotics for Efficient Big Data Processing in Cooperative Multi-Robot Systems*, IEEE Transactions on Industrial Informatics, Volume: 17, Issue: 6, Pages: 4181-4191, ISSN: 1551-3203, 2021.
- [28] Li, Y., Ma, L., & Zhang, Z., *Cloud-based Control and Coordination of Industrial Robots for Smart Manufacturing*, IEEE International Conference on Robotics and Automation (ICRA), pp. 588-593, ISSN: 1050-4729, 2018.
- [29] Lu, Y., Xu, X., & Xu, W., *Secure Cloud Computing for Collaborative Robots in Industry 4.0*, IEEE International Conference on Industrial Informatics (INDIN), Pages: 45-50, ISSN: 1932-8184, 2019.
- [30] Wang, P., & Jiang, H., *Optimizing Robot Task Execution Using Cloud-based Artificial Intelligence*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Pages: 1327-1332, ISSN: 2153-0858, 2020.
- [31] Z. Zhang, Y. Li, H. Wang, *Web-based Control and Monitoring System for Industrial Robots*, International Journal of Advanced Manufacturing Technology, Volume: 96, pp. 3379-3390, ISSN: 0268-3768, 2018.
- [32] H. Wang, L. Ma, X. Zhang, *A Cloud-based Web Application for Industrial Robot Programming and Control*, Robotics and Computer-Integrated Manufacturing, Volume: 61, pp.1018-1027, ISSN: 0736-5845, 2020.
- [33] J. K. Lee, S. Park, M. Choi, *A Web-based Interface for Remote Control of Industrial Robots*, IEEE International Conference on Robotics and Automation (ICRA), pp. 1372-1377, ISSN: 1050-4729, 2017.
- [34] A. Müller, T. Wendt, R. Anderl, *Web Services for Industrial Robot Integration in Smart Factories*, IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1-8, ISSN: 1946-0740, 2019.
- [35] X. Liu, D. Zhang, Y. Shen, *Developing a Web-based Monitoring System for Industrial Robot Health Management*, IEEE International Conference on Automation Science and Engineering (CASE), pp. 1024-1029, ISSN: 2161-8089, 2018.
- [36] M. Wang, T. Li, J. Huang, *Interactive Web Applications for Collaborative Robot Programming and Operation*, IEEE International Conference on Human-Robot Interaction (HRI), pp.160-167, ISSN: 2167-2121, 2021.
- [37] C. G. Jensen, D. C. Karnopp, *Industrial Robotics: Programming, Simulation and Applications*, InTechOpen, ISBN: 978-1-78923-672-6, 2020.
- [38] Kłosowski, T., *Digital Twin Technologies and Smart Cities*, Springer, ISBN: 978-3-030-38790-7, 2020.

- [39] Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C., *A Digital Twin for Smart Manufacturing*, Journal of Manufacturing Systems, Volume: 48, pp. 157-168, ISSN: 0278-6125, 2018.
- [40] Qi, Q., & Tao, F., *Digital Twin-driven Smart Manufacturing and Cyber-physical Systems*, Journal of Manufacturing Systems, Volume: 48, pp. 19-26, ISSN: 0278-6125, 2018.
- [41] Uhlemann, T. H.-J., Lehmann, C., & Steinhilper, R., *Implementation of Digital Twin for Real-time Monitoring of Robotic Systems*, : IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 876-881, ISSN: 2157-3611, 2017.
- [42] Lee, J., Bagheri, B., & Kao, H.-A., *Digital Twin for Predictive Maintenance in Industry 4.0*, IEEE International Conference on Automation Science and Engineering (CASE), pp. 155-160, ISSN: 2161-8089, 2015.
- [43] Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W., *Towards the Integration of Digital Twins with Robotics for Industrial Applications*, IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1-8, ISSN: 1946-0740, 2018.
- [44] Leng, J., Zhang, H., Yan, D., Liu, Q., Chen, X., & Zhang, D., *Digital Twin-based Framework for Remote Control of Industrial Robots*, IEEE International Conference on Robotics and Automation (ICRA), pp. 4889-4894, ISSN: 1050-4729, 2019.

Advancements in Robotic Intelligence: The Role of Computer Vision, DRL, Transformers and LLMs

Lejla Banjanović-Mehmedović^{*1}, Anel Husaković², Azra Gurdić Ribić³,
Naser Prljača¹, Isak Karabegović⁴

Abstract: In recent advancements in robotics, Artificial Intelligence (AI) methods such as Deep Learning, Deep Reinforcement Learning (DRL), Transformers, and Large Language Models (LLMs) have significantly enhanced robotic capabilities. Key AI models driving advancements in robotic vision include Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), the DETection Transformers (DETR), the YOLO family of algorithms, segmentation techniques, and 3D vision technologies.

Deep Reinforcement Learning (DRL), an AI technique where agents learn optimal behaviors through trial and error interactions with their environment, enables robots to perform complex tasks autonomously. Transformers, originally developed for natural language processing, have been adapted to robotics for tasks involving sequence prediction and data understanding, improving perception and decision-making processes. LLMs leverage vast amounts of text data to enhance robot-human interaction, enabling robots to understand and generate human-like language, thus improving their communicative and collaborative abilities in various applications. The integration of these AI methods enhances the adaptability, efficiency, and overall performance of robotic systems, paving the way for more sophisticated and intelligent autonomous agents.

Keywords: Artificial Intelligence, Robotics, Computer Vision, Deep Learning, Deep Reinforcement learning, Transformers, Large Language models, Perception, Control, Decision Making, Sequence Prediction

1. Introduction

The field of robotic intelligence has witnessed significant advancements in recent years, driven by the integration of cutting-edge technologies such as Deep Learning (DL), Deep Reinforcement Learning (DRL), Transformers, and Large Language Models (LLMs)[1-4]. These technologies are revolutionizing the way robots learn, perceive, and interact with their environments, leading to unprecedented levels of autonomy and efficiency.

^{*1}University of Tuzla, Faculty of Electrical Engineering, Tuzla, Bosnia and Herzegovina

²Eacon doo, Zenica, Bosnia and Herzegovina

³Foundation for Innovation, Technology and Transfer of Knowledge, Bosnia and Herzegovina

⁴Academy of Sciences and Arts of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina
E-mail: lejla.mehmedovic@untz.ba, naser.prljaca@untz.ba, anel@eacon.ba, azra.gurdic@fet.ba, isak1910@hotmail.com

Deep Reinforcement Learning (DRL) has emerged as a pivotal technique in robotic learning, combining the principles of reinforcement learning with the powerful representation capabilities of deep neural networks. DRL allows robots to learn optimal behaviors by interacting with their environments and receiving feedback in the form of rewards or penalties, facilitating the trial-and-error learning process. This approach has proven particularly effective in complex and dynamic tasks, such as robotic manipulation, navigation, and autonomous.

Originally developed for natural language processing (NLP), transformers have been adapted for various applications in robotics because of their capability to handle sequential data and understand long-range dependencies. The self-attention mechanism in transformers enables robots to concentrate on significant elements of their sensory inputs, making them highly suitable for tasks requiring the integration of multimodal information. This capability enhances robots' understanding and decision-making processes, enabling more sophisticated and context-aware interactions with their environments driving [5-8].

Large Language Models (LLMs) are transforming the interface between humans and robots by providing advanced natural language understanding and generation capabilities. LLMs enable robots to comprehend and execute complex instructions given in natural language, allowing for more intuitive and flexible human-robot interactions [4,9-10]. These models can interpret context, generate detailed action plans, and integrate information from multiple sources, significantly improving the robots ability to perform intricate tasks autonomously.

2. Fundamental Modules in Robotics

In robotics, fundamental capabilities include *perception, control, decision-making, planning and interaction*.

Perception is a crucial skill for robots, involving the use of sensors to collect environmental data. This capability enables robots to recognize objects, interpret visual information, and understand spatial relationships [11]. By leveraging advanced algorithms and AI models, computer vision enables robots to process visual information, identify objects, understand scenes, and make informed decisions. Techniques such as Convolutional Neural Networks (CNNs) [12] and Vision Transformers (ViTs) [6] have revolutionized this field, providing robust and accurate solutions for tasks like image recognition, object detection, and scene segmentation. As a result, robots equipped with sophisticated vision capabilities are increasingly capable of performing complex tasks in dynamic and unstructured environments [13], from industrial automation and autonomous driving to healthcare and service robotics [14].

Transformers have transformed robotic perception by improving the ability to process and interpret sensory data. Initially designed for natural language processing, transformers have been adapted for vision tasks [13], such as object

detection and scene understanding [15-16], through Vision Transformers (ViTs). ViTs can handle large amounts of visual data and extract meaningful features, which is crucial for robots to recognize and respond to their surroundings effectively. This advanced perception allows robots to perform tasks such as autonomous navigation and object manipulation with greater accuracy and reliability [6].

A variety of multi-modal models have been developed to tackle tasks like visual question answering, image captioning, and speech recognition [17]. The related technologies for perception used in robotics are Large Vision Models (LVMs), that can “see” clearly, but are “blind” in reasoning [18], Multimodal Large Language Model (MLLM), where LLMs and LVMs work together for discriminative and generative tasks and Vision-language-action (VLA) model, where pre-existing vision-language models, are trained without any new parameters to output text-encoded robot actions [19].

The control module is vital for managing robotic actions. Reinforcement Learning (RL) significantly contributes to robot control by enabling robots to determine optimal actions through trial and error. Deep Reinforcement Learning (DRL) algorithms allow robots to develop control policies that maximize cumulative rewards over time. This learning process is particularly beneficial for tasks requiring precise motor control and adaptation to changing environments. DRL has been successfully applied in robotic arms for tasks like picking and placing objects, where the robot learns to control its movements to achieve the desired outcomes efficiently [20]. The paper explores enhancing human-robot team performance by using Q-learning algorithms to adjust task loads based on real-time physiological data analysis [21].

The planning module in robotics is essential for enabling autonomous robots to make decisions and execute tasks effectively [22]. It involves the creation of strategies or sequences of actions that allow robots to achieve specific goals while navigating their environments [23]. Key components of robotic planning include *path planning*, *motion planning*, and *task planning*. The paper [24] presents the use of AI methods such as machine learning algorithms, heuristic search techniques, and optimization algorithms to enhance the efficiency and effectiveness of autonomous disinfection robots in navigating complex environments to combat the spread of COVID-19. The application of the Grey Wolf Optimization algorithm is presented in the paper [25] to significantly enhance the efficiency and effectiveness of service robot path planning, improving their operational performance in various environments.

Large Language Models (LLMs), like GPT-4, have advanced robotic planning by enabling robots to understand and execute complex instructions given in natural language. LLMs can translate high-level commands into detailed action sequences, making it easier for robots to plan and perform tasks autonomously.

This capability is enhanced when combined with visual input, allowing LLMs to create visual-semantic plans that integrate information from multiple modalities. Such integration improves the robots' ability to generate precise and context-aware plans, essential for tasks like autonomous navigation and complex manipulations. [26-27].

The decision-making module is a key component of robots, allowing them to make informed choices and plan tasks based on their current state and environment. As the core of a robot, decision-making integrates information from the perception module to produce suitable actions, bridging the gap between previous inputs and future responses. Sequence Precision involves executing tasks in a specific order with high accuracy, crucial for applications like assembly lines and surgical robots where precise timing and order are essential for success. Multimodal Decision Making involves for example LLMs, which can integrate information from various sources (e.g., visual, auditory, and textual data) to make context-aware decisions. This integration allows robots to generate detailed and accurate action sequences based on comprehensive situational understanding [4, 27].

Interaction is a crucial module that allows robots to engage with both their environment and humans. To improve their capability to interact effectively in the physical world, robots typically undergo extensive training [28-30].

The use of LLMs has also significantly improved human-robot interaction. LLMs' advanced natural language understanding allows robots to engage in more intuitive and meaningful conversations with humans. This ability to comprehend and respond to nuanced language makes robots more user-friendly and capable of performing tasks based on verbal instructions. Furthermore, the integration of LLMs with other AI technologies enables robots to provide personalized and context-aware interactions, enhancing their utility in various applications, from customer service to personal assistance [4].

3. AI Transform Robotics

AI is revolutionizing robotics by allowing robots to perform complex tasks with enhanced autonomy, accuracy, and flexibility [31]. Machine learning (ML) algorithms enable robots to learn from data and improve their performance over time, while computer vision systems provide them with the ability to perceive and understand their environment.

Natural language processing (NLP) helps robots communicate with humans more effectively, enhancing their usability in diverse applications. This transformation is paving the way for a future where intelligent robots seamlessly integrate into everyday life, providing innovative solutions and enhancing productivity across various sectors.

3.1. Advancements in Computer Vision in Robotics Based on AI Models

Convolutional Neural Networks (CNNs) have played a crucial role in the advancement of image recognition [32]. These networks are built to automatically and adaptively learn the spatial hierarchies of features from input images, which makes them extremely effective for tasks like image classification and object detection. CNNs are widely used in robotics for tasks like sorting objects, quality control in manufacturing, and facial recognition in security systems [33].

Vision Transformers (ViTs) represent a newer approach, leveraging transformer architectures that have proven highly effective in capturing long-range dependencies in data. ViTs have demonstrated superior performance in various image recognition tasks and are increasingly being adopted in robotics for applications requiring high accuracy and efficiency [16].

The DEtection TRansformer (DETR) is a novel approach that uses transformers for object detection. It has shown promising results in accurately detecting and recognizing objects in images, making it suitable for complex robotics applications [16].

The YOLO (You Only Look Once) family of algorithms has revolutionized real-time object detection and recognition [34]. YOLO models divide images into grids and predict bounding boxes and class probabilities for objects within these grids simultaneously, allowing for rapid and accurate object detection. This capability is crucial in robotics for tasks such as autonomous navigation [35], where robots need to detect and respond to objects in their environment in real time.

Segmentation involves partitioning an image into multiple segments to simplify its representation and make it more meaningful. Segmentation techniques, including Fully Convolutional Networks (FCNs), U-Net, Mask_R-CNN which provide detailed pixel-wise classification, essential for applications such as robotic surgery and automated inspection systems [36-37].

3D vision technologies enable robots to perceive and interact with their environment in three dimensions. This is achieved through techniques such as stereo vision, depth sensors, and LiDAR. In the realm of 3D vision, algorithms like PointNet and VoxelNet [38-39] process 3D point clouds to enable precise object detection and spatial understanding, facilitating tasks such as robotic grasping, 3D mapping, and autonomous driving.

3.2. Reinforcement Learning Algorithms

Reinforcement Learning (RL) aims to solve the sequential decision-making problem by taking actions that maximize expected rewards. Essentially, the agent follows a policy to make a series of decisions (i.e., take actions) in different states of the environment [40-41]. The sequence of these states and

actions creates a trajectory. Each decision within this policy is evaluated based on the accumulated rewards over the trajectory to determine the policy's effectiveness. By evaluating these trajectories, the agent improves the policy by increasing the likelihood of decisions that yield higher expected rewards. Through this iterative process of trial and error, the agent continually refines the policy until it reaches an optimal state.

To improve the policy, we utilize the Bellman optimality equations to update the value functions by choosing the action that provides the highest value, rather than considering all possible actions.

Various RL techniques have been proposed from different perspectives to optimize the policy. A detailed classification of RL algorithms is presented in Figure 1.

(Deep) Reinforcement Learning (RL) can be divided into model-free and model-based algorithms [41,42]. The main distinction between these categories lies in whether the agent has access to a model of the environment, specifically the transition function and the reward function. *Model-based algorithms*, like AlphaZero, obtain or learn a model of the environment to predict future values or states.

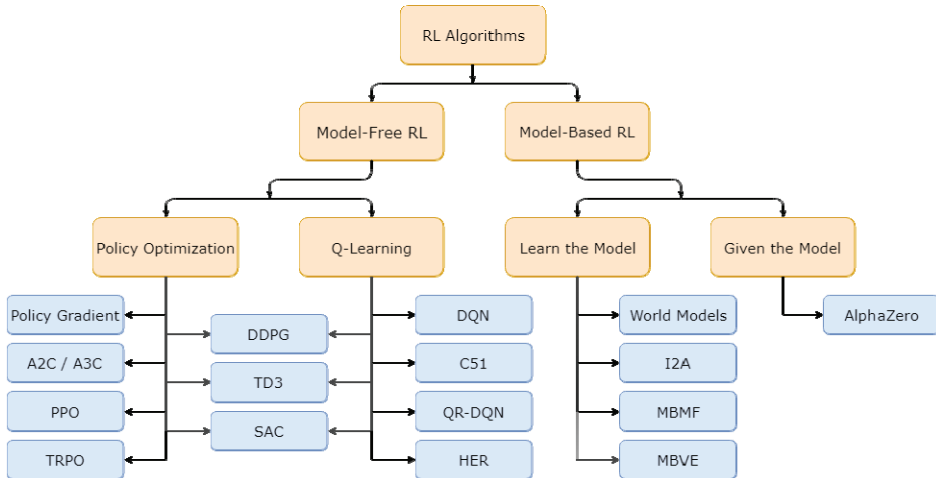


Figure 1. Classification of Reinforcement Learning [42]

In many RL scenarios, the reward and transition functions are unknown due to the environment's complexity and intricate mechanisms. As a result, agents often use *model-free methods*, learning the policy solely from the experience gained through interactions with the environment.

In its early stages, reinforcement learning (RL) was suited to scenarios with discrete and limited state and action spaces, enabling agents to record information in tables. However, modern tasks like playing Go or autonomous

driving involve large discrete state-action spaces or continuous values, rendering table-based methods impractical. To overcome this challenge, Deep Reinforcement Learning (DRL) integrates RL and deep learning (DL), with RL setting the problem and optimization objectives, while DL models the policy and expected rewards. Depending on the role of deep neural networks (DNN) in DRL, it can be classified into three categories [43].

In value-based methods, deep neural networks (DNN) learn a value function for discrete action spaces to assess each potential action. DNN do not participate in policy decision-making; instead, they are used to estimate the policy's performance. Figure 2 presents the schematic of the DQN model [44].

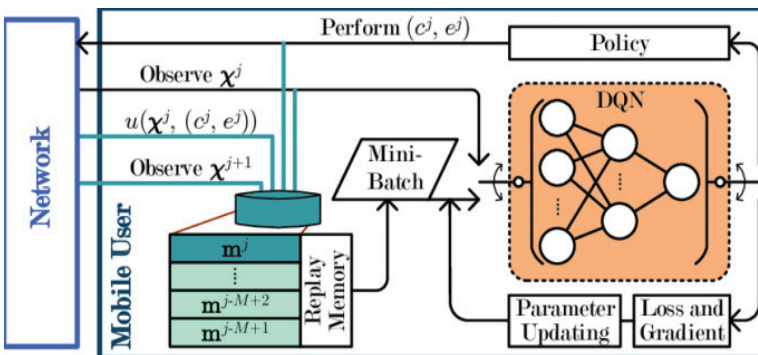


Figure 2. Schematics of the DQN model [44]

In policy-based methods, deep neural networks (DNN) are directly involved in selecting actions. Compared to value-based methods, policy-based algorithms generally offer better convergence and can learn stochastic policies, whereas value-based methods select actions deterministically based on the maximum Q value. *Policy-based algorithms*, like *Proximal Policy Optimization (PPO)*, offer a continuous action space and aim to directly map states to actions by developing a representation of the actual behavior policy. Figure 3 illustrates the network structure utilized for the PPO [45].

Value-based methods tend to be less stable and exhibit poorer convergence compared to policy-based approaches, as they rely on approximating the Q-function. However, value-based methods are more sample-efficient. On the other hand, policy-based methods are more prone to getting stuck in local optima due to the large search space.

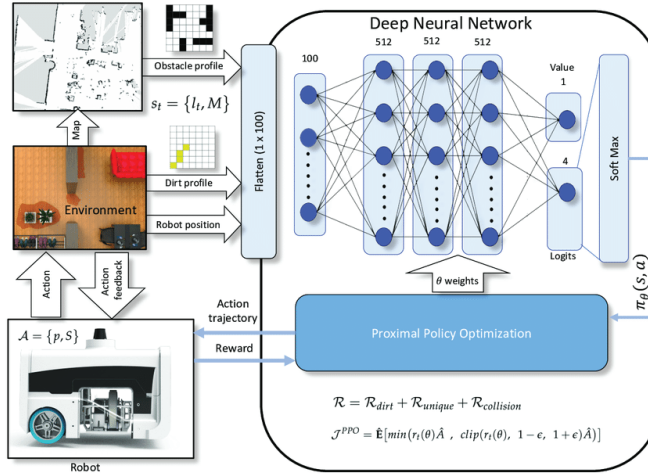


Figure 3. Schematic of the PPO model [45]

The hybrid actor-critic model integrates both approaches, using two distinct deep neural networks (DNNs) named the actor and the critic. During each training iteration, the actor evaluates the current state and the policy to decide on an action. The environment then transitions to a new state and provides a reward. The critic updates its parameters based on this feedback and rates the actor's action. Subsequently, the actor adjusts its policy based on the critic's evaluation. Recent examples of algorithms that employ this model include *Deep Deterministic Policy Gradient (DDPG)* and *Asynchronous Advantage Actor-Critic (A3C)*, as depicted in Figure 4 [43].

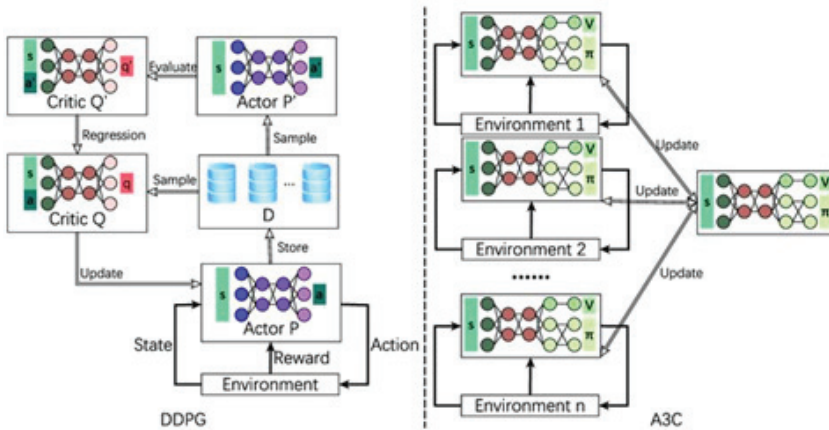


Figure 4. Schematic of the DDPG and A3C models. [43]

DDPG incorporates the concepts of target deep neural networks (DNNs) and experience replay memory [46,47]. Despite these improvements, DDPG's performance is hindered by the operation of experience replay. This limitation can be overcome by *asynchronous* training approach in A3C (Asynchronous Advantage Actor-Critic) [48]. Namely, multiple replicas of the global network interact with their respective environments. Each local deep neural network (DNN), which includes both the actor and the critic, is trained independently (as shown on the right side of Figure 4). During training, the local DNNs do not update their own parameters directly but instead modify the global model. The local models synchronize with the global DNN after several steps. This setup allows the global network to be refined through the independent, concurrent training of local models, thereby accelerating the training process.

A3C's multithreaded implementation also greatly boosts learning efficiency. However, A3C struggles in complex environments due to its fixed learning rate, resulting in less robust performance.

To tackle this problem, *DPPO (Distributed Proximal Policy Optimization)* was introduced [49], as shown in Figure 5. DPPO incorporates a penalty term, which mitigates the impact of an inappropriate learning rate by ensuring a more balanced update proportion.

The algorithm features a global network alongside multiple local networks and employs a centralized learning approach combined with decentralized execution. The global network updates the actor and critic parameters, while the local networks gather sample data. During training, the local networks interact independently and simultaneously with the environment based on the global network's strategy.

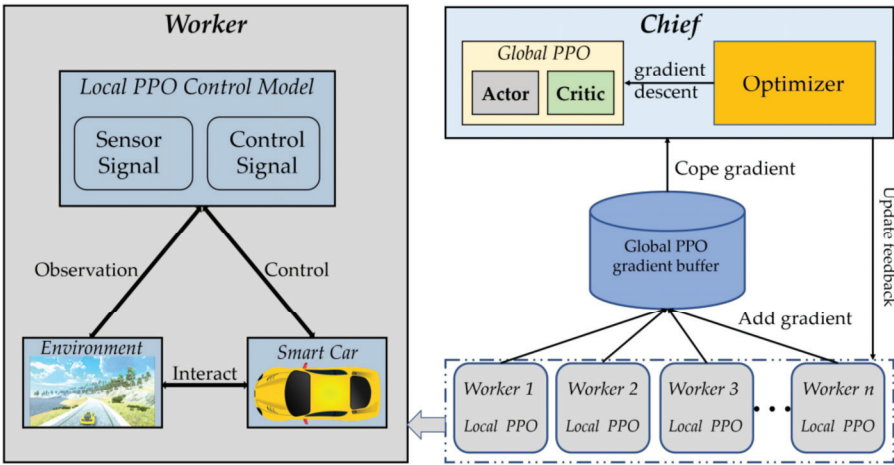


Figure 5. Schematic of the DPPO model [49]

They collect data, compute the policy gradient and store it in a shared gradient area. Once this shared area accumulates enough data, the global network retrieves gradient information for learning and updates its parameters. Subsequently, the updated policy parameters are shared with the local networks, which continue to gather data in their respective environments and repeat the process until the maximum number of training steps is reached. This setup involves multiple agents in parallel training within the simulation environment, allowing them to communicate and exchange feedback, thus significantly enhancing training efficiency and addressing issues like slow convergence.

3.3. Large Language Models and their Applications in Robotics

3.3.1 The working principle of LLM

A Large Language Model (LLM) is a foundational system designed to comprehend, interpret, and generate text in human language [50]. It accomplishes this by analyzing datasets and recognizing patterns and grammatical structures within the data to create text in a conversational manner. LLMs consist of numerous layers and contain millions or even billions of parameters. They are trained on extensive amounts of data, which allows them to understand complex relationships between words and predict the next word in a sentence.

These models utilize self-supervised learning, continuously processing the data until achieving a high level of accuracy. The performance of a language model is heavily influenced by the quality of its training data. The specific steps of the LLM model are shown in Figure 6.

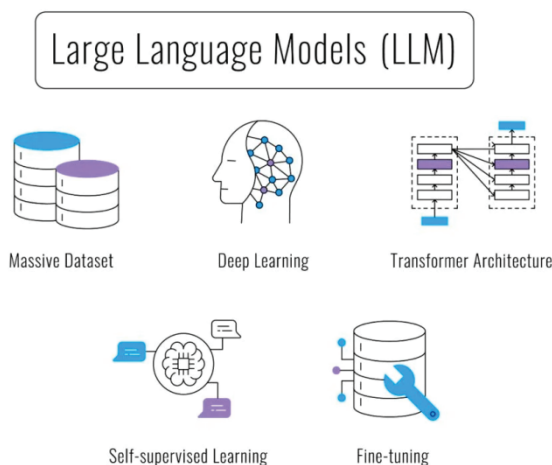


Figure 6. The custom steps of the LLM structure [50]

The first step in creating a large language model is to determine the type of LLM you intend to develop. This process starts with gathering a vast and diverse dataset of text from multiple sources, which serves as the foundation for training the model.

After collection, the text data undergoes *preprocessing*, which involves tasks such as *tokenization* (breaking text into words or subwords), converting text to lowercase, removing punctuation, and encoding the text into numerical formats suitable for machine learning. During this phase, each token (word or subword) is converted into a vector representation known as an *embedding*. Embeddings capture the semantic information of words, allowing the model to understand and learn the relationships between them.

Large Language Models generally use neural network architectures known as *transformer architectures*, which employ self-attention mechanisms to capture relationships between words irrespective of their positions in the input sequence. Since transformer architectures do not inherently account for the order of words, *positional encodings* are added to provide information about each token's position in the sequence, enabling the model to understand the sequential structure of the text.

Training a large language model (LLM) involves feeding sequences of tokens into the model and adjusting its parameters to minimize the difference between predicted and actual next tokens. These models are trained on extensive datasets over numerous epochs, progressively enhancing their performance. After the initial training phase, *fine-tuning* may be performed on more specific tasks or domains to customize the model for particular applications.

3.3.2 Examples of LLMs

To date, numerous foundational models or LLMs have been developed. Notable examples include BERT, Roberta, GPT-3, GPT-4, LLaMA, OPT (an open-source pre-trained transformer language model from Meta), Falcon 2 (an open-source LLM and VLM), Bloom (an open-source LLM from BigScience), and Mistral-7B. These models represent significant advancements in the field of artificial intelligence (AI) in recent years.

- *GPT*– OpenAI's Generative Pretrained Transformer (GPT) is perhaps the most renowned large language model (LLM). Among these, GPT-3 stands out with its 175 billion parameters, enabling it to generate coherent and contextually relevant text across a wide range of domains, including translation, question-answering, and cloze tasks. Additionally, it excels in tasks requiring real-time reasoning or domain adaptation, such as unscrambling words, using novel words in sentences, or performing three-digit arithmetic [52]. GPT-4(V), released on March 14, 2023, represents a significant advancement in the evolution of language

models. A notable innovation in GPT-4 is its multimodal capabilities, allowing it to process images as input (GPT-4V) and generate detailed descriptions, classifications, and analyses across various media types. This multimodal functionality expands the model's versatility and enhances its ability to understand and create content across diverse formats. GPT-4 Turbo and GPT-4(V) extended the context window from 32K to 128K. The latest OpenAI version GPT-4o (o for omni) processes and generates output across text, audio and image modalities in real time.

- *BERT* (Bidirectional Encoder Representations from Transformers) – Developed by Google, BERT introduced a novel approach of pre-training and fine-tuning for language understanding tasks. This innovation has resulted in exceptional performance in applications such as question answering and text classification.
- *T5* (The Text-to-Text Transfer Transformer) – Also developed by Google, T5 is a versatile language model that can be fine-tuned for a wide array of natural language processing tasks, including summarization, translation, and text generation.
- *LaMDA* – Developed by Google, LaMDA powers Google's conversational chatbot, Bard, enhancing its ability to engage in more natural and meaningful conversations.
- *PaLM-2* (successor of LaMDA), LLM designed to handle complex language tasks across different languages and contexts. PaLM is known for its large scale, extensive training data, and ability to perform a variety of NLP tasks. It is used on Microsoft BARD Chat bot.
- *LLaMA*– Developed by Meta, LLaMA is a generative text model pre-trained and fine-tuned with parameter counts ranging from 7 to 70 billion. It eliminates the absolute position embedding and introduces rotational position embedding at each layer of the network. The version, LLaMA 3, was integrated into Meta AI in 2024.

Applying large language models to the field of robotics has significant research implications and practical value [53]. By utilizing pre-trained language models, robots can better comprehend user intentions and needs. Additionally, other research focuses on employing LLMs for natural language generation in robots. From various perspectives, large language models-based robotics is one of the most promising pathways to achieving embodied intelligence in the future.

However, integrating LLMs with robotics presents several challenges. Training and deploying LLMs require significant computing resources and data, which can be a limitation for resource-constrained robotic platforms.

Generally, the applications of LLM in robotics hold tremendous potential. LLM models provide new paradigms and methods for robot control, perception, decision-making, and path planning. Below is a summary of LLM-based robotics [53]:

- *PaLM-SayCan* is an advanced system developed by Google, introduced in 2022. PaLM-SayCan is a sophisticated system that combines the Pathways Language Model (PaLM) with the SayCan framework to enable robots to execute tasks based on natural language instructions. It exemplifies the potential of LLMs to enhance human-robot interactions by translating complex language inputs into actionable tasks, supported by a value function to guide decision-making.
- *PaLM-E* (Pathways Language Model for Embodied AI) [54], introduced in 2023, is an *advanced multimodal AI* model developed by Google that integrates language (PaLM) and vision (ViT) capabilities for embodied AI applications. By extending the PaLM model's language processing abilities into the realm of robotics, PaLM-E enables sophisticated interaction with the physical world through natural language and visual inputs. It can perform tasks such as object manipulation, navigation, and interactive responses based on both visual and textual data. Largest model (2024) is PaLM-E-526B with 526B parameters.
- *PaLI-X* is an *advanced multimodal model* developed by Google, introduced in 2023. It represents a significance in the integration of language and vision capabilities, expanding the scope of tasks that large language models (LLMs) can perform across different modalities. By combining advanced language and vision capabilities, PaLI-X enables new types of interactions between humans and machines, and supports a wide range of practical applications.

LLM and LVM (Large Vision Model) run towards each other, leading to the new field of Multimodal Large Language Model (MLLM). Formally, it refers to the LLM-based model with the ability to receive, reason, and output with multimodal information [18]. Figure 7. shows the MLLM timeline.

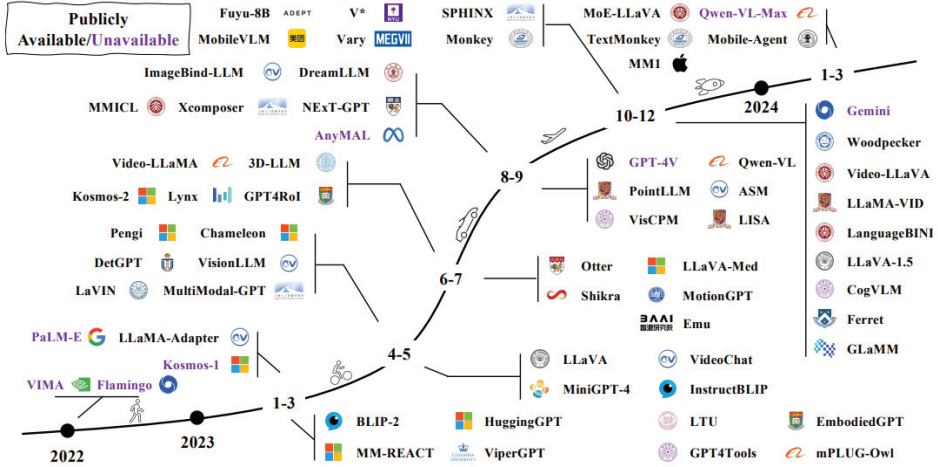


Figure 7. A timeline of representative MLLMs [18]

3.4. Transformer Models: Foundations and Their Role in Robotics

3.4.1 Transformer Models Architecture

Natural Language Processing (NLP) tasks can be categorized into discriminative tasks, and generative tasks. *Discriminative tasks* are used for sentiment analysis, text classification, for example if the sentence is given, one can classify the sentiment of the sentence (like positive/negative). *Generative tasks* are used for language modeling, machine translation, summarization, like where one needs to predict the next word based on the input sentence (context).

In the pre-transformer era (before 2017 [5]), standard ways to solve those tasks were Recurrent Neural Networks (RNNs), Long short-term memory (LSTM), and Convolutional neural networks(CNNs).

Each of those models have some problems. Problems faced with RNNs and LTMs are dependency across hidden layers, where the current hidden state in the network depends on the previous hidden states and hard to model long-term relationships, because languages may not have locality, unlike images where it takes $O(\text{sequence length})$ steps to model the interaction between two tokens. This limits the training parallelism.

For CNNs there are no dependency problems between tokens, which leads to better scalability, but there is limited context information, because of fixed window size of convolution that results in worse modeling capability.

Transformers are a type of model introduced by Vaswani et al. in their paper titled "Attention Is All You Need." [5]. They marked a significant shift from previous architectures like RNNs and LSTMs, which were previously dominant in sequence modeling tasks. Transformer models have introduced significant

advancements in language modeling, overcoming previous challenges with several key features:

- *Attention Mechanisms*: Transformers employ self-attention mechanisms to evaluate the significance of each word in a sentence in relation to the others.
- *Parallelization*: Unlike RNNs and LSTMs, transformers process all words in a sequence simultaneously, allowing for faster training and better scalability.
- *Scalability*: Transformers can be scaled up with more layers and parameters to handle complex tasks and large datasets.

The transformer architecture consists of two main parts: the *encoder* and the *decoder*, Figure 7. The encoder, situated on the left side of the Transformer architecture, is responsible for mapping an input sequence to a series of continuous representations, which are subsequently passed to the decoder. The decoder, located on the right side of the architecture, generates an output sequence by using the encoder's output in conjunction with the decoder's output from the previous time step. Both components are composed of multiple layers, and each layer contains specific subcomponents.

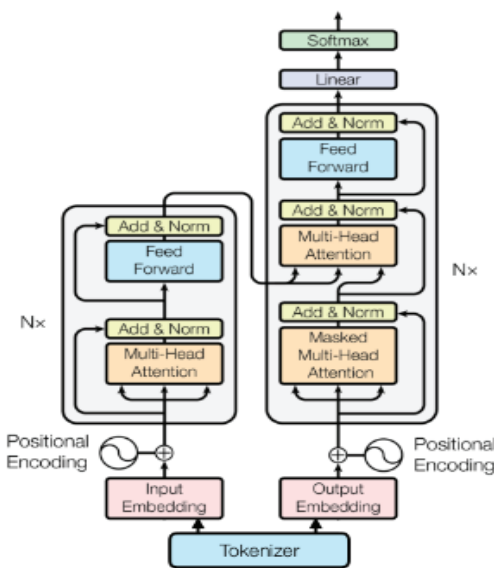


Figure 8. The encoder-decoder structure of the Transformer architecture [5]

The self-attention mechanism is crucial for a transformer's capacity to process sequences. Multiple self-attention heads operate in parallel, each capturing different aspects of the relationships between words. Since transformers do not inherently recognize the order of tokens, *positional encodings* are incorporated to convey information about the positions of tokens within the sequence.

3.4.2 Key Applications of Transformer Models in Robotics

Transformer models in robotics enhance perception, decision-making, and interaction capabilities, enabling robots to perform complex tasks with greater efficiency and accuracy.

Transformers improve a robot's ability to perceive and understand its environment:

Visual Recognition: Vision transformers (ViTs) apply transformer architectures to image data, enhancing object detection, classification and scene understanding.

Multimodal Perception: Models like PaLM-SayCan integrate visual and textual data, allowing robots to interpret and respond to visual cues combined with natural language instructions.

Transformers enhance the planning and navigation capabilities of robots:

- *Path Planning:* Transformers can process image data to help robots plan their movements by predicting the outcomes of different action sequences. This capability is crucial for autonomous robots that need to navigate dynamic environments.
- *Task Planning:* By understanding high-level goals and breaking them down into actionable steps, transformers assist robots in efficiently completing tasks.
- *Motion Prediction:* By analyzing sequences of past actions, transformers enable robots to anticipate future states and make informed decisions about their next moves.

The Control Transformer (CT) is a Transformer framework designed to model conditional sequences produced by robot actions [5,53]. It approaches the CT problem as a sequence modeling challenge with a goal-oriented perspective, allowing for learning from data collections obtained through sampling. Essentially, CT processes involve auto-regressively predicting actions within a sequence.

Decision Transformer (DT) for offline RL takes a sequence of returns, observations, and actions as input and outputs action predictions [55].

Trajectory Transformer (TT) also models the trajectory as a sequence of states, actions and rewards, while discretizing each dimension of state/actions [AH55].

Generalized DT unifies a family of algorithms for future information matching using transformers [55]. Transformers can also be used as world models for model-based RL [55,56].

The Q-Transformer [57] is designed to combine offline reinforcement learning with the Transformer architecture [5], facilitating the use of Q-values for each action dimension. This method involves discretizing each action dimension and representing them as distinct tokens using Q-values. By leveraging large and

diverse robot datasets, this approach improves the efficiency and effectiveness of the reinforcement learning process. Tested on real world experiments with dataset of 58k demonstrations on more than 700 tasks, with a fleet of 13 robots.

Robotics Transformer 1 (RT-1) [53,58] is designed to encode high-dimensional input and output data, such as images and instructions, into compact tokens that can be efficiently handled by Transformer [5]. It features real-time operation capabilities, making it well-suited for applications that demand quick processing and response times.

RT-1 showed impressive generalization capabilities in experimental evaluations. Its architecture includes FiLM, a conditioned EfficientNet, a TokenLearner, and a Transformer. Tested on 700+tasks (130k demonstrations/episodes), with fleet of 13 robots over 17 months, 35M parameters and compared with BC-Z and Gato (Generalist agent) models. Despite its advanced features, RT-1 is not an end-to-end model.

Robot Transformer 2 (RT-2) [59] is a model that utilizes the fine-tuning of a vision-language model (VLM) to develop an end-to-end system capable of directly mapping robot observations to actions. Research demonstrated how text-encoded 6 DoF actions from vision-language alignments (VLAs) are integrated into the robot's closed-loop control.

RT-2 is trained on a web-scale dataset to enable generalization and semantic awareness for new tasks. Specifically, it uses the WebLI dataset, which includes 1 billion image-text pairs across 109 languages and incorporates low-level action-related text tokens such as Cartesian end-effector commands. This model can be classified as a visual-language-action model (VLA) [59]. The largest version, RT-2-PaLI-X-55B, has 55 billion parameters and operates at 1-3 Hz.

Robotics Transformer X (RT-X) [60] is divided into two branches: RT-1-X and RT-2-X. RT-1-X utilizes the RT-1 architecture and is trained using the X-embodiment repository, while RT-2-X builds on the strategy architecture of RT-2 and is trained on the same dataset. Experiments have shown that both RT-1-X and RT-2-X demonstrate improved capabilities.

There are some notable examples of transformers in robotics:

- **PaLM-SayCan.** PaLM-SayCan integrates transformer models to allow robots to process natural language instructions and perform corresponding physical tasks.
- **LM-Nav, 2023:** Developed to enhance communication between users and robots using language, the LM-Nav system includes three key components: a vision-navigation model (VNM), a vision language model (VLM), and a large language model (LLM).
- **Expedition A1, 2023:** Developed by AGIBot, Expedition A1 showcases the company's commitment to integrating advanced AI into robotics and fostering seamless collaboration between humans and machines.

3.5. Comparison of Advanced Transformers Models in Robotics

The rapid evolution of artificial intelligence has profoundly impacted robotics, with transformer models emerging as a transformative force in the field.

This comparative analysis, presented in Table 1, explores the application of advanced transformer models in robotics, focusing on their performance, efficiency, and suitability for diverse robotic tasks.

Table 1. *Comparison of Advanced Transformer Models in Robotics*

Model	Developer	Release Year	Core Functionality	Primary Focus	Key Features	Applications
PaLM-SayCan	Google	2022	Combines NLP with robotic control	Task execution based on natural language	Multimodal capabilities, semantic understanding, physical embodiment	Robots interpreting and executing human commands in real-world environments
PaLM-E	Google	2023	Multimodal integration for robotic tasks	Enhanced perception and versatility	Enhanced perception, integrated embeddings, versatility	Situationally aware robots capable of complex and varied tasks
PaLM-X	Google	2024	Cross-modal language model	Multimodal understanding and generation	Advanced cross-modal integration, extensive knowledge base	General-purpose multimodal understanding and task execution
Q-transformer	OpenAI	2023	Integrates Q-learning with transformer architecture	Reinforcement learning and decision making	Combines Q-learning with transformers, optimized for sequential decision-making tasks	Complex decision-making in dynamic environments
Decision Transformer	MIT	2021	Combines sequence modeling with reinforcement learning	Reinforcement learning	Uses transformers for RL, sequence modeling, flexible policy generation	Decision-making in complex tasks, reinforcement learning
RT-1	Google DeepMind	2022	Real-time learning and adaptation for robotic control	Real-time learning	Reinforcement learning, scalable architecture, efficient training	Autonomous robots learning and adapting in real-time
RT-2	Google DeepMind	2023	Utilizes reinforcement learning within a transformer framework	Efficient learning and control	Reinforcement learning, scalable architecture, optimized for quick adaptation	Enhanced robotic performance in dynamic environments
RT-X	Google DeepMind	2024	Next-generation robotics transformer	Advanced real-time learning and control	Cutting-edge transformer architecture, superior reinforcement learning capabilities	Sophisticated robotic applications requiring high adaptability

ViT (Vision Transformer)	Google	2020	Image classification and processing	Visual tasks and perception	Transformer architecture applied to visual data, attention mechanisms	Image recognition, object detection, visual understanding
--------------------------	--------	------	-------------------------------------	-----------------------------	---	---

4. Examples of AI Methods in Robotics

AI methods in robotics encompass a range of techniques, from machine learning to reinforcement learning and natural language processing. Transformers and LLMs have revolutionized the field of robotics by enhancing robots capabilities in understanding, processing, and generating human-like language.

Below are some notable examples of how deep learning, reinforcement learning, transformers and LLMs are being applied in robotics to improve perception, interaction, and decision-making.

4.1. Computer Vision in Applications

Convolutional Neural Networks (CNNs) have transformed image recognition by learning spatial hierarchies from input images, making them ideal for tasks such as image classification and object detection. They are widely used in robotics for applications such as sorting and quality control. Vision Transformers (ViTs) and the DEtection TRansformer (DETR) represent newer approaches, with ViTs excelling in capturing long-range dependencies and DETR showing promising results in object detection. YOLO algorithms enable real-time object detection essential for autonomous navigation, while segmentation techniques offer detailed pixel-wise classification, crucial for applications like robotic surgery.

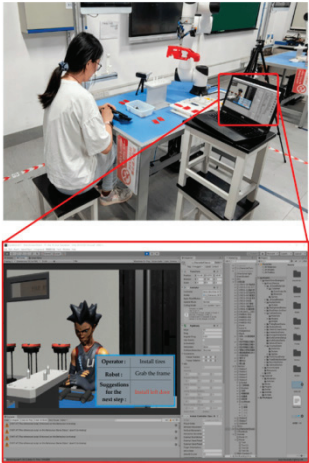


Figure 9. Computer vision based on CNN neural network in robotics [61]

Example of CNN-based computer vision used in human-robot collaboration in assembly tasks is presented in Figure 9 [61], while computer vision based on VLA model is presented in Figure 10 [59].



Figure 10. Computer vision based on VLA model in robotics [59]

4.2. DRL in Applications

Deep Reinforcement Learning (DRL) has been increasingly utilized to enhance human-robot collaboration across various industrial applications: *pick-and-place tasks, safe interaction in industrial settings, dynamic task allocation, assembly line collaboration, etc.* [62].

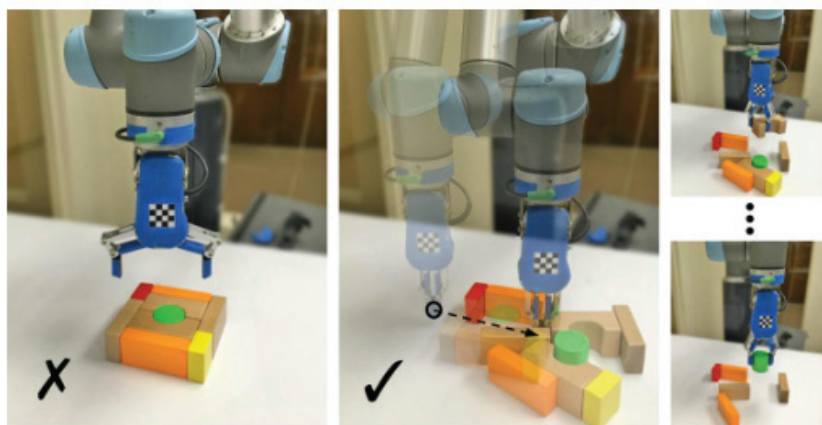


Figure 11. Pick and place tasks [63]

DRL is used to control collaborative robots (cobots) in pick-and-place tasks, allowing them to adapt to dynamic environments and handle objects that were not part of their initial training. This is achieved by integrating vision systems and DRL algorithms to improve object recognition and manipulation capabilities. As seen in Figure 11, KUKA developed, together with Roboception, a 3D vision system to efficiently perform bin picking using artificial intelligence [63].

In manufacturing, DRL algorithms optimize robot movements to ensure safe and efficient interaction with human workers. This involves path planning and collision avoidance, leveraging DRL to dynamically adjust the robot's actions based on the worker's movements and task requirements. With the use of DRL methods, robots are able to track the movement of humans and shape their path, in order to avoid any collisions, as seen in Figure 12 [64].



Figure 12. Safe interaction in industrial settings [64]

DRL helps in dynamically allocating tasks between humans and robots, improving efficiency and reducing the cognitive load on human workers. For instance, in flexible manufacturing systems, DRL algorithms can learn optimal task distributions based on real-time data, enhancing the overall workflow [62]. DRL is used in assembly line settings where robots and humans work side by side. Robots equipped with DRL can learn to perform precise tasks such as fastening or welding, adapting their actions based on the real-time feedback from human coworkers and environmental changes. When DRL is used with robots in welding processes, the quality of the seam can be quite improved due to the ability to adapt the welding parameters throughout the process, as seen in Figure 13, as well as enhance safety due to a number of sensors able to respond to human presence [65].

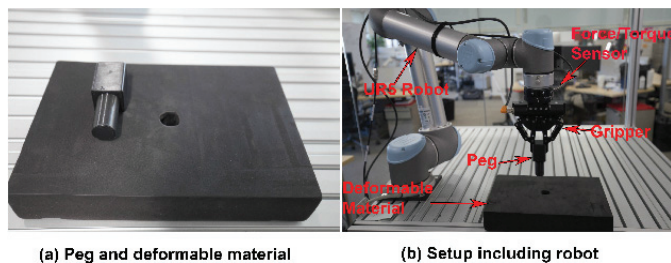


Figure 13. Assembly line collaboration [65]

These applications highlight how DRL can significantly enhance the adaptability, safety, and efficiency of human-robot collaboration, leading to more integrated and productive work environments.

4.3. Transformers in Applications

Traditional neural networks have shown limitations due to their lack of flexibility in executing actions based on inputs that do not match the training parameters, meaning they struggle to adapt to both minor and major changes in the system. Transformer models have revolutionized addressing this issue, primarily in the field of Natural Language Processing (NLP), where they have demonstrated excellent performance. Increasingly, transformer models are also being developed to solve various problems in robotics.

Task execution based on inputs. Transformer models combine computer vision and natural language processing to be used in a form where the robotic system performs specific tasks based on an input, usually in the form of a question. The robot takes the questions expressed in natural language and based on the images it gets, processes the task and performs the correct output [58]. Figure 14 shows such implementation, where the robotic arm using the RT-1 transformer model executes an instruction given by the operator.

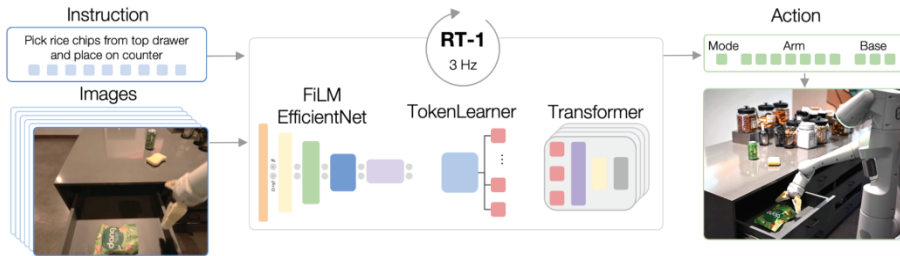


Figure 14. Task Execution based on input [58]

Motion Planning. Previously used methods in motion planning, when used in goal-targeted tasks, usually show inefficiency if they lack the time or pretraining. Transformer models are well-suited for solving planning tasks due to their capability to make long-horizon connections [66]. Additionally, they take large planning spaces and split them into discrete sets and carefully choose sampling regions, which allows the generation on near-optimal paths while reducing the planning time. The Vector Quantized-Motion Planning Transformer (VQ-MPT) is a transformer-based model that employs vector quantization to discretize the planning space into a set of distributions. These possibilities are illustrated in Figure 15 and Figure 16.



Figure 15. Motion path planning using Transformer model [66]

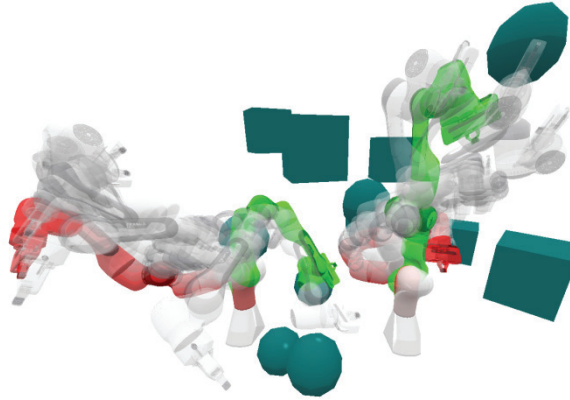


Figure 16. Sample paths planned by Vector Quantized-Motion Planning Transformer [66]

Adaptive Systems. Standard neural networks are trained on a specific dataset, creating a response matrix that lacks the ability to adapt to changes in the environment. However, adaptability is a desirable trait in artificial intelligence processes because it increases network flexibility and enables the solution of complex tasks. AdaTape exemplifies the application of Transformer models in adaptive systems by using an Adaptive Tape Reading mechanism to select candidates from a token bank. This selection process can involve direct picking from an input-driven token bank or additional computation based on potential candidates from a learnable token bank. [67]

Learning from Demonstration. Learning from Demonstration (LfD) is a process where a robot observes and records human behavior to mimic it. Human behavior can be unpredictable, complex, and sometimes unclear, posing challenges for traditional neural networks. However, Transformer models can address these issues due to their distinct learning approach. Transformers approach LfD by incorporating a large number of inputs, such as images from various angles, which the robot can analyze, break down into smaller segments, and use to determine the corresponding output actions [68]. One example learning from demonstration is presented in Figure 17 [69].

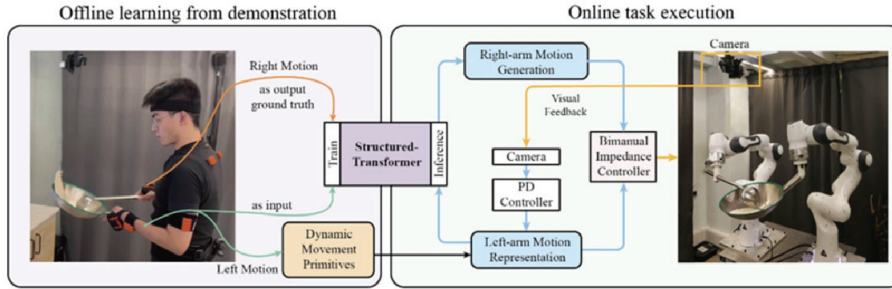


Figure 17. Learning from Demonstration [69]

4.4. LLMs in Applications

LLMs are increasingly being integrated into robotic applications to enhance their capabilities. By leveraging advanced natural language processing, For instance, LLMs can facilitate more intuitive human-robot communication, allowing robots to comprehend natural language instructions and respond in a contextually relevant manner. Additionally, LLMs can be employed to analyze large datasets for task planning, decision-making, and adaptive learning, ultimately leading to more flexible and intelligent robotic systems.

Enhanced Human-Robot Interaction. Robotics, although one of the most used modern technologies, unfortunately are unable to perform any actions outside of the ones it was programmed for. When used in human-robot interaction, the lack of understanding of the surrounding, or worse the human, whom it is working with, can result even in a dangerous interaction. Joining LLMs and robotics makes quite an impact on enhancing the HRI by having the best of both sides combined in different tasks [70]. Effective human-robot interaction is based on constant feedback and intention sharing between the two sides to perform actions that work best for the given environment, as shown in Figure 18.

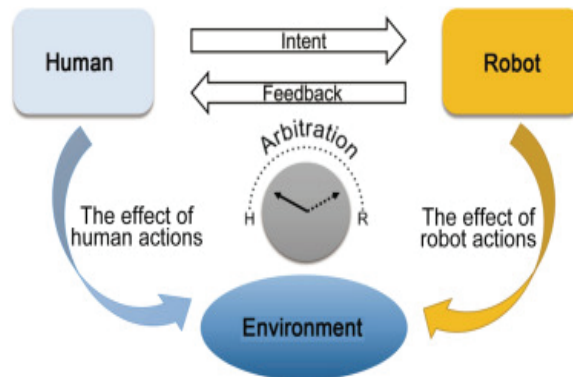


Figure 18. Enhanced Human-Robot Interaction [70]

Programming Assistance. Robot programming requires deep knowledge in both robotic movement and kinematics and programming languages, and thus learning to program robots can take years of gathering experience. LLMs can help in assisting new operators to quickly re-program certain paths or create totally new programs by “translating” human language commands, either written or verbally, into understandable robot action.

RoboDK has presented its Virtual Assistant which can be used in such tasks, while saving time and also help learn robotics in another way [71]. Figure 19 shows Robot Assistance in practice, where based on given commands, the robot transforms them into particular actions that can be performed [72].

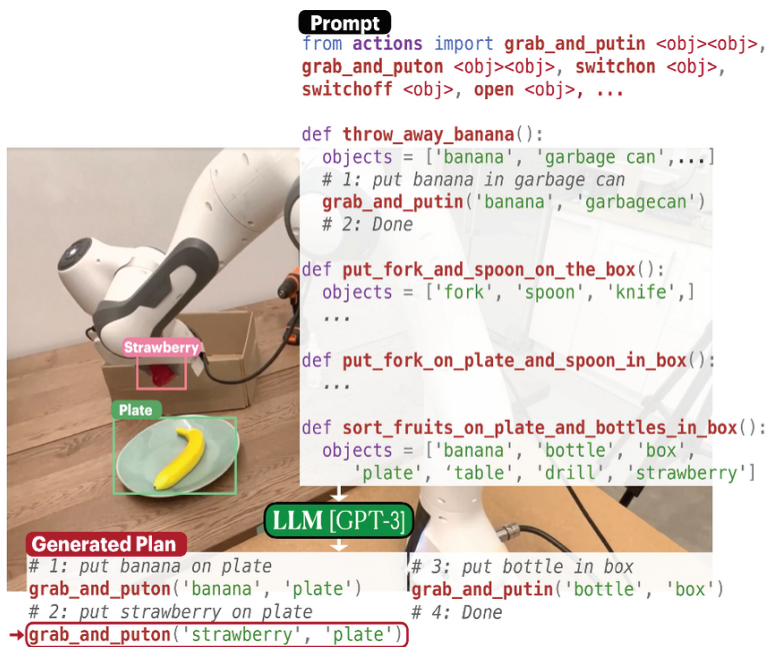


Figure 19. Programming Assistance [72]

Ethical Reasoning. Making decisions based on clear, logical situations is not an issue to any properly trained neural network, but when ethics and emotions come in the way, certain problems can arise. Robots with LLMs have shown their use and effectiveness in decision making that can be ambiguous for the robot itself and demand compassion, but with proper emotional weight management and various circumstances, robots made decisions based on emotion rather than on logic. [73,74]. To explain furthermore ethical reasoning and the steps that the robot can go through while dealing with an ethical problem, Figure 20 is used for additional explanation.

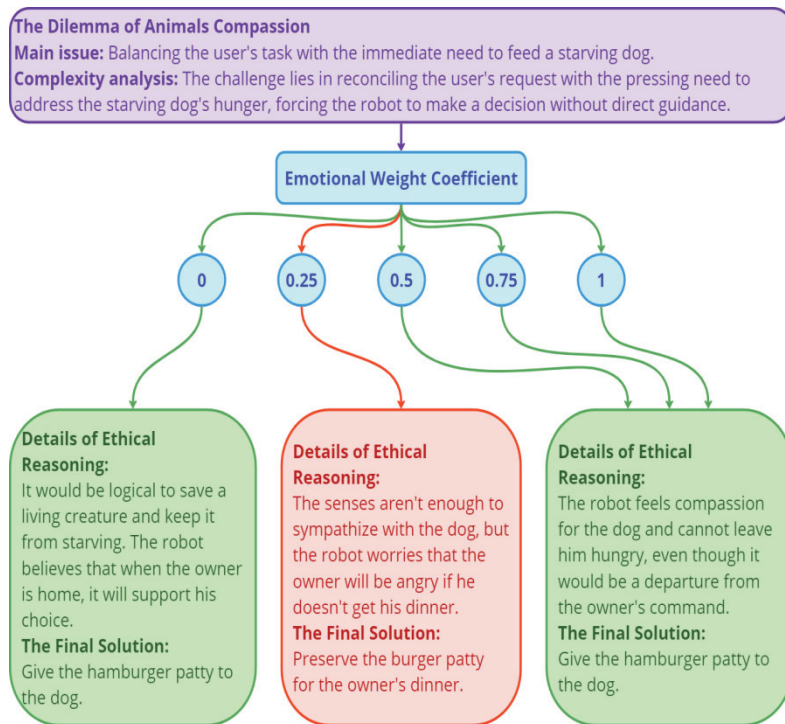


Figure 20. Ethical Reasoning [73]

5. Conclusion

The integration of computer vision, Deep Reinforcement Learning (DRL), Transformers, and Large Language Models (LLMs) in robotics has significantly advanced the field, enhancing the capability, autonomy, and versatility of robotic systems. Computer vision, powered by deep learning, allows robots to perceive and interpret their environment with remarkable accuracy, facilitating tasks such as object recognition, navigation, and manipulation. DRL has shown promise in enabling robots to learn complex tasks through trial and error, achieving superhuman performance in specific scenarios.

Transformers, initially designed for natural language processing, have been successfully adapted for various vision tasks, leading to the development of Vision Transformers (ViTs) which excel in image classification, object detection, and segmentation. These models have provided a more scalable and efficient alternative to traditional convolutional neural networks (CNNs).

LLMs have opened new horizons in human-robot interaction, allowing robots to understand and generate human-like text, facilitating more natural and intuitive

communication. This capability is crucial for developing robots that can assist in education, healthcare, and customer service.

Looking ahead, the convergence of these technologies offers a plethora of opportunities for the future of robotics. Key future directives include:

1. **Enhanced Human-Robot Collaboration:** Developing more sophisticated models for understanding human intentions and emotions, improving the synergy between humans and robots in various applications.
2. **Autonomous Decision-Making:** Advancing DRL techniques to allow robots to make real-time decisions in dynamic and unstructured environments, essential for applications such as autonomous driving and disaster response.
3. **Multimodal Learning:** Integrating vision, language, and other sensory inputs to create more holistic and context-aware robotic systems, enabling them to perform complex tasks with minimal supervision.
4. **Ethical and Safe AI:** Ensuring the development of ethical guidelines and safety protocols to prevent misuse and ensure that robotic systems operate reliably and responsibly in human-centric environments.
5. **Scalable Deployment:** Reducing the computational and energy requirements of these advanced models to facilitate their deployment in various real-world applications, including those with limited resources.

In conclusion, the amalgamation of DRL, computer vision, transformers, and LLMs is revolutionizing robotics, pushing the boundaries of what robots can achieve. Continued research and development in these areas will pave the way for more intelligent, capable, and human-friendly robots, profoundly impacting industries and daily life.

6. Acknowledgment

The authors express their sincere thanks for the funding support they received from Federal Ministry of Education and Science of Bosnia and Herzegovina (Grant for project: "Research and Development of Collaborative Intelligence in Service Robots for Industrial Applications", 2023-2024).

7. References

- [1] LeCun, Y., Bengio, Y., Hinton, G., *Deep learning*. Nature, 521(7553), 436-444, 2015. <https://doi.org/10.1038/nature14539>
- [2] Sutton, R. S., Barto, A. G., *Reinforcement Learning: An Introduction*. (2nd ed.). MIT Press, 2018 <http://incompleteideas.net/book/the-book-2nd.html>
- [3] Chen, M., Tworek, J., Jun, H., Yuan, Q., de Oliveira Pinto, H. P., Kaplan, J., ... Zaremba, W., *Evaluating large language models trained on code*, 2021 <https://arxiv.org/abs/2107.03374>
- [4] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D., *Language models are few-shot learners*. Advances in Neural Information Processing Systems, 33, 1877-1901, 2020 <https://arxiv.org/abs/2005.14165>
- [5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I., *Attention is all you need*. Advances in Neural Information Processing Systems, 30, 5998-6008, 2017 <https://arxiv.org/abs/1706.03762>
- [6] Chen, T., Moreaux, Z., Lee, K., Xu, Z., *Multi-modal transformer for robotics trajectory learning*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1643-1648, 2021 <https://arxiv.org/abs/2104.07176>
- [7] Tsai, Y. H. H., Bai, S., Liang, P. P., Kolter, J. Z., Morency, L. P., Salakhutdinov, R., *Multimodal transformer for unaligned multimodal language sequences*. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 6558-6569, 2019 <https://arxiv.org/abs/1906.00295>
- [8] Jain, A., Lample, G., Denoyer, L., *Multimodal state estimation for robust deep reinforcement learning of vision-based robotic manipulation tasks*. Conference on Robot Learning (CoRL), 147-156, 2019 <https://arxiv.org/abs/1911.08265>
- [9] Thomason, J., Gordon, D., Bisk, Y., Jett, J., Mihaylova, T., Zettlemoyer, L., *Improving grounded natural language understanding through human-robot dialog*. IEEE Transactions on Robotics, 36(3), 667-684, 2020 <https://arxiv.org/abs/1910.03581>
- [10] Fang, M., Zhao, W., Li, T., Xie, X., Yin, H., Lyu, L., *Multi-source information fusion for autonomous driving: A survey*. IEEE Transactions on Intelligent Transportation Systems, 2022 <https://arxiv.org/abs/2201.08311>
- [11] Siciliano, B., Khatib, O. (Eds.), *Springer Handbook of Robotics*. Springer, (2016) <https://doi.org/10.1007/978-3-319-32552-1>
- [12] LeCun, Y., Bengio, Y., Hinton, G., *Deep learning*. Nature, 521(7553), 436-444m, 2015 <https://doi.org/10.1038/nature14539>

- [13] Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., Shah, M., *Transformers in vision: A survey*. ACM Computing Surveys (CSUR), 54(10), 1-41, 2021 <https://arxiv.org/abs/2101.01169>
- [14] Banjanović-Mehmedović, L., *Artificial Intelligence Advancement in Service Robots Applications*. In Book: Service Robots: Advances in Research and Application (Editors: I. Karabegović, L. Banjanović-Mehmedović), Nova Science Publisher, USA, 2021
- [15] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... Houlsby, N., *An image is worth 16x16 words: Transformers for image recognition at scale*. International Conference on Learning Representations (ICLR), 2021, <https://arxiv.org/abs/2010.11929>
- [16] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S., *End-to-End Object Detection with Transformers*. European Conference on Computer Vision, 213-229, 2020 doi:10.48550/arXiv.2005.12872.
- [17] Ma, Y., Song, Z., Zhuang, Y., Hao, J., King, I., *A Survey on Vision-Language-Action Models for Embodied AI*, 2023 <https://arxiv.org/abs/2303.14280>
- [18] Shukang Y, Chaoyou F, Sirui Z., Ke L., Xing S, Tong X, Enhong C., *A Survey on Multimodal Large Language Models*. IEEE Transactions on pattern analysis and machine intelligence, 2023 <https://arxiv.org/pdf/2306.13549>
- [19] Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Chen, X., Choromanski, K., ... Zitkovich, B., *Rt-2: Vision-language-action models transfer web knowledge to robotic control*. Google DeepMind, 2023.
- [20] Gu, S., Holly, E., Lillicrap, T., Levine, S., *Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates*. 2017 IEEE International Conference on Robotics and Automation (ICRA), 3389-3396, 2017 doi:10.1109/ICRA.2017.7989385.
- [21] Korivand, S., Galvani, G., Ajoudani, A., Gong, J., Jalili, N., *Optimizing Human-Robot Teaming Performance through Q-Learning-Based Task Load Adjustment and Physiological Data Analysis*. Sensors 2024, 24, 2817, 2024 <https://doi.org/10.3390/s24092817>
- [22] Khatib, O., Laumond, J. P., Siciliano, B., *Robot Motion Planning and Control*. Springer Science & Business Media, 2008 doi:10.1007/978-1-4471-4005-4.
- [23] Balatti, P., Ozdamar, I., Sirintuna, D., Fortini, L., Leonori, M., Gandarias, JM., Ajoudani, A., *Robot-Assisted Navigation for Visually Impaired through Adaptive Impedance and Path Planning*. IEEE International Conference on Robotics and Automation (ICRA), 2023.

- [24] Banjanovic-Mehmedovic, L., Karabegovic, I., Jahic, J., Omercic, M., *Optimal path planning of a disinfection mobile robot against COVID-19 in a ROS-based research platform*. Advances in Production Engineering & Management, Volume 16, Number 4, December 2021, pp 405-417 <https://doi.org/10.14743/apem2021.4.409>
- [25] Husaković, A., Banjanović-Mehmedović, L., Konjić, T., *Efficiency Boost: Service Robot Path Planning with Grey Wolf Optimization*. Proceedings of IEEE 23rd International Symposium INFOTEH-JAHORINA, Bosnia and Herzegovina, 2024.
- [26] What are Large Language Models (LLMs) and how will they be used in 2024?, <https://www.dataquest.io/blog/what-are-large-language-models-llms-and-how-will-they-be-used-in-2024/>, [Accessed: 16.7.2024]
- [27] Khandelwal, P., Zhang, S., Mazumder, A., Zhang, P., Mehta, R., Stone, P., *Multimodal Sensor Fusion using Deep Learning for Autonomous Indoor Robots*. AAAI Conference on Artificial Intelligence, 2017
- [28] Goodrich, M. A., Schultz, A. C., *Human-robot interaction: a survey*. Foundations and Trends in Human-Computer Interaction, 1(3), 203-275, 2007 <http://dx.doi.org/10.1561/1100000005>
- [29] Alibegović, B., Prljača, N., Kimmel, M., Schultalbers, M., *Speech recognition system for a service robot – a performance evaluation*. The 16th International Conference on Control, Automation, Robotics and Vision ICARCV 2020, Shenzhen, China, 2020 DOI: 10.1109/ICARCV50220.2020.9305342
- [30] Cao, H., Elprama, S.A., Scholz, C., Siahaya, P.L., Makrini, I.E., Jacobs, A., Ajoudani, A., Vanderborght, B., *Designing interaction interface for supportive human-robot collaboration: A co-creation study involving factory employees*. Comput. Ind. Eng., 192, 110208, 2024
- [31] Russell, S., Norvig, P., *Artificial Intelligence: A Modern Approach*, 2020
- [32] LeCun, Y., Bengio, Y., Hinton, G., *Deep learning*. Nature, 521(7553), 436-444, 2015 doi:10.1038/nature14539
- [33] Banjanovic-Mehmedovic, L., Gurdić, A., *Object Classification in an Intelligent Robotic Cell using Deep Learning*. In I. Karabegovic (Ed.): New Technologies, Development and Application IV (NT 2021), LNNS, Springer, pp. 101–112, 2021
- [34] Redmon, J., Farhadi, A., *YOLOv3: An incremental improvement*, 2018 arXiv preprint arXiv:1804.02767.
- [35] Hodžić, M., Prljača, N., *Missile Guidance using Proportional Navigation and Machine Learning*. Special Issue on Computing, Engineering and Sciences, JENRS, 2024 DOI: <https://dx.doi.org/10.55708/js0303003>
- [36] Long, J., Shelhamer, E., Darrell, T., *Fully convolutional networks for semantic segmentation*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431-3440, 2015

- [37] He, K., Gkioxari, G., Dollár, P., Girshick, R., *Mask R-CNN*. In Proceedings of the IEEE international conference on computer vision pp. 2961-2969, 2017
- [38] Qi, C. R., Su, H., Mo, K., Guibas, L. J., *PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 652-660, 2017
- [39] Zhou, Y., Tuzel, O., *VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4490-4499, 2018
- [40] Cai, Q., Cui, C., Xiong, Y., Wang, W., Xie, Z., Zhang, M., *A Survey on Deep Reinforcement Learning for Data Processing and Analytics*. IEEE Transactions On Knowledge And Data Engineering, Vol. 35, No. 5, 2023
- [41] Panzer, M., Bender, B., *Deep reinforcement learning in production systems: a systematic literature review*, International Journal of Production Research, 60:13, 4316-4341, 2022 DOI: 10.1080/00207543.2021.1973138
- [42] Part 2: Kinds of RL Algorithms.
https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html
[Accessed. 12.7.2024.]
- [43] Wang, Y., Friderikos, V., A Survey of Deep Learning for Data Caching in Edge Network, Informatics 7(4):43, 2020 DOI: 10.3390/informatics7040043
- [44] Chen, X., Zhang, H., Wu, C., Mao, S., Ji, Y., Bennis, M., *Optimized Computation Offloading Performance in Virtual Edge Computing Systems Via Deep Reinforcement Learning*. In IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4005-4018, 2019, doi: 10.1109/IIOT.2018.2876279.
- [45] Pathmakumr, P., Elara, R., Gómez, B., Ramalingam, B A., *Reinforcement Learning Based Dirt-Exploration for Cleaning-Auditing Robot*. Sensors2021, 21, 8331.
- [46] Tai, L., Zhang, J., Liu, M., *A survey of deep reinforcement learning for robotic motion control*, 2017 arXiv preprint arXiv:1610.00696.
- [47] Wang, X., Xie, J., Guo, S., Li, Y., Sun, P., Gan, Z., *Deep reinforcement learning-based rehabilitation robot trajectory planning with optimized reward functions*. Advances in Mechanical Engineering. 2021;13(12). doi:10.1177/16878140211067011
- [48] Mnih, V., Badia, A. P., Mirza, M., Graves, A., Harley, T., Lillicrap, T. P., ... Kavukcuoglu, K., *Asynchronous Methods for Deep Reinforcement Learning*. In Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016

- [49] Lin, J., Zhang, P., Li, C., Zhou, Y., Wang, H., Zou, X., *APF-DPPO: An Automatic Driving Policy Learning Method Based on the Artificial Potential Field Method to Optimize the Reward Function*. Machines 2022, 10, 533. <https://doi.org/10.3390/machines10070533>
- [50] What is a Large Language Model (LLM)? Definition, Examples, Use Cases, <https://em360tech.com/tech-article/large-language-model>, [Accessed: 16.7.2024]
- [51] Zeng, F., Gan, W., Wang, Y., Liu, N., Yu, P. S., *Large language models for robotics: A survey*, 2023 arXiv preprint arXiv:2311.07226.
- [52] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D., *Language Models are Few-Shot Learners*, 2020 <https://arxiv.org/abs/2005.14165>
- [53] Zeng, F., Gan, W., Wang, Y., Liu, N., Lu, P. S., *Large Language Models for Robotics: A Survey*. 2023 <https://arxiv.org/abs/2304.12109>
- [54] Smith, A. M., Johnson, B. L., Lee, C. K., *PaLM-E: Pathways Language Model for Embodied AI*, 2023 Google AI Research.
- [55] Sun, Y., Ma, S., Madaan, R., Bonatti, R., Huang, F., Kapoor, A., *SMART: Self-supervised Multi-task pretraining with control Transformers*. Proceedings of the 11th International Conference on Learning Representations (ICLR 2023), 2023
- [56] Chen, C., Wu, Y. F., Yoon, J., Ahn, S., *Transdreamer: Reinforcement learning with transformer world models*, 2022 arXiv preprint arXiv:2202.09481.
- [57] Chebotar, Y., Vuong, Q., Irpan, A., Hausman, K., Xia, F., Lu, Y., Kumar, A., Yu, T., Herzog, A., Pertsch, K., et al., *QTransformer: Scalable offline reinforcement learning via autoregressive q-functions*, 2023 arXiv preprint, arXiv:2309.10150 .
- [58] Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Dabis, J., Finn, C., Gopalakrishnan, K., Hausman, K., Herzog, A., Hsu, J., et al., *RT-1: Robotics transformer for real-world control at scale*. Robotics: Science and Systems XIX, 2023
- [59] Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Choromanski, K., Ding, T., Driess, D., Dubey, K.A., Finn, C., Florence, P.R., Fu, C., Arenas, M.G., Gopalakrishnan, K., Han, K., Hausman, K., Herzog, A., Hsu, J., Ichter, B., Irpan, A., Joshi, N., Julian, R.C., Kalashnikov, D., Kuang, Y., Leal, I., Levine, S., Michalewski, H., Mordatch, I., Pertsch, K., Rao, K., Reymann, K., Ryoo, M.S., Salazar, G., Sanketi, P.R., Sermanet, P., Singh, J., Singh, A., Soricut, R., Tran, H., Vanhoucke, V., Vuong, Q.H., Wahid, A., Welker, S., Wohlhart, P., Xiao, T., Yu, T., Zitkovich, B., *RT-2: Vision-Language-*

- Action Models Transfer Web Knowledge to Robotic Control*. 2023. ArXiv, abs/2307.15818.
- [60] O'Neill, A. et al. *Open X-Embodiment: Robotic learning datasets and RT-X models*. DeepMind, 2023 <https://robotics-transformer-x.github.io/paper.pdf>.
- [61] Gao, Z., Yang, R., Zhao, K., Yu, W., Liu, Z., Liu, L., *Hybrid Convolutional Neural Network Approaches for Recognizing Collaborative Actions in Human–Robot Assembly Tasks*. Sustainability 2024, 16, 139. <https://doi.org/10.3390/su16010139>
- [62] Gomes N.M., Martins F.N., Lima J., Wortche H., *Reinforcement Learning for Collaborative Robots Pick-and-Place Applications: A Case Study*, Automation, 3 (1) , pp. 223-241, 2022
- [63] Zeng, A., Song, S., Welker, S., Lee, J., Rodriguez, A., Funkhouser, T., *Learning synergies between pushing and grasping with self-supervised deep reinforcement learning*. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4238-4245, 2018 <https://arxiv.org/abs/1803.09956>.
- [64] Liu, Q., Liu, Z., Xiong, B., Xu, W., Liu, Y., *Deep reinforcement learning-based safe interaction for industrial human-robot collaboration using intrinsic reward function*. Advanced Engineering Informatics. 49, 2021 <https://doi.org/10.1016/j.aei.2021.101360>
- [65] Luo, J., Solowjow, E., Wen, C., Aparicio Ojea, J., M. Agonino, A., *Deep Reinforcement Learning for Robotic Assembly of Mixed Deformable and Rigid Objects*. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2062-2069. 2018. <https://api.semanticscholar.org/CorpusID:57755511>
- [66] Johnson, J. J., Qureshi, A. H., Yip, M. C., *Learning sampling dictionaries for efficient and generalizable robot motion planning with transformers*. IEEE Robotics and Automation Letters, 2023
- [67] Xue, F., Likhoshervstov, V., Arnab, A., Houlsby, N., Dehghani, M., You, Y., *Adaptive computation with elastic input sequence*. In International Conference on Machine Learning (pp. 38971-38988). PMLR, 2023 <https://arxiv.org/abs/2301.13195>
- [68] Tianci, G., *Transformer-XL for Long Sequence Tasks in Robotic Learning from Demonstration*, 2024 <https://arxiv.org/abs/2405.15562>
- [69] Liu, J. et al., *Robot Cooking With Stir-Fry: Bimanual Non-Prehensile Manipulation of Semi-Fluid Objects*. In IEEE Robotics and Automation Letters, vol. 7, no. 2, pp.5159-5166. 2022, doi: 10.1109/LRA.2022.3153728.
- [70] Ceng, Z., Junxin, C., Jiatong, L., Yanhong, P., Zebing, M., *Large language models for human–robot interaction: A review*, Biomimetic Intelligence and Robotics, Volume 3, Issue 4, 2023, 100131, ISSN 2667-3797, <https://doi.org/10.1016/j.birob.2023.100131>

- [71] Chen, J. T., Huang, C. M., *Forgetful large language models: Lessons learned from using LLMS in robot programming*. In Proceedings of the AAAI Symposium Series, Vol. 2, No. 1, pp. 508-513, 2023<https://arxiv.org/abs/2310.06646>
- [72] Singh, I., Blukis, V., Mousavian, A., Goyal, A., Xu, D., Tremblay, J., ... Garg, A., *Progprompt: Generating situated robot task plans using large language models*. In 2023 IEEE International Conference on Robotics and Automation (ICRA) pp. 11523-11530 2023<https://arxiv.org/abs/2209.11302>
- [73] Lykov, A., Cabrera, M. A., Gbagbe, K. F., Tsetserukou, D., *Robots Can Feel: LLM-based Framework for Robot Ethical Reasoning*, 2024 arXiv preprint arXiv:2405.05824.,<https://arxiv.org/abs/2405.05824>.
- [74] Markelius, A., *An Empirical Design Justice Approach to Identifying Ethical Considerations in the Intersection of Large Language Models and Social Robotics*. 2024, <https://arxiv.org/html/2406.06400v2>

Use of Artificial Intelligence in Traffic Technology and Transport

Sadko Mandžuka^{*1}, Edouard Ivanjko¹, Krešimir Vidović²

Abstract: *This article presents the application of various artificial intelligence technologies in the field of intelligent transport systems. Examples of research into the application of expert systems based on ANFIS technology in motorway traffic safety control and urban mobility assessment are described. In addition, the results of scientific research on the application of learning agents in traffic control are presented.*

Keywords: *Intelligent Transport Systems, Artificial Intelligence, Expert Systems, ANFIS, Intelligent mobility, Traffic control, Learning agents*

1. Introduction

The use of artificial intelligence (AI) in the field of Intelligent Transport Systems (ITS) has become increasingly important. Today's traffic and transport systems generate vast amounts of data in real time, where AI algorithms offer innovative solutions to improve traffic flow, reduce congestion, and increase overall transport efficiency, [1, 2, 3]. In addition, one of the key advantages of artificial intelligence is its ability to analyze real-time traffic data from various sources, such as cameras, sensors, GPS devices, etc. By leveraging this data, AI algorithms can identify behavioral patterns and process trends. These real-time analyses enable traffic management systems to make decisions and implement dynamic management strategies to improve traffic flow, especially in the most congested areas, [4]. This includes adaptive traffic control systems as another application of artificial intelligence. These systems use AI algorithms to adjust the timing of control signals based on current traffic conditions. This proactive approach to traffic control has proven to be very effective in reducing congestion and improving the overall driving experience. In addition to real-time data analysis and adaptive signal control, predictive analytics algorithms are also used to predict traffic patterns and identify potential critical areas of congestion, [5, 6]. By analyzing historical data and trends, these algorithms can predict future traffic conditions, allowing proactive measures to be taken to alleviate congestion before it occurs. By using effective AI algorithms to analyze data in

^{*1}Faculty of Traffic and Transport Sciences, Zagreb

²Ericsson Nikola Tesla, Zagreb

E-mail: smandzuka@fpz.unizg.hr, eivanjko@fpz.unizg.hr, kresimir.vidovic@ericssonnikolatesla.com

real time, optimize traffic flow and predict traffic patterns, it enables more sustainable traffic systems.

The Expert System is branch of artificial intelligence (AI) which is utilized to address complex problems by utilizing human knowledge on various areas for analysis and decision making. By definition, an expert system is a computer program that manipulates facts, knowledge, and reasoning to solve problems efficiently and effectively in a narrow problem area that normally requires various expensive sources of knowledge and human experts. Expert systems operate similar to human experts, which use symbolic logic and heuristic rules to propose result, and they are characterized by some advantages compared to humans.

This paper presents two applications of expert systems based on ANFIS architecture: Decision support system for motorway safety management and Expert system for urban mobility estimation,[7, 8].ANFIS architecture is based on fine-tuning of fuzzy causal IF-THEN rules and enables work with large data sets. The advantage of ANFIS is that it combines fuzzy decision-making ability with the learning ability provided by the neural network to model the dynamics of different nonlinear systems (universal approximation), [9, 10].ANFIS usually is created as a five-layer network of multilayer perceptron (MLP), Fig. 1.

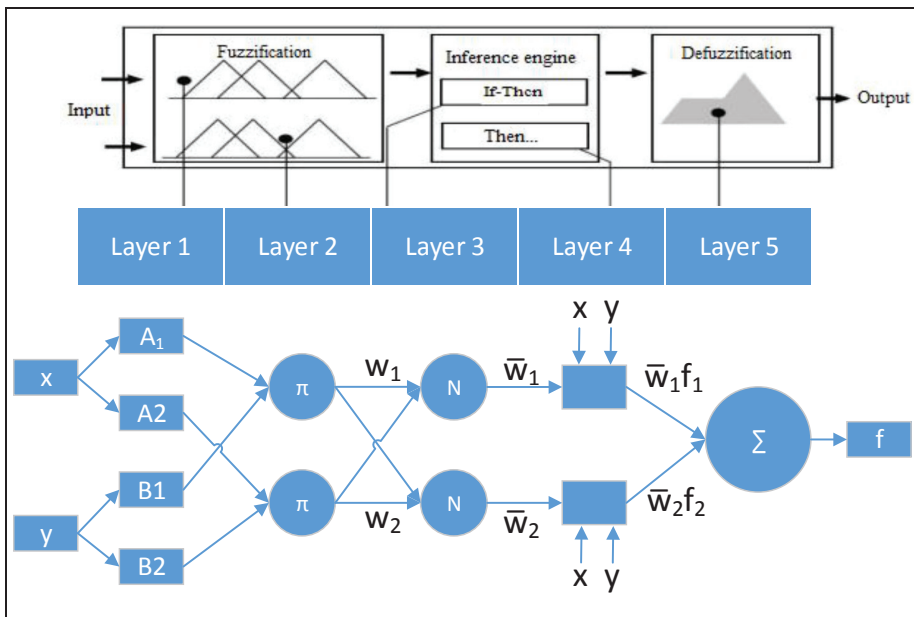


Figure 1. ANFIS network architecture [7]

One area where AI is applied very successfully is traffic control. It is an important service in the scope of Intelligent Transportation Systems (ITS). Today, traffic control must solve the problem related to everyday congestion, and significant short or long term changes in traffic demand. Especially in urban environments where urban motorways [11] and signalized intersections [12] are the two most important use cases for ITS traffic control services. An often applied learning approach to obtain a control law capable of solving significant changes in traffic demand is Reinforcement Learning (RL). The concept of learning agents and their application in traffic control on urban motorways and for signalized intersections is elaborated in forth chapter.

Generally, an agent in AI is a structure that can learn from previous experience and improve its performance during operation. Usually, it is applied as software connected with sensors to measure the state of its working environment containing internal logic to create output via actuators to change the state of the working environment according to a defined performance criteria function. There exists a class of cooperative agents that have communication capabilities to exchange information with other (neighboring) agents to improve global performance when dealing with larger complex systems like urban transportation networks.

2. Use of Expert Systems to Motorway Safety Management

In this approach, the fundamental part of the real-time motorway safety management system is ANFIS, serving as a pivotal machine learning technology. Leveraging predefined algorithms and assimilated rules derived from the Crash potential model [7], the system generates recommendations for optimizing motorway signalizations. Essentially, there are two primary systems influencing traffic: variable message signs (VMS) and the implementation of suitable variable speed limits (VSL) signs. With its capacity for learning, the system adapts and applies the most effective solutions based on past experiences.

The architectural proposal for the decision support system in motorway safety management is depicted in Fig. 2 [7]. Thorough testing on simulation models and selected trials is imperative for the future implementation of these technologies. In this context, analyzing driver behavior regarding adherence to speed limits is particularly crucial, serving as a fundamental prerequisite for the successful adoption of this motorway safety management approach.

During the training phase, ANFIS learns from a dataset by adjusting the parameters of the fuzzy IF-THEN rules and the neural network. This is typically done using a technique like backpropagation or gradient descent to minimize the error between the actual and predicted outputs. Once trained, ANFIS can make predictions or decisions based on new input data. It uses the fuzzy IF-THEN

rules to fuzzify the input variables and propagate this information through the neural network to produce crisp output values. One of the key strengths of ANFIS is its ability to adapt and refine its rules and parameters based on new data or changes in the environment. This adaptability allows it to continuously improve its performance over time.

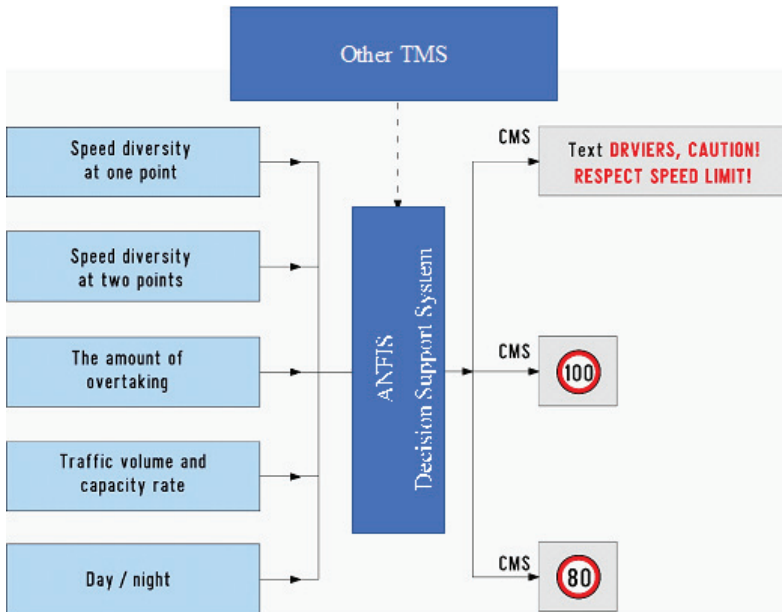


Figure 2. Decision support system architecture [7]

The true benefits of the approach outlined in this paper can be fully realized through cooperative traffic management systems [5]. By employing the cooperative intelligent vehicle speed adaptation system, issues related to driver compliance with speed limits are effectively addressed and mitigated.

3. Expert System for Urban Mobility Assessment Based on the Application of Artificial Intelligence Technologies

Urban mobility is defined as the movement of people between sources and destinations, at various times, by different means of transport and modes of travel for achieving different goals.[13] The issue with urban mobility estimation is that currently proposed methods are characterized by number of shortcoming, since each research area (city) has its own characteristic that has to be taken into account. Therefore, since the mobility assessment should be based on objectively measured parameters, it requires expert knowledge that can quantify the assessment value regarding the state of urban mobility based on measured

parameters. Without embedded expert knowledge, the system itself is incapable of provision of conclusion if the state of the mobility in observed area is satisfactory and by which grade it can be assessed. The goal of the expert system is to be applicable on as large number of cities as possible.

To build ANFIS based expert system for urban mobility assessment expert knowledge is required, that will be applied over objectively measured mobility indicators. Those mobility indicators have to be selected in the way that they are universally accessible and available, in order to mitigate the impact of diversity and variety of mobility data that may come from different sources. Therefore, the public mobile telecommunication network has been identified as primary data sources, since it's potential as data source for urban mobility related application has been proved by number of researchers[14]. Both telecom network Call data records (database storing records for billing purposes) and signaling data (communication data among different network elements, obtained by probes) have been used for anonymized population migration analysis, resulting in following indicators: origin destination matrixes data (representing the volume of migrations in a unit of time); distance matrix (average distance per user migration among two spatial units, represented by either Euclidian distance, or road distance od most likely itinerary obtained from external source); and travel time matrix (average travel time per user migration among two spatial units), [15]. In order to generate expert opinion, those parameters have to be fuzzified based on following principle. Indicator trip duration is divided into three categories: Short trip duration (up to 33% of the longest lasting trip), Medium trip duration (from 34% to 66% of the longest lasting trip) and Long trip duration (from 67% to 100% of the longest lasting trip). The same principle is applied to other indicators.

ANFIS process overview is presented in Fig. 3.

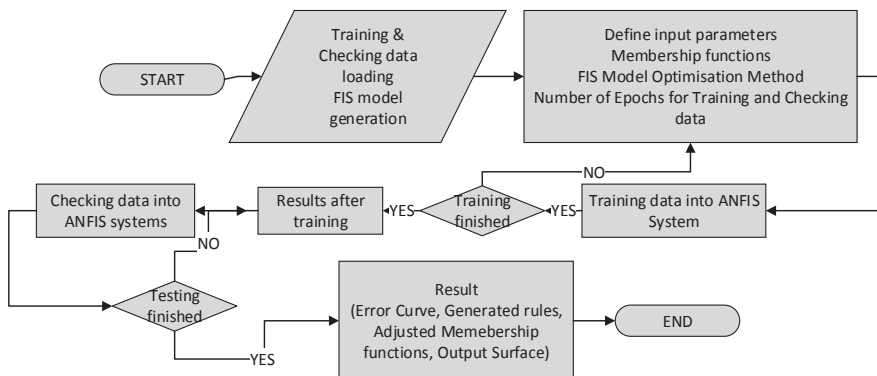


Figure 3. ANFIS model creation process, based on [9], [10]

The experts (urban mobility professionals, i.e., academics, experts from the private sector, city planners...) are then utilized and presented with a number of scenarios that they have to evaluate in terms of mobility assessment. Their feedback reflects the relationships between the fuzzified values of the mobility indicators and the mobility assessment which are the used for model training. The model is described with rules, each including respective interval of the mobility estimate output values, depending on the values of the input variables. Before the model learning initial rule base is set and the values of the target variable (mobility assessment values) are unknown. Following the model learning process (selection of model with lowest prediction error), ANFIS enables the creation and fine-tuning of rules which describe the behavior of a complex system and enables presentation of mobility assessment as a single number, considering the values of all indicators. Since it is based on generally accessible data source and universally applicable methodology, this AI based model is capable on calculation of urban mobility estimate grade that is comparable to itself (during different time periods) and to other observed areas (different cities).

4. Traffic Control Based on Learning Agents

4.1. Concept of Learning Agents

The underlying controlled process can be defined as a Markov Decision Process (MDP) with the tuple (S, A, T, R) , where S is the environment states set, A is the action set the controller can take, R is the reward function and T represents the transition function between states. In such a case, the RL concept can be applied with the Q-Learning (QL) algorithm as an important representative:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_{t+1} + \gamma \max_{a' \in A} Q(s_{t+1}, a') - Q(s_t, a_t)), \quad (1)$$

where $Q(s_t, a_t)$ is the Q value for a state-action pair in training iteration t , α is the learning rate at which the Q value is updated, r_{t+1} is the reward received from the environment after a particular action was selected and γ is the discount factor to include the impact of possible future rewards.

To train a QL-based traffic controller, all possible state-action pairs must be visited enough times to differentiate the best action to be taken in a particular state. This leads to convergence issues because, for a good controller, the state-action pair representation must be fine enough creating the curse of dimensionality.

4.2. Urban Motorways

On urban motorways two most common traffic control approaches are applied: Ramp Metering (RM) and Variable Speed Limit (VSL) [11]. In RM, vehicles from on-ramps are allowed to join the motorway main traffic flow if the occupancy or density near the on-ramp is under the critical value to postpone congestion build-up (left part of Fig. 4). Similarly, in VSL, the speed of vehicles coming to a motorway area with increased occupancy or density approaching the critical value is decreased to resolve or at least to postpone the congestion build-up (right part in Fig. 4).

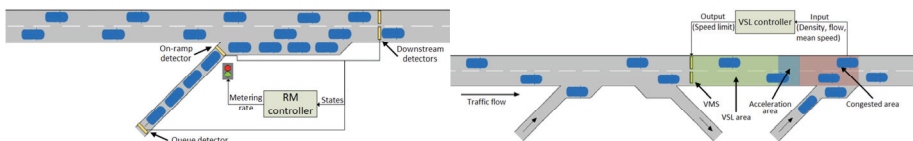


Figure 4. Control block schemes for RM (left) and VSL (right) [11]

To set up an agent for RM, the set of states should contain the notation about the on-ramp traffic light phase, average speed on motorway main traffic flow, and average on-ramp queue length. The action set should contain the possibility to stay in the same traffic light phase, change the phase, or turn RM off in the case of low traffic demand. The reward function should be linked to the queue length (should be minimized) and main traffic flow speed (should be maximized). When such a setup is applied on a simple motorway simulation model with one on-ramp, an improvement in the main traffic flow traveling time of around 10% can be expected. More details are available in [16].

To set up an agent for VSL, the sets for states and actions, and the reward function must also be appropriately defined. For VSL, also large oscillations of posted speed limits between control steps should be avoided. Thus, the state must include the past behavior of the main traffic flow and past actions. A good solution is to build the state vector containing two past actions (posted speed limits), and traffic density and speed in motorway areas affected by VSL. The action set should contain the allowed speed limits starting from the lowest to the highest. The reward function should positively emphasize a higher value of main traffic flow speed, minimize Total Time Spent (TTS), and penalize large oscillations of posted speed limits between consecutive control steps. With such a setup on a motorway simulation model with two on-ramps, one off-ramp and one VSL zone an improvement of 4% regarding average main traffic flow travel time including gradual change of speed limit can be expected. More details are available in [17].

4.3. Signalized Intersections

On signalized intersections, Adaptive Traffic Signal Control (ATSC) can be applied to alleviate congestion. The main idea of ATSC is to adapt the signal program according to the current traffic situation. Thus, RL must learn the control law for signal program adaptation. For intersection state description, queue lengths on all inbound roads can be used. The problem is the very large number of potential states in comparison to motorways negatively affecting the training convergence. However, the analysis done in [18] revealed that in the case of an independent intersection similar states can be grouped and the same signal program can be effectively used for the whole traffic states group. Thus, by using a Self-Organizing Map (SOM) the classic RL framework can be augmented improving training convergence (left part in Fig. 5), [18]. The drawback of such a SOM-based framework is the needed apriori off-line analysis of collected queue length measurements. By using the Growing Neural Gas, combined simultaneous training of the state-action representation and control law for ATSC can be achieved (right part in Fig. 5) [19]. Thus, the control system can be implemented more easily with improved training convergence compared to the classic QL algorithm implementation. For setting the reward function, the delay difference between two action executions was used. The GNG based approach has comparable convergence and control quality despite the simultaneous state-action representation and control law training. More details are available in [18, 19].

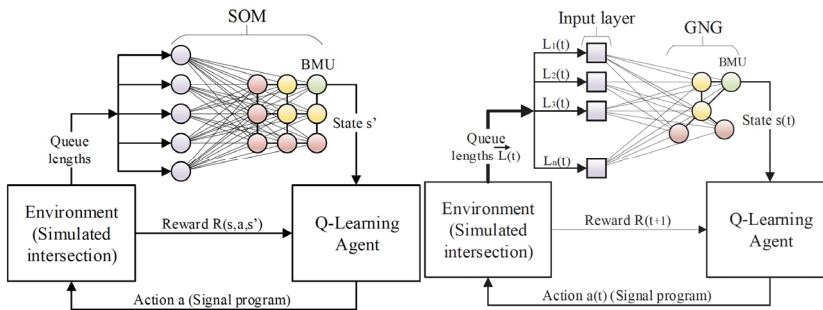


Figure 5. Applying SOM (left) and GNG (right) for QL based ATSC [18, 19]

5. Conclusion

One important application of AI in traffic technology and transport is solving the traffic control problem as a crucial ITS service. The application of expert systems as one of the most significant applications of artificial intelligence already plays a significant role in intelligent transport systems today. It has been

shown that the ANFIS architecture is very useful for these applications, both in real-time applications and in traffic planning applications.

Control structures based on learning agents can cope with significant daily changes in traffic demand and update their control laws during operation evaluating the deployment of such systems. The drawback is the needed accurate state-action representation complexity resulting with the curse of dimensionality which reduction is currently researched to improve training convergence.

Soon, a significant breakthrough in the application of AI in the field of traffic incident management is expected, [20]. Also, some new AI based applications in the area of ITS cyber security, especially for cooperative intelligent transport systems [21], are expected.

6. References

- [1] Abduljabbar, R., Dia, H., Liyanage, S., Bagloee, S. A., *Applications of artificial intelligence in transport: An overview*, Sustainability, 11(1), 189., 2019.
- [2] Miles, J. C., Walker, A. J., *The potential application of artificial intelligence in transport*, IEE Proceedings-Intelligent Transport Systems, Vol. 153, No. 3, pp. 183-198, 2006.
- [3] Mandžuka, S. *Intelligent transport systems - Selected Lectures*, Fakultet prometnih znanosti, Zagreb, 2015.
- [4] Bharadiya, J., *Artificial intelligence in transportation systems a critical review*, American Journal of Computing and Engineering, 6(1), 34-45., 2023.
- [5] Mandžuka, S., Ivanjko, E., Vujić, M., Škorput, P., Gregurić, M., *The use of cooperative ITS in urban traffic management*, Intell. Transp. Syst. Technol. Appl, 272-288., 2015.
- [6] Gregurić, M., Buntić, M., Ivanjko, E., Mandžuka, S., *Improvement of highway level of service using ramp metering*, Proceedings of the 21st international symposium on electronics in transport-ISEP, pp. 25-26, 2013.
- [7] Mandžuka, S., Dedić, L., Kos, G., Šoštarić, M., *Multicriteria decision support system for motorways safety management*, International Conference "New Technologies, Development and Applications" (pp. 624-630). Cham: Springer International Publishing, 2022.
- [8] A. A. Mohammed, K. Ambak, A. M. Mosa, and D. Syamsunur, "Expert system in engineering transportation: A review," J. Eng. Sci. Technol., vol. 14, no. 1, pp. 229–252, 2019.
- [9] J. S. R. Jang and C. T. Sun, "Neuro-Fuzzy Modeling and Control," Proc. IEEE, vol. 83, no. 3, pp. 378–406, 1995, doi: 10.1109/5.364486.

- [10] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," IEEE Trans. Syst. Man Cybern., vol. 23, no. 3, pp. 665–685, 1993, doi: 10.1109/21.256541.
- [11] Vrbanić, F., Ivanjko, E., Kušić, K., Čakija, D., *Variable Speed Limit and Ramp Metering for Mixed Traffic Flows: A Review and Open Questions*, Applied sciences (Basel), Vol. 11, No. 6, 2021.; <https://doi.org/10.3390/app11062574>
- [12] Miletić, M., Ivanjko, E., Gregurić, M., Kušić, K., *A review of reinforcement learning applications in adaptive traffic signal control*, IET Intelligent Transport Systems, Vol. 16, No. 10, pp. 1269-1285, 2022.; <https://doi.org/10.1049/itr2.12208>
- [13] K. Vidović, M. Šoštarić, S. Mandžuka, G. Kos, "Model for Estimating Urban Mobility Based on the Records of User Activities in Public Mobile Networks," Sustainability, vol. 12, no. 3, p. 838, 2020, doi: 10.3390/su12030838.
- [14] C. Chen, J. Ma, Y. Susilo, Y. Liu, M. Wang, "The promises of big data and small data for travel behavior (aka human mobility) analysis," Transp. Res. Part C, vol. 68, pp. 285–299, 2016, doi: 10.1016/j.trc.2016.04.005.
- [15] K. Vidović, M. Šoštarić, aD. Budimir, "An Overview of Indicators and Indices Used for Urban Mobility Assessment," PROMET - Traffic&Transportation, vol. 31, no. 6, pp. 703–714, 2019, doi: <https://doi.org/10.7307/ptt.v31i6.3281>.
- [16] Ivanjko, E., KoltovskaNečoska, D., Gregurić, M., Vujić, M., Jurković, G., Mandžuka, S., *Ramp metering control based on the Q-learning algorithm*, Cybernetics and Information Technologies, Vol. 15, No. 5. 2015, pp. 88-97.; <https://doi.org/10.1515/cait-2015-0019>
- [17] Kušić, K., Ivanjko, E., Gregurić, M., *A Comparison of Different State Representations for Reinforcement Learning Based Variable Speed Limit Control*, Proceedings of MED-2018, 2018. pp. 266-271.; <https://doi.org/10.1109/MED.2018.8442986>
- [18] Miletić, M., Kušić, K., Gregurić, M., Ivanjko, E., *State Complexity Reduction in Reinforcement Learning based Adaptive Traffic Signal Control*, Proceedings of ELMAR-2020, 2020. pp. 61-66.; <https://doi.org/10.1109/ELMAR49956.2020.9219024>
- [19] Miletić, M., Ivanjko, E., Mandžuka, S., Koltovska-Nečoska, D., *Combining Neural Gas and Reinforcement Learning for Adaptive Traffic Signal Control*, Proceedings of ELMAR-2021, 2021. pp. 179-182.; <https://doi.org/10.1109/ELMAR52657.2021.9550948>
- [20] Mandžuka, S., Kljaić, Z., Skorput, P., *The use of mobile communication in traffic incident management process*. Journal of Green Engineering, 4, 413-429, 2011.

- [21] Škorput, P., Vojvodić, H., & Mandžuka, S., *Cyber security in cooperative intelligent transportation systems*. In 2017 International Symposium ELMAR (pp. 35-38), 2017.

Acknowledgment

The presented research has been carried out within the activities of:

- a) Projects "*Proposal of a model of intelligent traffic safety management on a section of a multi-lane highway*" and "*Proposal of a model of a proactive messaging system for drivers*" approved as part of the National Road Traffic Safety Plan of the Republic of Croatia for 2021/30.
- b) Centre of Research Excellence for Data Science and Cooperative Systems supported by the Ministry of Science and Education of the Republic of Croatia. This work has been supported by the Croatian Science Foundation under the project IP-2020-02-5042.

Artificial Intelligence and its Application in Manufacturing

Vidosav D. Majstorović^{*1}, Sonja Dimitrijević², Vladimir Simeunović²,
Dragan Stošić²

Abstract: Artificial intelligence (AI) has become the most important element of the Industry 4.0 model today. It has wide application possibilities in the entire new value chain. Its history is about eight decades long, and a special area of its research and development is manufacturing, in which AI has been applied since the mid-1980s. Expert systems (ES) were the first AI tools applied in manufacturing. The goal of this paper is to perform a systematic analysis of the state of development and application of AI in manufacturing, which is originally used as an aid to the engineer, planner and designer of various engineering products. It is also used to manage processes and systems in mechanical engineering. Starting from that, the paper is structured in such a way as to provide answers to the following questions: what is AI and how did it develop, how were AI models created and how were they developed in technological systems, what are today's models and prospects for applying AI in them, as well as possible directions of future research in this area. As a special point of this paper, some results of our research in this area are presented.

Keywords: Artificial intelligence, Machine learning, Manufacturing systems, Manufacturing, Planning, Design, Management.

1. Introduction

Artificial intelligence (AI) is a branch of intelligence science, which broadly encompasses two areas: natural and artificial. Natural intelligence is the science of discovering the processes and models of intelligent behavior in living systems, while artificial intelligence, or AI, is both the science and engineering of creating intelligent software systems and machines. These two areas of research are connected and have contributed to each other during the past eight decades of development.

On the other hand, AI is the basis for the development and application of smart manufacturing. What is smart manufacturing? It represents the use of advanced technologies, while performing data analysis to improve manufacturing processes, increase efficiency and optimize manufacturing [1]. Also, it integrates

^{*1}Faculty of Mechanical Engineering, Belgrade

²"Mihajilo Pupin" Institute, Belgrade

E-mail: vidosav.majstorovic@sbb.rs, sonja.dimitrijevic@pupin.rs

various technologies: AI, Internet of Things (IoT), Big Data Analytics (BDA), robotics and cloud computing into the manufacturing environment.

The key features of smart manufacturing are: (a) system integration - Connecting different manufacturing systems and processes through IoT and other technologies to create a cohesive and synchronized manufacturing environment, for online operation, (b) big data analytics. Here, AI is used to analyze the huge amounts of data generated during manufacturing processes. This helps identify patterns of occurrence of events, predict maintenance needs, optimize manufacturing schedules and improve overall efficiency, (c) automation and robotics - the application of robotics and automation to perform tasks, which are repeated with high precision and reliability, thus reducing human error and increases productivity, (d) IoT and networking. Using IoT devices, data is collected in real time from machines, equipment, processes and products in the factory. This data is used for monitoring, management and decision-making, (e) advanced manufacturing technologies - are additive manufacturing (3D printing), digital twins (virtual models of physical assets) and augmented reality (AR) for design, simulation and maintenance purposes, (f) supply chain integration - improving supply chain management through real-time visibility and coordination, enabling faster response to market demands and shortening delivery times, and (g) energy efficiency and sustainability - optimizing energy use and resource consumption to reduce environmental impact and improving the sustainability of manufacturing.

Smart manufacturing is a new paradigm of technological systems, and its establishment and application in manufacturing is based on AI tools and techniques, which is especially evident in the Industry 4.0 model.

Smart manufacturing has several definitions, and one of the common ones is [2,3]: "*Smart manufacturing uses cognitive computing, industrial IoT and advanced analytics to optimize manufacturing processes in ways that were not possible before.*" *It helps organizations improve key business metrics such as productivity, delivery reliability, quality, business security and profitability, while reducing downtime and operating costs.*"

Smart manufacturing technologies look deeply into the manufacturing process and business environment like never before to extract information that has tangible value for the manufacturer.

The aim of this work is to systematically analyze models of development and application of AI in manufacturing, and it contains the following units: (i) historical development of the application of AI in manufacturing, (ii) detailed analysis of today's models of application of AI, (iii) some results of our research in this areas, (iv) conclusions and future research.

2. Historical Development of the Application of AI in Manufacturing

The history of the development of the application of AI in manufacturing is more than seventy years long, and includes five stages, table 1 [3-7].

Table 1. Historical overview of the development of AI in manufacturing (supplemented according to [3-7])

<i>Period of time</i>	<i>Characteristics</i>	<i>Elements of AI application in manufacturing</i>
Early AI development and application (1950s - 1970s) (level 1).	1950s: The term "artificial intelligence" (AI) was defined and the fundamental concepts of AI were developed.	-
	1960s: First AI research and development of models for problem solving and symbolic thinking.	Basic application of numerical control (NC) tools on control units (CDUs) of machines.
	1970s: Expert systems emerge as a prominent AI application. They use knowledge-based rules to solve specific problems.	They provide decision support in manufacturing processes.
Robotics and automation (1980 - 1990) (level 2).	1980s: Industrial robots began to be integrated into manufacturing processes, controlled by computer systems.	Early AI algorithms for tasks such as welding, painting and assembly.
	1990s: AI techniques such as machine learning began to be applied in manufacturing.	They are used for quality control, process optimization and predictive maintenance. There have also been advances in computer vision for automated inspection.
Knowledge-based systems (1990 - 2000) (level 3).	Expert systems have continued to evolve, helping to define and apply domain-specific knowledge in manufacturing contexts (knowledge engineering).	These systems provided extensive rule- and heuristic-based decision support (product design, process planning, and manufacturing management).
Data-driven AI (2000 to present) (level 4).	The 2000s marked a shift towards data-driven AI models, fueled by the availability of large data sets and advances in computing power.	
	Machine learning techniques such as neural networks, support vector methods, and decision trees have become increasingly popular for tasks such as predictive maintenance, anomaly detection, and optimization of manufacturing processes.	
	AI-based analytics and optimization tools have begun to integrate into manufacturing operations, enabling real-time insights and continuous improvement.	
Industry 4.0 and smart manufacturing (2010s to present) (level 5).	The concept of Industry 4.0 emphasizes the integration of AI, IoT, cloud computing and cyber-physical systems in manufacturing.	
	AI plays a key role in the development of smart factories, where machines and systems are interconnected, communicate and make decisions autonomously based on real-time data and analytics.	
	The application of AI includes: design and planning of products, processes and systems, predictive analytics for maintenance, adaptive manufacturing processes, autonomous robots and supply chain management.	

AI in manufacturing has evolved from early automation and ES, to sophisticated data-driven applications that increase productivity, quality and agility in manufacturing environments. Today's advances in AI technologies continue to shape the future of manufacturing, leading to more intelligent, efficient and adaptive manufacturing systems.

3. Overview of the Application of AI in Manufacturing - Literature Analysis

The last two decades of development and application of AI in manufacturing represent the "golden" age in this field. Thanks to the development of ICT and its application in manufacturing, they created the basis for extensive research and application of AI in manufacturing, as the following analysis shows.

This system analysis was performed according to the PRISMA methodology [7], and our questions were: Q1. What AI techniques and tools are being applied in manufacturing, and Q2. How AI improves decision-making, planning, analysis and management of manufacturing processes. The following were defined for them: time period of analysis, type of study, AI models, search methodology and assessment of study quality. The sample included 286 papers, and 35 papers met the set criteria. Broadly speaking, a systemic analysis of the application of AI in manufacturing is performed here from the following angles: (i) application of agents as an AI approach, (ii) concepts and models of application of AI in manufacturing, (iii) learning models, (iv) generative AI and its application, and (v) application of ChatGPT in manufacturing. The following analysis refers to all the mentioned areas.

3.1. Application of Agents as an AI Technique in Manufacturing

Agent-based distributed manufacturing management was a popular AI technique applied at the beginning of this century. In the traditional centralized manufacturing management systems of the time, the supervisor made decisions on manufacturing planning, scheduling and resource allocation. In contrast, in the agent model, decisions are distributed among several agents, who communicate with each other, in order to achieve optimal manufacturing results. Traditional optimization methods often consider only one domain (process planning or scheduling) and ignore the constraints of the other domain. This can lead to sub-optimal or even invalid manufacturing plans, due to the dynamic nature of real-time machine loads and plant conditions. That is why the application of AI agents is a much better solution. In this analysis, two characteristic overviews of approaches to the application of this AI technique are presented, Table 2.

Table 2. Overview of the use of agents

Area	Ref. / yr.	Method	The goal	Use
Manufacturing planning and scheduling	[8] 2006	Application of agents	Increasing the accuracy of manufacturing planning and scheduling	Management of manufacturing in the plant
Reconfigurable technological system	[9] 2009	A multi-agent system	Intelligent manufacturing management	Distributed manufacturing management

In agent-based manufacturing planning and scheduling systems, applied agent negotiation protocols require individual agents to respond to incoming bids. Therefore, in this model, it is extremely important that the knowledge base, reasoning and reasoning mechanism are developed [8]. Also, due to the autonomous and cooperative nature of the agents, planning and scheduling functions can be integrated from higher to lower levels. By applying the agent model, an innovative, agile and reconfigurable model technological system [9]. Thus, an intelligent distributed manufacturing management model was obtained. Finally, for this period of analysis, the first decade of this century, we can say that the agent model was the dominant approach to the application of AI in manufacturing.

3.2. Concepts and Models of Application of AI in Manufacturing

The second decade of this century saw a transformation of the world economy and economy, based on digitization and AI. In manufacturing, AI has produced major changes in the entire value chain, especially in: decision-making, product design, manufacturing planning and management, maintenance, quality control, while reducing downtime and manufacturing costs, increasing operational business parameters. This chapter analyzes in detail studies that have considered various aspects of the application of AI in manufacturing, at the national, regional or global level. This can help us to determine our situation, comparing it with them, and to define our policy in this area based on that. The first study [10] presented an analysis of the application of AI in 28 countries, Table 3, and the way it was defined was the GMI Summit, where dozens of Panels were held on

Table 3. Overview of studies on the application of AI (policies, national programs, good practice)

The way	Ref. / yr.	Method	The goal	Use
Global Manufacturing and Industrialization Summit (GMIS)	[10] 2020	The context of Industry 4.0	Guidelines for the application of AI	Defining the four levels of AI deployment
Foundation for autonomous manufacturing	[11] 2020	Autonomous manufacturing	Intelligent machines	In-house manufacturing and maintenance
Maturity of Industry 4.0	[12] 2020	Integration with factory functions	Model factory of the future	Wide application of deep learning models
AI as support for smart manufacturing	[13] 2020	Smart manufacturing	Improvement of the manufacturing model	National policy and strategy for the application of AI
Convergence of AI and manufacturing technologies	[14] in 2021	Modeling nonlinear AI problems with deep learning	Increase the efficiency and quality of manufacturing technologies	Application to different types of products

EU leader in the application of AI	[15] in 2021	In the Industry 4.0 model	Competitive advantage of EU manufacturing	High-tech factories
Deep learning models	[16] in 2022	CNN	Smart products	Smart manufacturing
Deep learning	[17] in 2022	CNN, RBM, RNN	Smart manufacturing	Maintenance and quality control
Analysis of ML techniques	[18] in 2023	Examples of good practice	Reduce application risks	Smart manufacturing
White paper - AI in EU manufacturing	[19] in 2023	Maturity model of AI in manufacturing	Assessment of EU companies for the application of AI	Definition of AI KPI parameters
Challenges in applying AI	[20] in 2023	Application project model	Successful application of AI	Six aspects that affect the success of an AI project in the factory
DT for the factory	[21] 2024	IDEFO model	Factory as DT	Applying AI through DT for the factory
Application of generative AI	[22] 2024	Encouraging producers	Policy of application of generative AI	Application in manufacturing

various aspects of AI not only applied in manufacturing. She also defined a multi-level model for the application of AI in manufacturing: (i) AI champions, completed digital transformation with a clear strategy for further application of AI, (ii) AI innovators, a clear digital transformation strategy, including AI, implemented in stages. The project enters full implementation, (iii) AI and digital companions, are organizations that have horizontal digital integration (sales, procurement, engineering, manufacturing), but not vertical integration (CAD, CAPP, CAM, MES, ERP). The same applies to elements of AI in application, and (iv) AI and digital first-timers are those organizations that have not yet entered the transformation project. Management has no strategy for this Project, their survival in the market is questionable. The document [11] of the Foundation for Autonomous Manufacturing starts from the premise that machines in manufacturing should be made intelligent for decision-making in relation to all issues of manufacturing and maintenance. This means that the Industry 4.0 model was established and applied, and supported by AI models, this manufacturing model is characterized as follows: (i) continuous optimization of the manufacturing process, with its online diagnostics, (ii) quick reaction to changes in manufacturing quality parameters, (iii)) close coordination of machines in the process, (iv) human interventions that can be a source of errors, remove, for example reprogramming, (v) improve the machine self-diagnosis process, and (vi) apply tighter control tolerance limits. AI represents the maturity of the application of Industry 4.0 in practice [12]. This paradigm is based on the following facts: (i) predictive maintenance. By applying deep machine learning (DML) and big data analysis (BDA) models, large savings are achieved and downtime is reduced, (ii) quality assurance and inspection. Using computer vision and the DML model, the analysis and recognition of errors on products is performed, after manufacturing operations and on the assembly line, (iii) optimization of supply chains. AI models can be effectively used here for inventory management, demand forecasting and

delivery scheduling, (iv) application of generative AI in engineering activities. Rapid prototyping and product improvement, (v) digital twins and simulation. AI models in this area reduce maintenance costs, predict failures and optimize energy consumption, and (vi) AI models for information security. Far Eastern countries (Japan, South Korea, China, Taiwan and India) adopted national programs for AI at the end of the last decade, which included its application in manufacturing [13]. These Programs defined the Smart Manufacturing model, as a framework for the AI-based factory model. At the moment, the application of AI in manufacturing covers a wide range of machine learning models, where the key to success is pattern recognition for highly nonlinear data, analysis of unstructured data, robustness of repetitive tasks, and high interoperability [14]. To illustrate the aforementioned facts, the following examples are given: autonomous vehicle control, assembly robot in the automotive industry, predictive maintenance of wind generators, manufacturing quality management in a steel plant, and semiconductor manufacturing quality management. All the above examples speak of the high potential of convergence of AI and various manufacturing problems that can be solved. The analyzes presented in [15] show that the EU applies more AI in manufacturing than the industrialized countries of the Far East, as well as North America, on average about 30%. These facts speak of the importance of the symbiosis of research, development and application of AI in manufacturing through the Industry 4.0 model, which as a model was "born" in the EU. Further analysis shows that AI in the EU is mostly applied in intelligent maintenance (MES), quality control and inspection (MES) and manufacturing planning and management (ERP). The study [16], discusses the challenges and possibilities of applying deep learning models in smart manufacturing. The analysis shows that CNN is the best model, and the challenges are: data quality, data security, and learning model reliability. All these analyzes were performed for the application of AI in maintenance, quality control, collaborative robots and supply chains. An in-depth analysis of the application of deep learning models (DML) in smart manufacturing is presented in the study [17]. The latest research shows that in smart manufacturing models, only 2% of the data generated is relevant for use and decision-making, so these approaches are also called data-driven systems based on AI, supported by deep learning (CNN, RBM, RNN). All this is illustrated with examples from maintenance and quality control. The study [18] provides a detailed overview of the application of machine learning techniques with application in smart manufacturing. However, examples of good practice also have major challenges in application, which should be taken into account: data collection and management, human resources (ML experts), infrastructure (hardware, software), information security risks and supplier business models. A maturity model of AI application in EU companies is presented in [19]. KPI parameters are defined for it, at the process, CPS, factory level, which is connected to the

RAMI model of horizontal and vertical integration. The maturity model evaluates the process, CPS and factory in five dimensions: productivity, time to market, worker role, resilience, society and environment. In the study [20], various aspects of the application of AI in smart manufacturing were analyzed, and some of the most important challenges are: data quality for deep learning models, integration of AI models into factory functions and systems, application costs, readiness of the factory environment for the application of AI models, legal and ethical rules and restrictions, and employee resistance to change. All this means that Project AI in the factory should be handled very carefully. A particular challenge for implementing AI in the factory is to do it using the intelligent DT factory model [21]. The factory was modeled as a hierarchic system through the IDEFO model, and then a virtual counterpart was developed for all components. The entire system was tested in the laboratory on one example. A study [22] discusses the application of generative AI in manufacturing in the US, by the US Society of Manufacturing Engineers. Namely, it is considered that generative AI is the immediate future of the application of AI in manufacturing, so a policy and strategy should be established at the national level to encourage its application, which is detailed in the study.

Therefore, concluding the analysis of this chapter, we can state the following: (i) industrially developed countries have their own national and regional strategies and programs for the application of AI in manufacturing, (ii) the Industry 4.0 model is the best base and framework for the application of AI in manufacturing, and especially the model of smart manufacturing, which is developed from this concept, and (iii) everything said says that Serbia should also seriously address these trends and approaches, all with the aim of developing and building a new industry.

3.3. Analysis of Learning Models Applied in Manufacturing

The essence of today's application of AI in manufacturing is the application of various learning models, especially those related to deep learning, to solve various engineering, business and managerial problems. That is why this area is separated as a unit, and examples of the application of these models are given in Table 4.

Supervised (deep learning) and unsupervised learning are previously widely used learning models for the application of AI in manufacturing. One example of the application of these models is given in [23], and it refers to the prediction of the quality of the technological process in a steel rolling mill. In order to model the rolling process online, the following are monitored: rolling force, speed and temperature of the process, and all these parameters are monitored via six characteristics.

Table 4. Overview of the application of AI learning models in manufacturing

The way	Ref. / yr.	Method	The goal	Use
Three process parameters	[23] 2013	4 Learning models	Better quality of rolled steel	Process quality management
45 process parameters	[24] 2013	Method of support vectors	The same quality of every product	Quality monitoring of 4 technological processes
Paradigms of intelligent manufacturing	[25] 2017	Deep learning	Intelligent manufacturing	Model-driven collaborative manufacturing
Development of the application of ML in smart manufacturing	[26] 2018	ANN/CNN	State of ML Application in US Industry	Digital manufacturing
Deep learning model development	[27] 2018	Extreme Gradient Boost (XG Boost)	Improvement of manufacturing quality	Smart manufacturing at Bosch
Application of elements of Industry 4.0	[28] 2018	CNN, support vector method, random forest model, recurrent NN and Bayesian classifiers	Intelligent decision making	Smart manufacturing
Industrial intelligent model	[29] 2018	"ABCDE" model	Building an AI platform for smart manufacturing	Intelligent CPS (Machine Tool)
Application in the MES model	[30] 2019	CNN	Development of intelligent manufacturing	Intelligent plant management (MES)
Detection of assembly errors	[31] 2019	XG Boost, random forest model and support vector method	Error detection	Volkswagen assembly line
Predictive maintenance	[32] 2019	Online status monitoring	A new maintenance model	Machine care

The following learning techniques were used: deep learning - Bayesian classifiers, decision tree, K nearest neighbor and support vector method, and learning - K means. The sample for defining learning included 470 examples of measurement, and the results in application were excellent, which was demonstrated. In the study [24], a cluster model is described for the improvement of product quality through all its stages of manufacturing (casting, processing by plastic deformation, processing by cutting and thermal treatments). As a deep learning model, the support vector method was used, where the learning sample included 45 parameters with 360 samples. Thus, a model was established for online monitoring of product quality through the specified stages of manufacturing. The AI-supported intelligent manufacturing model is based on four paradigms [25]: (i) model-driven intelligent collaborative manufacturing, (ii) knowledge-based enterprise cloud service, (iii) human-machine-material cooperative cloud workshop, and (iv) autonomous intelligent manufacturing units (CPS). So, for example, the first paradigm means that cloud is a model of all kinds of manufacturing resources/capacity, and then intelligent cloud technology should be used to automatically match product requests and resource/service requests. Based on this, define an intelligent infrastructure for the product - support. Also, those operational centers of intelligent manufacturing platforms in the cloud must support model-driven collaborative activities of the total product life cycle: research and development,

manufacturing, manufacturing management, logistics and support services. All the above elements are supported by various AI learning models. The study, presented in [26] included the analysis of over 4000 references from the field of advanced manufacturing, from the aspect of application and possibility of applying machine learning models. Recommendations are given for the faster application of these models in advanced manufacturing, namely: management support in application and their greater use for decision-making, establishment and use of a digital knowledge base, application to the product life cycle and the basis from which to start for the successful application of ML is digital manufacturing. In manufacturing, an abundance of data is generated, from different sources and in different forms, so they should be organized and prepared for use in learning models [27]. Therefore, the construction of a learning model for smart manufacturing consists of four steps: (i) collection of raw data, (ii) their processing and editing, (iii) development and testing of the learning model (70% is used for the development of the learning model, and 30% for its testing (verification)), and (iv) application of the model in practice, through its validation. It (validation) includes the following steps: accuracy, sensitivity, precision and F-measure. All this is illustrated with an example of a deep learning model from Bosch smart manufacturing. The study [28] provides an overview of the application of deep learning models in manufacturing. The elements of Industry 4.0, such as the Internet of Things (IoT), big data (BDA), digital twins (DT) and cloud computing (CC), form an ideal framework for the application of deep learning models, such as: CNN, the method of support vectors, random forest model, recurrent NN and Bayesian classifiers, in smart manufacturing. These deep learning approaches in this study are illustrated with examples from: quality control and inspection, predictive maintenance, and condition diagnostics. One approach to the development of industrial AI models is defined as the "ABCDE" model, and is presented in [29]. The meaning of the labels in the model is: A - technology analytics, forms the core of the AI eco manufacturing system model (data, software/hardware platform, CPS, B - big data technology (BDA), C - cloud and cyber technology, D - knowledge domain, E - evidence Big data technology and cloud are essential elements of the data source that make up the AI platform Domain knowledge and evidence are key elements due to: (i) understanding the problem and focusing the power of AI on solving it, (ii) disseminating the system and collecting real data, (iii) understanding the physical meanings of the parameters and their relationship in the model, and (iv) communicating the change of these parameters from one CPS to another The development and application of the Industry 4.0 model is focused on the intelligent development of the MES model, and for its application it is necessary to use the deep learning model, especially the CNN model shown in [31]. It is a Volkswagen assembly line, where the sample included 18148 units of products with 29 characteristics. The developed model enabled the

detection of assembly errors with an accuracy of 98.25%, which was satisfactory. The study [32] shows the application of a deep learning model for predictive maintenance (online monitoring of machine condition and intervention as needed). The paper showed that convolutional neural networks (CNN) and recurrent neural networks (RNN) are mostly used for state monitoring (machine care), deep learning models. It is expected that research in this area will develop in the following directions: (i) for large open-source data sets, CNN models up to 150 layers, with 10 hidden layers each, will be developed, (ii) greater use of the knowledge domain (not only process knowledge), (iii) visualization of learning data, (iv) deep learning knowledge transfer, (v) balancing knowledge classes about the state (health) of machines.

At the end of the analysis of learning models applied in manufacturing, we can draw the following conclusions: (i) deep learning models are widely applied today, especially in maintenance and quality control, and (ii) the wider application of Industry 4.0 elements is expected to encourage further development and application of deep learning models in manufacturing.

3.4. Generative AI and its Application in Manufacturing

Generative AI (GenAI) models typically use machine learning algorithms to analyze patterns and relationships in existing data and then generate new content that is often difficult or impossible for humans to create, an approach that also applies to manufacturing. Thus, generative AI models are being researched and developed today for: product design and design for additive manufacturing (AM), materials science, optimization of business and technological processes, robotics and automation, predictive maintenance, quality control and supply chain management. Limitations in the application of these AI models can be: data quality, interoperability, scalability and cyber security. An overview of selected examples of generative AI applications is given in Table 5.

Table 5. Overview of the application of generative AI (GenAI) models in manufacturing

The way	Ref. / yr.	Method	The goal	Use
Business and technological processes	[33] 2023	GenAI analytics	Process transformation	In the Industry 4.0 model
A three-level model	[34] 2023	Integration	Autonomous system	Factories of the future
GenAI model	[35] 2024	CAD model	A new design model	Product design
GenAI model architecture	[36] 2024	Four levels	Business Improvement	Smart factory
GenAI as a support for Industry 4.0	[37] 2024	Integration	Application of the GenAI concept	Factory with the Industry 4.0 model
Identification of 10 areas of innovation management	[38] 2024	A new model of innovation management	Faster implementation of innovations in practice	Innovation management

GenAI has been applied to several elements of Industry 4.0, namely: forecasting demand and marketing strategy, designing new products, optimizing the workforce and their skills, improving quality control and predictive maintenance [33]. GenAI enables managers to transform manufacturing by optimizing processes, improving product design, improving quality products, thus raising the application of innovations in manufacturing. At the same time, GenAI technologies facilitate predictive analytics for demand forecasting as well as improvement marketing strategies and identifying market trends. All of the above is illustrated with several examples. A particularly interesting analysis was presented in the study [34], which referred to the place and role of GenAI in the factories of the future. A three-level model is proposed: assistance, recommendation and autonomous systems. Assistance means that GenAI should generate the program code for the machine tool, and the engineer only reviews it, while recommendation means that GenAI guides the maintenance technician to select the necessary maintenance instructions in a predictive model, step by step, including spare parts. And finally, autonomous systems in the factory of the future supported by GenAI, are those that will have self-regulating capabilities and adapting to unknown situations. For the factory of the future, AI will be the cornerstone of its business, and the key model will be GenAI [35]. Here's how the new approach - GenAI looks at product design. GenAI brings revolutionary changes in product design, bringing a unique blend of creativity and efficiency. More specifically, generative text-to-image tools help designers bridge the gap between concepts and manufacturing-ready models.

The job of a product design engineer is to define specific design goals, taking into account design metrics, such as: sustainability goals, manufacturing costs, product compliance with customer (market) requirements and manufacturing conditions. GenAI systems generate different design options based on these predefined parameters. In the study [36], the architecture of the GenAI model is shown, which includes four levels: (i) hardware infrastructure, (ii) work on the IoT (Industrial Internet of Things) platform PTC ThingWorx that manages business (sales, procurement), engineering (CAD, CAM, PLM) and manufacturing information system (ERP, MES), (iii) machine learning models for GenAI - Azure OpenAI Service, and (iv) application development for specific problems of the organization. The study [37] talks about how the Industry 4.0 model can be the basis for the application of the GenAI concept. In this sense, it is suggested that each organization, which starts this way, must define a special integration project. The study [38] describes future research opportunities related to the application of GenAI in innovation management, where ten areas of application of this model in innovation are identified.

Concluding this analysis, we can conclude the following: (i) GenAI is a new model of applying AI in manufacturing, whose possibilities and application examples are still being explored, and (ii) it can be said that the first experiences about it indicate that this approach is more user-oriented than other deep learning models, which represents a good basis for its faster penetration into manufacturing.

3.5. Application of ChatGPT in Manufacturing

ChatGPT is a powerful AI language model, which can be applied in various industries to automate tasks, improve efficiency and improve user experience. So for example in a smart factory, this AI tool can be useful for us to: (i) help MES by generating reports, provide real-time feedback and enable remote monitoring and control of manufacturing processes, (ii) help ERP systems by generating reports, by providing personalized customer support and optimizing supply chain logistics, (iii) assisting QMS by analyzing data from quality control processes to identify defects and causes of errors, and (iv) assisting PLM by generating digital twins of products or systems, enabling tracking and real-time simulation of manufacturing and maintenance. An overview of the application of this AI model is shown in Table 6.

Table 6. Overview of the application of the ChatGTP model in manufacturing

The way	Ref. / yr.	Method	The goal	Use
Collection and analysis of knowledge	[39] 2023	Searching for knowledge	Defining knowledge of terms (26)	Knowledge about Industry 5.0
IoT	[40] 2023	Analysis of information from manufacturing	Report generation	Manufacturing
Flag presentation	[41] 2023	Answers to the questions	Engineer training	Help the engineer
Platform of business and technological knowledge	[42] 2023	Online management	Support for smart manufacturing	Industry 4.0 model

Therefore, we can state that the application of AI in manufacturing is increasingly moving from algorithmic to linguistic intelligence, where interactive activities between humans and machines play an active and important role online in real time. ChatGPT has proven its effectiveness in providing comprehensive information and knowledge about Industry 5.0 [39]. For the defined 26 keywords (AI for advanced automation, metaverse learning and optimization,...), satisfactory answers were given from the large knowledge base that was used. The study [40] talks about the challenges of applying this AI model in different areas, but the context of applying IoT in manufacturing is interesting for us. More precisely, how the information coming from the manufacturing can be used for the application of this tool, which was already stated in the introduction of this chapter. Manufacturing includes a large number

of complex tasks, which require extensive knowledge and experience to perform. With the rapid development of AI, especially with the emergence of language models such as ChatGPT, new opportunities for knowledge assurance are opening up through the conversation [41]. In this study, it was shown that ChatGPT, as a generative model, provides promising creations in a comprehensive, creative, and objective manner, thereby demonstrating its potential to support works of summarization, synthesis, and creation. The study also presented a three-layer model developed for the needs of manufacturing and training of engineers. In the concept of Industry 4.0, ChatGPT will be a business and technological knowledge platform that will be of great help to engineers and managers to manage smart manufacturing online [42].

Concluding the analysis of this chapter, we can state that this model (ChatGPT) is in its infancy, so new experiences about its application in manufacturing are still expected.

At the end of this detailed analysis, we can conclude that AI in manufacturing enters a new phase of development and application, listed in Table 1 as level 5, which refers to smart manufacturing, and the basis of that application is the GenAI model.

4. Some Research Results of the Faculty of Mechanical Engineering in Belgrade

This analysis refers only to the first-named author's research results in this area. Namely, we can state with pleasure that in the mid-eighties of the last century, under the leadership of prof. Dr. Vladimir Milačić, at the Faculty of Mechanical Engineering in Belgrade, the research program - Intelligent Manufacturing Systems began to be implemented, where the first-named author had a significant place and role. As a young researcher, I completed and defended my PhD thesis in 1988, in the field of development of an Expert system for maintenance of machine tools (EXMAS), which was also verified by the most important references of that time [43 - 45]. Research in the field of ES continued in the field of inspection on CMM [46], with the development of a model for inspection of box parts on CMM (EXINS), also verified by the most important reference at that time. Later, these researches were continued, so that at the end of the second decade of this century, this author developed an original model of virtual manufacturing metrology (CPM³ - Cyber Physiacal manufacturug metrology model), also verified in a large number of references of the most important international journals, and in Mongrafia [47], which was published by Springer, and it received the Award of the City of Belgrade for Science, 2019, as the highest national recognition for science.

This brief overview tells us that at the Faculty of Mechanical Engineering in Belgrade, at the Department of Manufacturing Engineering, there is an enviable history of the development and application of AI in the aforementioned field.

5. AI in Manufacturing – What's next?

The future of manufacturing is smart manufacturing, or otherwise, the full application of the Industry 4.0 model in manufacturing. In such a sequence of events, the role of AI will be large and significant, which means that it will apply to [19,20,34]: (i) digital manufacturing, which will be based on AI, where it will create digital versions of physical products, enabling virtual prototyping, simulation and testing, before real manufacturing, (ii) design for manufacturing, AI will help to create products suitable for manufacturing, which will meet all the requirements of the customer, ecology while saving resources, (iii) optimization of the manufacturing process, AI can analyze manufacturing data to optimize manufacturing processes, reducing energy consumption, waste and costs, (iv) intelligent manufacturing planning, based on AI can optimize manufacturing resources, based on demand forecasting, supply chain constraints and manufacturing capacity, (v) optimization supply chain, using AI to optimize inventory, logistics and delivery routes, reducing costs and shortening delivery times, as well as risk management, (vi) automation and collaborative robots, use AI algorithms to optimize assembly, welding and inspection tasks. Cobots with AI increase productivity and reduce labor costs, (vii) augmented reality (AR) and digital twins (DT) supported by AI are used to simulate processes, machines and systems, and for their maintenance, (viii) condition monitoring and predictive maintenance, use machine learning algorithms to analyze equipment performance data, predicting when maintenance is needed and proactively schedule maintenance, reducing downtime and increasing overall equipment efficiency, (ix) quality control and assurance, AI is used to analyze product quality data in real time, detecting defects and errors, while enabling rapid corrective actions to improve product quality, (x) AI-based data analytics will analyze large amounts of manufacturing data (BDA) to identify trends, behavioral patterns and correlations, enabling data-driven decision making, and (xi) cyber - AI-powered security systems can detect threats and respond to them in real-time, protecting against cyber-attacks.

These are some directions of the future application areas of AI in smart manufacturing, and as AI technology (GenAI, ChatRTG) continues to develop, we can expect even more innovative applications to emerge in the future.

6. Conclusion

Since the basic context of this work is AI in manufacturing, in conclusion we can point to trends in the future development of AI, which will have a direct impact on its application in smart manufacturing. A few basic AI development trends are: (i) general AI (GAI), refers to hypothetical AI systems, which will possess human-like AI, capable of reasoning, learning and applying knowledge to a wide range of tasks. In that case, GAI could surpass human intelligence in many areas, (ii) edge AI, refers to the use of AI and machine learning in frontier areas, where data is generated and processed in real time, without the need for centralized processing, (iii) transfer learning, is a technique in which AI models can learn from one task or domain and apply that knowledge to another related task or domain, without the need for relearning, and (iv) quantum AI, similar to quantum computing, which has the potential to advance AI, providing new methods for machine learning, optimization and simulation, which could lead to breakthroughs in areas such as smart manufacturing, medicine, finance and climate modelling.

All the mentioned facts speak of the great perspective of the development and application of AI in smart manufacturing as well.

7. References

- [1] Lihui Wang, From Intelligence Science to Intelligent Manufacturing, Engineering, Volume 5, Issue 4, 2019, Pages 615-618, 10.1016/j.eng.2019.04.011.
- [2] Kiritsis, D., Hodkiewicz, M., Lazaro, O., Lee, J., Ni, J., eds. (2021). Data-Driven Cognitive Manufacturing - Applications in Predictive Maintenance and Zero Defect Manufacturing. Lausanne: Frontiers Media SA.
<https://doi.10.3389/978-2-88966-583-9>.
- [3] Vidosav D. Majstorović, Cognitive Manufacturing, Keynote Lecture, Conference Proceedings Mechanical Technologies and Structural Materials, Split, Croatia, 21 - 22 September 2023 [cited 2024, June 15]. Available at: <http://www.strojarska-tehnologija.hr>.
- [4] Julie Basello and Shannon Feeley , The History of AI in Manufacturing, [cited 2024, Jun 15]. Available at: <https://blog.radwell.com/the-history-of-ai-in-manufacturing>.
- [5] Evelyn Du Jack , AI in manufacturing: Where it's been and where it's going, [cited 2024, Jun 15]. Available at: <https://www.controleng.com/articles/ai-in-manufacturing-where-its-been-and-where-its-going/>.

- [6] The Rise of AI in Manufacturing: A Brief History, [cited 2024 Jun 15]. Available at: <https://www.khenda.com/post/the-rise-of-ai-in-manufacturing-a-brief-history>.
- [7] Systematic reviews and Meta-Analyses (PRISMA), [cited 2024 Jun 15]. Available at: <https://www.prisma-statement.org/>.
- [8] Shen, Weiming & Wang, Lihui & Hao, Qi. (2006). Agent-based distributed manufacturing process planning and scheduling: A state-of-the-art survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 36. 563 - 577. 10.1109/TSMCC.2006.874022.
- [9] Leitão, Paulo. (2009). Agent-Based Distributed Manufacturing Control: A State-of-the-Art Survey. Engineering Applications of Artificial Intelligence. 22. 979-991. 10.1016/j.engappai.2008.09.005.
- [10] dr. Bashar El-Jawhari, et al., PwC's An introduction to implementing AI in manufacturing, 2020 [cited 2024 Jun 15]. Available at: <https://www.pwc.com/gx/en/industrial-manufacturing/pdf/intro-implementing-ai-manufacturing.pdf>.
- [11] Brian McCarson, Eight Key Considerations when Implementing AI in Manufacturing, Planning for Success in our Journey to Implementing AI, 2020, [cited 2024 Jun 15]. Available at: <https://www.rtinsights.com/wp-content/uploads/2020/09/AI-in-Manufacturing-Ebook-092020-3.pdf>.
- [12] AI in Manufacturing, Moving from Aspirations to Implementations, White paper, [cited 2024, Jun 15]. Available at: <https://www.connection.com/manufacturing#form>.
- [13] AI Ensemblment on the way to smart manufacturing, 2020, [cited 2024 Jun 15]. Available at: <https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/cip/deloitte-cn-cip-ai-manufacturing-application-survey-en-200116.pdf>.
- [14] Kim, Sung & Kong, Jun & Lee, Jiho & Lee, Seungchul. (2021). Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review. International Journal of Precision Engineering and Manufacturing. 23. 10.1007/s12541-021-00600-3.
- [15] Scaling AI in Manufacturing Operations, 2021, [cited 2024, Jun 15]. Available at: <https://www.sogeti.ie/explore/reports/scaling-ai-in-manufacturing-operations/>.
- [16] Xu, J.; Kovatsch, M.; Mattern, D.; Mazza, F.; Harasic, M.; Paschke, A.; Lucia, S. A Review on AI for Smart Manufacturing: Deep Learning Challenges and Solutions. Appl. Sci. 2022, 12, 8239. 10.3390/app12168239.
- [17] Pawar, Prajakta & Adhav, Mr & Naik, Miss & Choudhari, Miss & Tonge, Miss. (2022). Methods and Applications of Deep Learning for Smart Manufacturing. GIS Science Journal. Volume 9, Issue 5, 2022, pp 641-655. [cited 2024, Jun 15]. Available at: <https://www.researchgate.net/>

- publication/362520050_ Methods_and_Applications_of_Deep_Learning_for_Smart_Manufacturing.
- [18] Plathottam, Siby Jose & Rzonca, Arin & Lakhnori, Rishi & Iloeje, Chukwunwike. (2023). A review of artificial intelligence applications in manufacturing operations. *Journal of Advanced Manufacturing and Processing*. 5. 10.1002/amp2.10159.
 - [19] Makris, S. Alexopoulos, et al., (2023). Artificial Intelligence In Manufacturing, White paper, Prepared by the Artificial Intelligence in Manufacturing Network - AIM - NET. 10.13140/RG.2.2.23010.13763.
 - [20] Balasubramanian, S & Scholar II, Research. (2023). Integration of Artificial Intelligence in the Manufacturing Sector: A Systematic Review of Applications and Implications. *International Journal of Production Technology and Management*. Vol. 14. Pages 1-11. 10.17605/OSF.IO/3XPWN.
 - [21] Urgo, Marcello, Walter Terkaj and Gabriele Simonetti. "Monitoring manufacturing systems using AI: A method based on a digital factory twin to train CNNs on synthetic data". *CIRP Journal of Manufacturing Science and Technology* 50 (2024): 249 - 268. 10.1016/j.cirpj.2024.03.005.
 - [22] Kathryn Wengel, Working Smarter: How Manufacturers Are Using Artificial Intelligence. White paper. 2024 [cited 2024, Jun 15]. Available at: <https://nam.org/wp-content/uploads/2024/05/NAM-AI-Whitepaper-2024-1.pdf>.
 - [23] Lieber, Daniel & Stolpe, Marco & Konrad, Benedikt & Deuse, Jochen & Morik, Katharina. (2013). Quality Prediction in Interlinked Manufacturing Processes based on Supervised & Unsupervised Machine Learning. *Proceeding CIRP*. 7. 193–198. 10.1016/j.procir.2013.05.033.
 - [24] Wuest, Thorsten & Irgens, Chris & Thoben, Klaus-Dieter. (2014). An approach to quality monitoring in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*. 25. 1167-1180. 10.1007/s10845-013-0761-y.
 - [25] Li, Bo & Hou, Bao-cun & Yu, Wen-tao & Lu, Xiao-bing & Yang, Chun-wei. (2017). Applications of artificial intelligence in intelligent manufacturing: a review. *Frontiers of Information Technology & Electronic Engineering*. 18. 86-96. 10.1631/FITEE.1601885.
 - [26] Sharp, Michael & Ak, Ronay & Hedberg, Thomas. (2018). A Survey of the Advancing Use and Development of Machine Learning in Smart Manufacturing. *Journal of Manufacturing Systems*. 48. 10.1016/j.jmsy.2018.02.004.
 - [27] Carbery, CM, Woods, R., & Marshall, AH (2018). A New Data Analytics Framework Emphasizing Preprocessing in Learning AI Models for Complex Manufacturing Systems. In Z. Yang, D. Yang, K. Li, M. Fei, & D. Du (Eds.), *Intelligent Computing and Internet of Things - First International*

- Conference on Intelligent Manufacturing and Internet of Things and 5th International Conference on Computing for Sustainable Energy and Environment, IMIoT and ICSEE 2018, Proceedings (pp. 169-179). (Communications in Computer and Information Science; Vol. 924). Springer-Verlag. https://doi.org/10.1007/978-981-13-2384-3_16.
- [28] Wang J, et al. Deep learning for smart manufacturing: Methods and applications. Volume 48, Part C, July 2018, Pages 144-156, 10.1016/j.jmsy.2018.01.003.
- [29] Lee, Jay & Davari, Hossein & Singh, Jaskaran & Pandhare, Vibhor. (2018). Industrial Artificial Intelligence for Industry 4.0-based Manufacturing Systems. *Manufacturing Letters*. 18. 10.1016/j.mfglet.2018.09.002.
- [30] Nascimento, Alexandre & Porto-Bellini, Carlo. (2018). Editorial - Artificial intelligence and industry 4.0: The next frontier in organizations. *BAR - Brazilian Administration Review*. 15. 10.1590/1807-7692bar2018180152.
- [31] Peres, Ricardo & Barata, J. & Leitão, Paulo & Garcia, Gisela. (2019). Multistage Quality Control Using Machine Learning in the Automotive Industry. *IEEE Access*. 7. 1-1. 10.1109/ACCESS.2019.2923405.
- [32] Zhao, Rui & Yan, Ruqiang & Chen, Zhenghua & Mao, Kezhi & Wang, Peng & Gao, Robert. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*. 115. 10.1016/j.ymssp.2018.05.050.
- [33] Doanh, Doung Cong, Dufek, Zdenek, Ejdys, Joanna, Ginevičius, Romualdas, Korzynski, Pawel, Mazurek, Grzegorz, Paliszkievicz, Joanna, Wach, Krzysztof and Ziemba, Ewa. "Generative AI in the Manufacturing Process: Theoretical Considerations" *Engineering Management in Production and Services*, vol.15, no.4, 2023, pp.76-89. 10.2478/emj-2023-0029.
- [34] *Daniel Cooper, et al.*, Generative AI's Role in the Factory of the Future, Boston Consulting Group, 2024 [cited 2024 Jun 15]. Available at: <https://www.bcg.com/publications/2023/gen-ai-role-in-factory-of-future>.
- [35] Generative AI in Manufacturing: 5 Industry-Transforming Use Cases, Amazuta, 2024 [cited 2024 Jun 15]. Available at: <https://www.azumuta.com/blog/generative-ai-in-manufacturing-5-industry-transforming-use-cases/>.
- [36] Amol Adgaonkar, et al., GenAI Ushers Human-Like Intelligence in Manufacturing, *ML Journal*, 2024 [cited 2024 Jun 15]. Available at: <https://manufacturingleadershipcouncil.com/genai-ushers-human-like-intelligence-in-manufacturing-36245/>.
- [37] Jeff Winter, How AI is speeding the Rise of "Industry 4.0", [cited 2024, Jun 15]. Available at: <https://social-innovation.hitachi/en-us/think-ahead/manufacturing/ai-for-smart-manufacturing-industry-4-0/>.

- [38] Mariani, Marcello & Dwivedi, Yogesh. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. *Journal of Business Research*. 175. 114542. 10.1016/j.jbusres.2024.114542.
- [39] Wang, Fei-Yue & Yang, Jing & Wang, Xingxia & Li, Yuanjuan & Han, Qing-Long. (2023). Chat with ChatGPT on Industry 5.0: Learning and Decision-Making for Intelligent Industries. *IEEE/CAA Journal of Automatica Sinica*. 10. 831-834. 10.1109/JAS.2023.123552.
- [40] Gill, Sukhpal Singh & Kaur, Rupinder. (2023). ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems*, Volume 3, 2023, Pages 262-271. 10.1016/j.iotcps.2023.05.004.
- [41] Wang, Xingzhi & Anwer, Nabil & Dai, Yun & Liu, Ang. (2023). ChatGPT for design, manufacturing, and education. *Procedia CIRP* 119 (2023) 7–14 10.13140/RG.2.2.35077.22244.
- [42] Javaid, Mohd & Haleem, Abid & Singh, Ravi. (2023). A study on ChatGPT for Industry 4.0: Background, Potentials, Challenges, and Eventualities. *Journal of Economy and Technology*. 1. 10.1016/j.ject.2023.08.001.
- [43] VD Majstorovic, VR Milacic, and IZ Race, "Building expert system for diagnosis and maintenance in FMS," *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, Philadelphia, PA, USA, 1988, pp. 1126-1129 vol.2, 10.1109/ROBOT.1988.12211.
- [44] VD Majstorovic, VR Milacic, An Expert System for Diagnosis and Maintenance in FMS, *CIRP Annals*, Volume 38, Issue 1, 1989, Pages 489-492, 10.1016/S0007-8506(07)62752-8.
- [45] VD Majstorović, Expert systems for diagnosis and maintenance: The state-of-the-art, *Computers in Industry*, Volume 15, Issues 1–2, 1990, Pages 43-68, 10.1016/0166-3615(90)90084-3.
- [46] PO Bojanic, DV Majstorović, RV Milačić, CAD/CAI Integration with Special Focus on Complex Surfaces, *CIRP Annals*, Volume 41, Issue 1, 1992, Pages 535-538, 10.1016/S0007-8506(07)61262-1.
- [47] Stojadinović, S., Majstorović, V. (2019). An intelligent inspection planning system for prismatic parts on CMMs, Springer International Publishing., ISBN 978-3-030-12806-7, 10.1007/978-3-030-12807-4.

Prompt Engineering

Samir Lemeš^{*1}

Abstract: *Prompt engineering is the process of designing, testing, and optimizing prompts that are sent to artificial intelligence (AI), especially large language models like GPT-4. The goal is to formulate the prompts in a way that allows the AI model to provide the most relevant, accurate and useful answers. This process involves understanding how the AI model interprets and responds to different query formulations, as well as tailoring those queries for different applications. Well-designed prompts can significantly improve the performance of AI systems, making them more useful and efficient for various applications. In the context of Industry 4.0, where AI is used for production optimization, equipment maintenance, data analytics and other critical functions, more efficient prompts can contribute to increased productivity and reduced costs. Prompt engineering is an interdisciplinary field that requires understanding of natural language, artificial intelligence, and domain-specific applications. As AI models become more sophisticated, the role of prompt engineering will become even more important in leveraging their full potential.*

Keywords: *Machine Learning, Large Language Models (LLM), Prompt, Industry 4.0*

1. Introduction

Prompt engineering is the process of designing, testing, and optimizing prompts that are sent to large language models such as GPT. It is a new technique that is interdisciplinary and requires a wide range of other skills that are not always necessarily limited to the field for which artificial intelligence is used. Prompt engineering is an increasingly popular topic for wide research in various fields. Gu et al. have attempted to provide a comprehensive overview of recent research on prompt engineering across three types of visual language models: multimodal-to-text generation models (e.g., Flamingo), image-to-text matching models (e.g., CLIP), and image-to-text generation models (e.g., Stable Diffusion) [1]. Although they refer to 216 literature sources, their research was limited to pre-trained models only.

Strobelt et al. have developed a workflow that allows users to focus on model feedback using small data before moving to a big data mode that allows for empirical grounding of prompts using quantitative measures of the task [2].

^{*1}University of Zenica, Polytechnic Faculty, Zenica, Bosnia and Herzegovina
ORCID: 0000-0002-3596-645X, E-mail: samir.lemes@unze.ba

They have developed an open-source system (<http://prompt.vizhub.ai/>) that can work with any available language model backend.

Marvin et al. in [3] provided a thorough understanding of prompt engineering, with relevant exercises for applying these engineering techniques in practice, current and future LLM trends, and prompt engineering research, including the rise of automatic instruction generation and query selection methods.

Sahoo et al. have addressed the lack of systematic organization and understanding of various prompt engineering methods and techniques by providing a structured overview of recent advances in prompt engineering, categorized by application area with a description of the methodology, its application, the models involved, and the datasets used [4]. They have also addressed the strengths and limitations of each approach, including a taxonomy diagram and table summarizing the datasets, models, and critical points of each technique.

Ekin in [5] provided a comprehensive guide to mastering prompt engineering techniques, tips and best practices for achieving optimal results for ChatGPT. He also covered best practices, including iterative refinement, balancing user intent, leveraging external resources, ensuring ethical usage, and advanced strategies such as temperature and token control, fast chaining, domain-specific customizations, and handling ambiguous inputs.

Henao, Franco-Cardona and Cadavid-Higueta introduced a methodology to optimize interaction with language models of artificial intelligence, such as ChatGPT, through prompt engineering [6]. They called the approach GPEI, which consists of four steps: define the goal, design the prompt, evaluate the response, and repeat. Their proposal includes two key aspects: incorporating data into the design of prompts for engineering applications and integrating explainable artificial intelligence principles to evaluate responses, increasing transparency.

Shin et al. investigated the effectiveness of GPT-4 LLMs with three different prompt engineering techniques (basic prompting, learning in context, and task-specific prompting) against 18 fine-tuned LLMs on three typical automated software engineering (ASE) tasks: code generation, code compression and code translation [7]. They concluded that GPT-4 with conversational prompting (i.e. when a human provides feedback and instructions to the model to achieve the best results) showed a drastic improvement compared to GPT-4 with automatic prompting strategies. They observed that participants tended to seek improvements, add more context, or provide specific instructions as conversational prompts, which went beyond typical and generic prompting strategies.

Korzynski et al. in [8] aimed to create a theoretical framework that would highlight optimal approaches in the field of prompt engineering for AI. The study revealed the profound implications of prompt engineering for the

application of artificial intelligence in various domains such as entrepreneurship, art, science and health. They have shown how efficient prompting can significantly improve the performance of large language models (LLMs), generating more accurate and contextually relevant results.

López-Riobóo-Botana, Gallent-Iglesias and Gonzalez-Vázquez presented QUA4I (Question Answering for the Industry 4.0), a chatbot or IVA (Intelligent Virtual Assistant) oriented application for the Industry 4.0 domain, mixing NLU and NLG techniques and using Rasa chatbot framework [9]. They designed a custom demo for answering questions and extracting information about Industry 4.0, including a dialogue system that can generate automatic responses in natural language. They have also included ASR (Automatic Speech Recognition) and TTS (Text to Speech) modules, so one can interact with the bot using Spanish language speech.

Tinnes et al. have explored the use of large language models for the automatic conversion of unstructured asset data into industry-relevant structures in the Asset Administration Shell (AAS) [10]. For some product categories, they showed that full automation is possible — in the worst case, 17% of the structures had to be corrected manually. They provided a new set of real-world industrial data as well as tools for comparing AAS structures for future studies.

Blasek et al. investigated the quality of LLM outputs to define requirements for Digital Twins [11]. The main contribution of their work is the results of an experiment that compared requests from DT air conditioners and ChatGPT and the prompt engineering conclusions that resulted from that experiment.

Howell investigated how artificial intelligence can be used in manufacturing processes [12]. He showed how prompt can help transform manufacturing processes, for example by improving predictive analytics to improve supply chain management through proactive risk detection and mitigation.

2. Large Language Models and Natural Language Processing

One of the most fascinating features of the technology that became freely and publicly available in November 2022 is its ability to understand different languages. Unlike previous search services, which required precise and correct entry of search terms, large language models, on which services such as OpenAI ChatGPT, Microsoft Copilot or Google (Bard) Gemini are based, can process even hybrid sentences in which different languages are mixed.

The terminology used by software engineers and other IT professionals is a kind of mixture of English and other languages, because some terms are simply untranslatable. Figure 1 shows a ChatGPT prompt made in such a hybrid language, and the response shows that the model understands such mixture of languages without any problems. The model recognizes even an incorrect transcription (the word "diplojamo" is derived from the English term "deploy",

which is not translated, but is already a common incorrect transcription that is perfectly understood by people involved in software development).

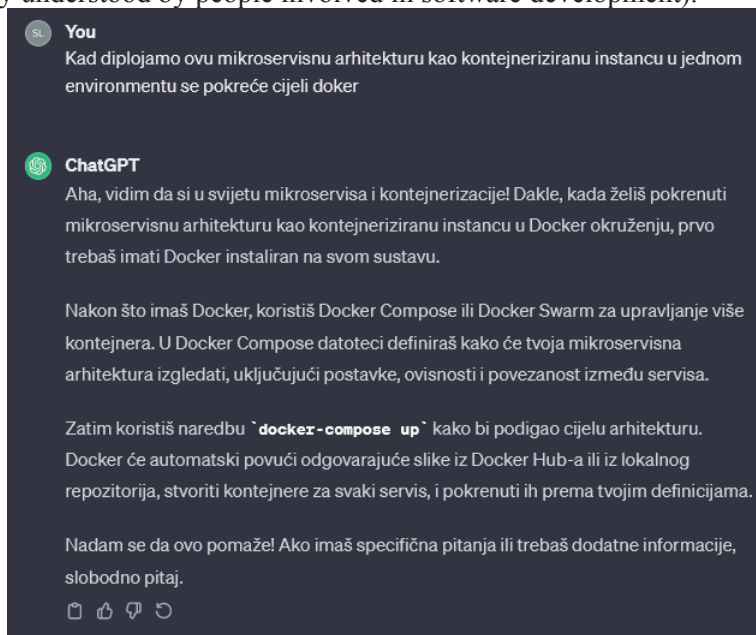


Figure 1. An example of a reply to a query made to ChatGPT with a mixture of incorrect English transcription and Bosnian language

Large Language Models (LLMs) are the result of decades of research in the field of Artificial Intelligence (AI) and Natural Language Processing (NLP). They were developed thanks to advances in machine learning, especially deep learning, and the vast amount of textual data available. Algorithms in early attempts (until the 1980s) of language processing using rules and small dictionaries were based on fixed rules (Rule-Based Systems). Statistical methods used in the 1990s, such as n-gram models, which used large amounts of data to learn the probabilities of word sequences, had limitations in understanding context. The introduction of neural networks, especially Recurrent Neural Networks (RNN) and Long-Term Memory Networks (LSTM), enabled a better understanding of sequences and contexts, or models based on deep learning. Transformer architecture, presented in 2017 [13] enabled parallel data processing, which significantly speeded up training and improved performance. The development of models such as OpenAI GPT-3 (Generative Pre-trained Transformer 3), with 175 billion parameters, is trained on huge collection of data from the Internet, which allowed them to generate high-quality text and answer complex queries.

Large language models use several techniques to recognize and to correct the misspelled text:

- Contextual understanding: Using transformers, models analyse the context of the entire text, which allows them to recognize irregularities and generate correct versions.
- Ability to generalize: They are trained on many different texts, including texts with errors. This helps them recognize and correct common mistakes.
- Correlation of sound and spelling: In poor transcription, models use phonetic similarities and context to infer which word was most likely intended to be used.
- Translation errors: When recognizing text that is the result of poor transcription or translation, the models use parallel sets of texts and their translations to identify and correct errors.

Today, large language models are used in numerous applications and can recognize and correct irregularities in text thanks to their ability to understand context and generalize from large bodies of data. These technologies continue to evolve, making interaction with machines more and more natural and efficient.

It is important to note that LLMs are not perfect and can sometimes make mistakes. However, they are constantly improving and getting better at recognizing and correcting errors in the text.

3. Possible Consequences of Using LLM and NLP

The use of large language models and NLP technologies in Industry 4.0 brings significant advantages, but also potential risks. Positive consequences include increasing productivity, improving user experience, improving quality and safety, and encouraging innovation. On the other hand, negative consequences include privacy concerns, ethical issues, labour substitution, and technical complexity. The role of prompt engineering is key to optimizing model performance and tailoring their responses to specific industry needs, which is essential for success and sustainability in the Industry 4.0 era.

The positive consequences of using LLM and NLP in Industry 4.0 include:

- Increased productivity through automation and faster decision-making: LLM and NLP enable the automation of routine and repetitive tasks, such as answering user queries, analysing data and generating reports. The integration of NLP technologies can speed up the analysis and decision-making process, enabling faster reactions to market changes.
- Improving user experience through personalization and ongoing support: LLMs enable personalized communication with users, providing responses tailored to individual needs and preferences. Chatbots and virtual assistants enable constant (24/7) support, improving customer satisfaction.

- Improving quality and safety: NLP can help analyse customer feedback and automatically identify product quality issues. Analysis of text data can identify potential security threats and risks.
- Innovation and research: Analysing large amounts of data can reveal market trends and needs, enabling the development of innovative products and services. NLP can speed up the process of searching and analysing scientific papers, facilitating research and development.

Of course, there are also potentially negative consequences of using these technologies:

- Threat to privacy: The use of LLM often requires the collection and analysis of large amounts of data, which may pose a risk to user privacy. There is a risk of misuse of collected data, which can lead to problems with user trust.
- Ethical issues reflected through bias and manipulation of information: LLMs can inherit biases from the data they are trained on, which can lead to discriminatory decisions or content. NLP technologies can be used to spread misinformation or manipulate public opinion.
- Labor replacement: One of the most frequently used arguments against the introduction of Industry 4.0 is the automation of jobs, which can lead to the loss of jobs, especially those involving routine tasks. Employees will need to adapt to new technologies, which may require additional training and education. Someone accepts it as a problem, and someone as an opportunity.
- Technical complexity as a challenge in implementation and maintenance: Integrating LLM and NLP technologies can be technically demanding and expensive, which can be a challenge for smaller companies. Complex technology systems can be prone to errors and breakdowns, which can affect operational efficiency.

Prompt engineering refers to the design and optimization of queries sent to large language models to obtain the desired responses. In the context of Industry 4.0, prompt engineering plays a key role in several aspects:

- Model performance optimization: Well-formed queries can significantly improve the accuracy and relevance of the answers generated by the model, which is crucial for automated decision systems.
- Improving user experience: Through prompt engineering, models can be tuned to provide responses tailored to specific user needs or industry requirements.
- More efficient data analysis: It enables the creation of specific focused queries to extract relevant information from large data sets, facilitating analysis and making informed decisions.

- Innovation and research: Prompt engineering can encourage innovative and creative ways of using language models to explore new ideas and develop new products and services.
- Customized training and development of employees: Using prompt engineering, language models can provide personalized training and education programs tailored to the specific needs of employees.

Prompt engineering has the potential to become not only a technique, but even one of the most sought-after professions in the future.

After the initial enthusiasm for the possibilities provided by generative models of artificial intelligence, the first concerns also appeared, especially in IT industry. Some occupations will surely die out, as generative AI models can not only generate new but also optimize existing code for software. Some predictions [14] claim that the job of writing software code will be taken over by generative models of artificial intelligence in the next few years.

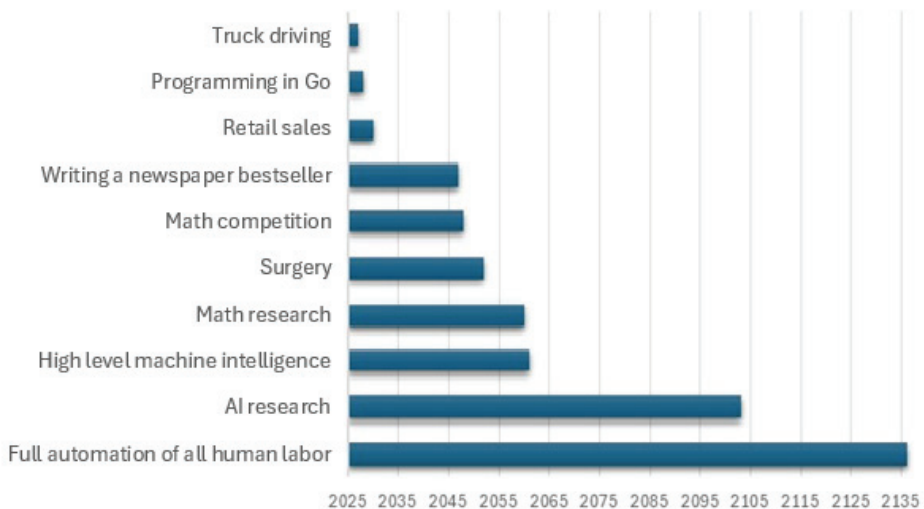


Figure 2. Predictions of when certain professions will be replaced by artificial intelligence [14]

4. Structure and Syntax of the Prompt

The results generated by AI depend not only on the size of the database on which language models are trained, but even more so on the formulation of the task. In short, prompt engineering is the skill of asking questions to a chatbot. It is a multidisciplinary skill, combining linguistics, logic, philosophy, engineering, etc. to achieve the best possible results from data generators based

on artificial intelligence. The precision and accuracy of querying with as much information and specific instructions as possible helps the model to generate the expected output, and to avoid generating irrelevant answers.

When defining a query (prompt), the rule "Garbage In = Garbage Out" applies, that is, a bad input usually generates a bad output. At the same time, the term "bad" does not refer to correct writing (spelling, grammar, language rules, terminology) at all, but rather to the formulation of the problem. This can be illustrated by an old programmer's joke, in which the wife instructs the programmer husband: "Go to the store, buy a butter, and if there are eggs, buy ten". The programmer came home with ten butters and the sentence "There were eggs". In this case, it is an insufficiently precisely formulated instruction that the programmer (and in all cases the software) does not critically analyse the instruction, but blindly listens to it in the form in which it was given. Mechanical response to insufficiently precise instructions cannot always be expected to produce the desired result.

The term "Garbage In = Garbage Out" in the case of applying generative AI models means that the quality of the output largely depends on the quality of the data corpus on which the model was trained. If that data is wrong, of course the result will also contain the error it inherited from the bad input.

However, the quality of the results can be influenced by knowing the correct structure of the prompt (Figure 3). In the introduction, a context should be set, which helps to give the AI model an imaginary 'role' to think about itself.

Phase	Example
Introduction	Act like a software engineer. You are an expert in Python and...
Task	I want you to develop software to manage your DVD collection.
Contextual information	I want it to be a web application written in Python.
Instructions	I want you to generate the source code.
Closing	I want it to be an AWS Lambda function.

Figure 3. Structure of a good prompt for generative AI models with examples

Separating instructions from context in a chatbot is an important step to achieve accuracy and relevance in responses. When asking questions or giving instructions, the user must be precise and clear. Specific language should be used, and ambiguity should be avoided. Key words in the instructions that are relevant to the desired task should be identified. The chatbot needs to analyse the context before executing the instructions.

Instructions are placed at the beginning of the prompt, denominators `###` or `"""` can be used to separate the instructions from the context, and the context is separated by an extra blank line. For example:

```
### Instructions ###
Write code to sort an array of names

""" A series of names: Sarajevo, Zenica, Tuzla,
Banja Luka, Mostar """
```

When formulating the prompt, user should be as precise as possible. Instead of "Write a poem about OpenAI" one should write "Write a short inspirational poem about OpenAI, focusing on the recent launch of DALL-E products in the style of {famous poet}". The desired output can be articulated through several examples. Giving examples can improve the quality of results. A positive prompt should be used (instead of "don't use a negative prompt"). Instructions should be clear and concise, avoiding ambiguity and unnecessary complexity. Context and specifics should always be considered. For example, the expressions "I'm sorry" and "My bad" do not sound the same in real life and at the funeral. A conversational style should be used instead of giving instructions. Rhetorical questions should be used and the active instead of the passive, for example instead of "An essay on the benefits of exercise should be written." (object + verb + subject) should be written "Can you write a persuasive essay on the benefits of exercise?" (subject + verb + object). For certain purposes, such as writing software code, there are generic queries that give good results (Figure 4).

<i>Scan the following code for potential problems</i>	Even if the code executes successfully, potential problems can be discovered
<i>Evaluate the following code and look for performance issues</i>	Improving code performance
<i>Write a test for the following {language} code</i>	Generating tests to examine the code
<i>Explain how {something} works in {language}.</i>	It explains how the borrowed code logic works
<i>Translate the following {first language} code into {second language}.</i>	I'm not learning a new programming language, translate C# to Java for me
<i>What is the correct syntax to {do something} in {language}?</i>	How do I send an HTTP header in Python?
<i>Write a function to {do something} in {language}.</i>	Write me a function to connect to a MongoDB database in Rust

Figure 4. Examples of well-worded software coding prompts

5. What is next?

The European Commission proposed the first regulatory framework for artificial intelligence in April 2021 [15]. This framework, known as the "Artificial Intelligence Act", sets the rules for the development, deployment and use of AI in the European Union to make AI systems safe, transparent, traceable and non-discriminatory, to ensure safety and fundamental human rights, and at the same time to encourage innovation.

ChatGPT introduced a completely new challenge to the education system – the paradigm of education must be changed. Writing essays and term papers can no longer be used to evaluate students, because it is practically impossible to determine the originality of such works, that is, such works do not prove the competence of those who claim to be the authors of those works.

The wide availability and prevalence of multiple generic AI models multiply the problems of plagiarism, wrong data and privacy violations. Some schools and states have tried to ban the use of new technologies in education, attempting to protect against misuse. The ban is not only ineffective, but even counterproductive. Instead of banning, it is necessary to learn how to use AI for personal and general benefit.

The development of technology is unstoppable, so we just need to learn how to use technology properly. The fear of losing jobs should be used to strengthen skills that help achieve better results at work with the support of modern technologies such as LLM, NLP and AI. It is essential to develop critical thinking and logic, which are disciplines that are almost unknown and even undesirable in our educational system. Without critical thinking and knowing the advantages and disadvantages of technology, there is no progress and catching up with developed countries.

6. Conclusion

Artificial intelligence has the potential to replace boring and repetitive jobs, just as robots have replaced manual labour. The role of the software developer is changing drastically with the advent of AI: instead of coding, the programmer will control artificially generated code, he/she will have to develop creative thinking, intuition, and strengthen expertise in a certain domain. All of these are new skills that must be included in the list of competencies for modern engineers on which Industry 4.0 rests. AI will not replace but complement software developers. It should be noted that LLM is not a real artificial intelligence, it is just an excellent autocomplete tool or a top search engine. True AI (Strong AI) emerges only when an autonomous system can surpass human abilities in most economically valuable tasks.

7. References

- [1] Gu, J., Han, Z., Chen, S., Beirami, A., He, B., Zhang, G., ... & Torr, P. (2023). A systematic survey of prompt engineering on vision-language foundation models. arXiv preprint arXiv: 2307.12980 <https://doi.org/10.48550/arXiv.2307.12980>
- [2] Strobelt, H., Webson, A., Sanh, V., Hoover, B., Beyer, J., Pfister, H., & Rush, A. M. (2022). Interactive and visual prompt engineering for ad-hoc task adaptation with large language models. *IEEE transactions on visualization and computer graphics*, 29(1), 1146-1156. <https://doi.org/10.48550/arXiv.2208.07852>
- [3] Marvin, G., Hellen, N., Jjingo, D., Nakatumba-Nabende, J. (2024). Prompt Engineering in Large Language Models. In: Jacob, I.J., Piramuthu, S., Falkowski-Gilski, P. (eds) *Data Intelligence and Cognitive Informatics. ICDICI 2023. Algorithms for Intelligent Systems*. Springer, Singapore. https://doi.org/10.1007/978-981-99-7962-2_30
- [4] Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. arXiv preprint arXiv: 2402.07927. <https://doi.org/10.48550/arXiv.2402.07927>
- [5] Ekin, S. (2023). Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. Authorea Preprints. <http://dx.doi.org/10.36227/techrxiv.22683919>
- [6] Velásquez-Henao, J. D., Franco-Cardona, C. J., & Cadavid-Higuaita, L. (2023). Prompt Engineering: a methodology for optimizing interactions with AI-Language Models in the field of engineering. *Dyna*, 90(230), 9-17. <https://doi.org/10.15446/dyna.v90n230.111700>
- [7] Shin, J., Tang, C., Mohati, T., Nayebi, M., Wang, S., & Hemmati, H. (2023). Prompt Engineering or Fine Tuning: An Empirical Assessment of Large Language Models in Automated Software Engineering Tasks. arXiv preprint arXiv:2310.10508. <https://doi.org/10.48550/arXiv.2310.10508>
- [8] Korzynski, P., Mazurek, G., Krzypkowska, P., & Kurasinski, A. (2023). Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. *Entrepreneurial Business and Economics Review*, 11(3), 25-37. <https://doi.org/10.15678/EBER.2023.110302>

- [9] López-Riobóo-Botana, I., Gallent-Iglesias, D., & Gonzalez-Vázquez, S. (2023). QUA4I: Question Answering for the Industry 4.0 Domain. An Application of Intelligent Virtual Assistants. <https://ceur-ws.org/Vol-3516/paper17.pdf>
- [10] Tinnes, C., Ristin, M., Hohenstein, U., Fathi, K., & van de Venn, H. W. (2024). From Unstructured Product Descriptions to Structured Data for Industry 4.0 with ChatGPT. International Conference on Industrial Cyber-Physical Systems (ICPS)
- [11] Blasek, N., Eichenmüller, K., Ernst, B., Götz, N., Nast, B., & Sandkuhl, K. (2023). Large language models in requirements engineering for digital twins. 2nd International Workshop on Digital Twin Engineering (DTE), Technical University Vienna, Vienna, Austria
- [12] Howell, J. (2024) How Prompt Engineering is Revolutionizing the Manufacturing Industry?, <https://101blockchains.com/prompt-engineering-in-manufacturing/> [Dostupno: 11.06.2024]
- [13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30. <https://doi.org/10.48550/arXiv.1706.03762>
- [14] Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will AI exceed human performance. Evidence from AI experts. JArtifIntellRes62, 729-754. <https://doi.org/10.1613/jair.1.11222>
- [15] European Commission (2021). AI Act <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> [Dostupno: 14.06.2024]

Machine Safety and Workplace Implications of Artificial Intelligence

Viktorio Malisa^{*1}

Abstract: *Artificial intelligence (AI) is giving an additional boost to the already rapid pace of technological progress. The rapid development of new technologies has an impact on machine safety and changes jobs. With a series of laws, directives and regulations, the EU is trying to get a grip on negative developments and use new technologies in a humane way. At the same time, industry is being called upon to acquire additional expertise in new areas of digitalisation. Established institutions around the world are working on developing new standards to facilitate the implementation of regulations. The new EU Machinery Regulation covers digitalisation and defines the use of AI or machine learning. Generative AI (GenAI) is driving prompt engineering and this technology is finding its way into machine development and other production processes. One challenge that remains is how to provide proof of safety for AI in automation technology and especially for autonomous systems.*

Keywords: *European legislation, machine regulation, artificial intelligence, AI norms, prompt engineering*

1. Introduction

Industry 4.0 has initiated the digitalization of everything from machines to the entire production process, the Internet of Things (IoT) has triggered the global networking of software, devices and machines, and now a new field of research is coming into focus: artificial intelligence. AI research reports new algorithms, models and use cases almost daily. New AI-optimized hardware architectures are also constantly being announced. With the Assisted, Augmented and Autonomous Intelligence groups, artificial intelligence is conquering all technological disciplines of Industry 4.0: from the small assembly line to AI that provides humans with supporting information for better decisions, right through to fully autonomous, AI-controlled systems.

Parallel to AI research and development, legal and normative regulations are being developed in various countries and by various institutes and associations. Of particular importance is European Union legislation that also affects global IT companies operating in the EU market, such as the EU-AI Act (see Table 1). Table 1 lists the most important EU legislation on digitalization that relates directly to AI or has an indirect impact on AI. AI is also software that collects

^{*1}F-AR Förderung der Automation und Robotik, Vienna, Austria
Email: malisa@f-ar.at

and processes data, and therefore regulations on the protection of personal data, cyber-attacks, dissemination of information, etc. are relevant.

Table 1. EU legislation on digitalisation that relates directly or indirectly to AI systems [2]

EU legislation	Content
General Data Protection 2016/679 Regulation (GDPR)	Strict rules for the collection, storage, and processing of personal data, rights of data subjects, obligations of data processors
Free Flow of Data Regulation (EU) 2018/1807 (FFoD Regulation)	Data transfer of non-personal data across national borders within the European Union, including data for (professional) users of data processing services
Public Sector Information Directive (EU) 2019/1024 (PSI Directive)	The Public Sector Information Directive regulates the use of publicly funded data for commercial and non-commercial purposes in the European Union.
Directive on Digital Content and Digital Services (EU) 2019/770 (DI Directive)	Harmonised regulatory framework for the distribution of digital content and services
Directive on measures for a high common level of cybersecurity (EU) 2022/2555 (NIS-2 Directive)	Sets requirements for cybersecurity measures and reporting for operators of essential services and digital service providers
Digital Markets Act (EU) 2022/1925 (DMA)	Targets large online platforms ("gatekeepers"), sets rules to prevent market abuse and unfair practices
Digital Services Act (EU) 2022/2065 (DSA)	Regulates handling of illegal content, liability of intermediaries, transparency of advertisements, protection of user rights in the digital space
Data Governance Act (EU) 2022/868 (DGA)	Framework for sharing data between businesses, public administrations, and citizens, considering data protection and security
Artificial Intelligence Act (EU) 2021/0106 (AI Regulation-E)	Classification of AI systems based on risk (minimal, limited, high, unacceptable) and specific requirements and obligations for the development and use of high-risk AI applications.
Cyber Resilience Act (EU) 2022/0272 (CRA-E)	Requirements for cybersecurity measures and reporting for operators of essential services and digital service providers, promoting resilience to cyberattacks.
Data Act (EU) 2022/0047 (DA-E)	Regulates fair access to and use of data generated by connected devices and machines, promoting innovation and competition
Machinery Regulation (EU) 2023/1230	Sets requirements for the safety and compliance of machinery, considers the integration of AI and autonomous systems, emphasizes the need for measures to maintain inherent machine safety

In addition to EU legislation, guidelines and standards from established institutions are particularly important for the implementation and use of AI, such as:

- ISO - International Organization for Standardization [5]
ISO/IEC JTC 1/SC 42 Artificial Intelligence is a standardization subcommittee of the Joint Technical Committee ISO/IEC JTC 1 of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). To date, this subcommittee has published a total of 28 ISO norms on AI, with a further 33 under development. Table 2 lists some examples of norms that deal with security and safety.[5]

Table 2. *Examples of published AI norms [5]*

Norm Reference	Topic
ISO/IEC 22989:2022	Information technology — Artificial intelligence — Artificial intelligence concepts and terminology
ISO/IEC TR 17903:2024	Information technology — Artificial intelligence — Overview of machine learning computing devices
ISO/IEC 23894:2023	Information technology — Artificial intelligence — Guidance on risk management
ISO/IEC TR 5469:2024	Artificial intelligence — Functional safety and AI systems
ISO/IEC 5338:2023	Information technology — Artificial intelligence — AI system life cycle processes
ISO/IEC TR 24028:2020	Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence
ISO/IEC TR 24030:2024	Information technology — Artificial intelligence (AI) — Use cases

This is an example of a technical norm developed by ISO in cooperation with the International Electrotechnical Commission (IEC):

International norm ISO/IEC 42001, which is currently undergoing the EN standard process, defines the requirements for the introduction, implementation and continuous improvement of an AI management system in organisations that offer or use AI-based products or services. The standard addresses challenges such as ethics, transparency and continuous learning. The standard provides a structured approach to the management of AI projects, from risk assessment to risk management. [adapted 6]

- IEEE SA - Institute of Electrical and Electronics Engineers Standards Association: The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) has the goal to create specifications for certification and marking processes that advance transparency, accountability, and reduction in algorithmic bias in autonomous and intelligent systems. ECPAIS intends to offer a process and define a series of marks by which organizations can seek certifications for their processes around the A/IS products, systems, and services they provide. [9]
- DIN Deutsches Institut für Normung e.V. is the independent platform for standardization in Germany and worldwide. It uses norms and standards to describe what is technically possible, economically viable and proven in practice. [1]
- VDA Association of the German Automotive Industry, which deals with the safeguarding of AI, i.e. how the use of AI in vehicles can be proven to be safe. [1]

2. EU Machinery Regulation

The new Machinery Directive 2023/1230 will come into force in all member states of the European Union on 20 January 2027. This means that the essential health and safety requirements and conformity assessment procedures in particular must be applied uniformly for all stakeholders in the EU and there is no scope for divergent implementation, as is currently the case with the Machinery Directives. The Machinery Directive 2006/42/EC [4] applies to the manufacture of machinery and will remain in force until 19 January 2027.

The new Machinery Regulation also takes into account new risks arising from the integration and connection of digital technologies with machinery, such as the implementation of software that utilises machine learning or software that ensures safety functions.

The new Machinery Regulation defines that if the machine is connected: [3]

- to the Internet;
- to a device (data carrier, e.g. USB stick or other devices);
- with remote devices, i.e. via the Internet;

this must not lead to dangerous situations.

Software that fulfils a safety function of the machine and is placed on the market separately is considered a safety component and is subject to the certification procedure for safety components.

The risk assessment of a machine in accordance with the new regulation must also take the following points into account: [3]

- effects of software updates on the safety of the machine;
- future updates to the software installed in the machine;
- use of artificial intelligence in the execution of a safety function;
- Training the AI with training data on the machine

If the AI controls a machine function that may affect the safety of the machine, the AI is monitored by the Safety Monitoring System. This system continuously monitors the operating parameters and the results of the AI and intervenes if the defined safety limits are exceeded in order to put the machine in a safe state. In the event of a machine stop triggered by the safety monitoring system, the reason that led to the stop should be displayed. An Explainable AI is to be integrated for this purpose in order to transparently display the reasons for the machine stop. One example of an AI function in an industrial robot is the optimisation of robot movements within a defined range. If the defined range is exceeded, the safety monitoring system must stop the machine. However, the certification of an AI for use in the area of functional safety is still a field of research at the current state of the art.

The new Machinery Regulation provides for a type examination of machinery by a notified body if [amended after 4].

- Safety components whose behaviour is wholly or partly self-developed by machine learning techniques and which guarantee safety functions.
- Machinery with embedded systems whose behaviour is wholly or partly self-developed through the application of machine learning techniques and which ensure the safety functions and which are not placed on the market separately but only together with the embedded system.

When AI is used in autonomous machines, e.g. autonomous mobile robots with automatically determined routes, the system must be designed in such a way that it does not cause

- machines or associated products to act beyond their defined task and range of movement;[4].
This means that the limit values defined for the machine must not be exceeded and that monitoring of the limit values must be provided.
- The record of the data relating to the safety-related decision-making process to demonstrate the conformity of the machinery must be kept at the reasoned request of a competent national authority;[4].
This means that the conformity of the machinery can also be subsequently withdrawn due to the negative influence of further developed or no longer safe artificial intelligence. In this case, the

machinery must be decommissioned until a new conformity procedure has been carried out.

- A correction of the machine or the associated product to maintain the inherent safety is possible at any time. [4]
In the case of autonomous robots, it must be possible to intervene in the machine control system at any time in order to change or overwrite the results of the AI. This can be done automatically by an external safety monitoring system or manually by an operator, for example. If manual correction of the AI results is planned, this requires that the operating personnel are trained for this workstation and are familiar with the machine, the process and the application of the AI.

3. Prompting Engineering

As prompt engineering is being learnt through its current use in search engines and through the use of generative artificial intelligence (GenAI), it is to be expected that this technology will also be widely used in the development of automation solutions. The ambiguity of human language, which often leads to undesirable results in generative AI when prompting, is clearer in mechanical and electrical engineering - not least due to pronounced legal and normative regulations - and is therefore more suitable for AI in technical applications. Prompts are already being used today in the use of technical documentation and in the creation of control programmes. It is therefore conceivable that the prompting technology used today will be integrated into the further life cycles of automation in an adapted form.

Prompt engineering can be used from the creation of the layout of a machine through to complete production in a factory. However, the definition of prompt engineering "is the process by which solutions are controlled with generative artificial intelligence (generative AI) to achieve the desired results" [8] must be adapted for reasons of safety and technical feasibility in the automation technology of a production, e.g: Prompt engineering in automation technology is the process in which solutions for the automation of a production are controlled with artificial intelligence in such a way that they achieve the desired results according to the state of the art.

The development of prompt engineering for the design of automation systems is about precise input and an AI with training data that can rule out incorrect input from the outset.

What applies to ChatGPT "*Express yourself precisely and give context*" [7] would apply to automation technology: "*Enter precise information with sufficient parameters.*"

Such a prompt during layout creation in automation technology could contain the following information:

- System consisting of the system elements (machine types, incomplete machine, sensors, protective elements, etc.)
- Process description with process parameters and automation corner points
- Workpieces and tools with their parameters
- Limits of the respective system element (e.g. machine with work area, performance, control system, etc.)
- Environmental parameters and internal company standards for defined machines and processes

Software tools for design, layout creation and simulation are particularly suitable for generating variants and therefore data that can be used for training AI.

In CAD or simulation tools, it is possible to generate models or initial layouts using prompts. In a CAD tool, such a prompt could look like this:

"Create an articulated arm robot on an area of 10 x 20 m: payload 25 kg, range 1800 mm, which picks up the workpieces: round blanks with a weight of 2 kg and the dimensions WxL 20 x 250 mm from a table height 730 mm with a work surface of 1 x 2 m with a gripper and positions them in the CNC machine".

The result should be a layout of a system in which the following system parts are drawn in a hall: an industrial robot with a gripper, a table in front of it and a CNC machine to the side, a protective grid with a door around it and a control cabinet for the robot and a control cabinet for the CNC machine outside the protective grid. Based on this layout, the technical draughtsman can then move or change each element and use the design for the meeting and then make further adjustments and changes.

For a simulation, the basic control programme of the robot, the CNC machine and the higher-level PLC, including the control panel with start and stop buttons, could be created automatically. The networking between the robot, CNC and PLC should also be created. In the safety part of the PLC, the emergency stop button and the safety switches on the safety gate should be integrated into the control programme.

The following applies to generative AI: *"Avoid bias in the prompts, as the results will usually contain bias."* [7] should be avoided in engineering for reasons of machine and occupational safety. Therefore, one challenge for the future is to develop an AI for automation systems in such a way that it can intercept such and similar incorrect inputs as *"Create a layout with a CNC machine for welding sheet metal on a 1 x 2 m table"*, as a CNC machine is not intended for welding.

It is also important that only trained automation technicians work with these tools and that the evaluation of the results of AI in automation may only be carried out by experienced automation technicians.

3. Conclusion

In recent years, the European Union has responded to the rapid development of digitalization with new regulations to define the impact on people, the economy and society as a whole. The new regulations also lead to a burden on the economy, not just the European economy, but also have an impact on globally operating corporations. The new EU Machinery Regulation takes digitalization and AI into account. However, how the safety of AI can be proven is still the subject of research. A solution is currently being sought in which the behavior of the AI is monitored externally. Generative AI is changing engineering through the introduction of prompt engineering. In particular, software tools with integrated prompts in design, simulation and programming make work easier and increase efficiency.

4. References

- [1] DIN, 2024. *Deutsches Institut für Normung*. Wikipedia, https://en.wikipedia.org/wiki/Deutsches_Institut_f%C3%BCr_Normung [accessed 11.07.2024]
- [2] Piltz Legal, 2024. *Wichtige aktuelle EU-Digitalgesetzgebung*. <https://www.piltz.legal/news-blog/eu-digitalgesetzgebung> [accessed 12.06.2024]
- [3] IHK, 2024. *CE-Kennzeichnung: Leitfaden neue EU-Maschinenverordnung*. <https://www.produktentwicklung.ihk.de/produktmarken/ce-kennzeichnung/leitfaden-eu-maschinenverordnung-5549614> [accessed 25.05.2024]
- [4] EU-Parlament und Rat, 2023. *Verordnung 2023/1230 über Maschinen*. <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32023R1230> [accessed 30.05.2024]
- [5] ISO/IEC JTC 1, 2024. *ISO/IEC JTC 1/SC 42 Artificial intelligence*. <https://www.iso.org/standard/81230.html> [accessed 18.06.2024]
- [6] ISO/IEC 42001:2023. *Informationstechnik – Künstliche Intelligenz - Managementsystem*. Online: <https://www.dinmedia.de/de/norm/iso-iec-42001/376160445> [accessed 18.06.2024]
- [7] Solis, T., 2023. *So erstellst du gute ChatGPT-Prompts*, Online: <https://www.scribbr.de/ki-tools-nutzen/gute-chatgpt-prompts/> [accessed 08.06.2024]
- [8] AWS, 2024. *Was ist Prompt-Engineering?*. Online: <https://aws.amazon.com/de/what-is/prompt-engineering/> [accessed 09.07.2024]
- [9] IEEE SA, 2024. *THE IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS*. Online: <https://standards.ieee.org/industry-connections/ec/autonomous-systems/> [accessed 11.07.2024]

Trend Innovation of Artificial Intelligence and Robotic Technology: Implementation in Advanced Robotic Systems

Isak Karabegović^{*1}, Ermin Husak², Samir Vojić², Edina Karabegović²,
Mehmed Mahmić²

Abstract: *In the last ten years, the development and research of advanced technologies, as well as their application in all segments of society, have led to major changes and reshaping of the new world. New innovations are occurring on a daily basis, but their application is not going fast enough due to the rigid infrastructure. However, in order to secure an optimal future, we all have to adapt to the changes that are coming. The developed countries have adopted the strict implementation of advanced technologies of Industry 4.0, some of which include: Internet of Things (IoT), Big Data, Cloud Computing, smart sensors, Radio Frequency Identification (RFID), 3D printing, advanced security systems, Virtual and Augmented Reality (VAR), etc. Robotics is the basic and first technology that has been implemented since the 60s of the last century, with artificial intelligence coming in the spotlight in the last ten years. Artificial intelligence is becoming a key to the development of advanced robots, as it enables them to adapt to unpredictable situations, to learn from experience and make intelligent decisions. Robots use AI to process sensor data, navigate, recognize objects, plan paths and interact with the environment. In short, artificial intelligence enables robots to be smart, whereas robotics uses AI to create autonomous and useful devices. This symbiosis contributes to progress in many industries, including healthcare, manufacturing and transportation. Artificial intelligence (AI) and robotics are two key fields that complement each other. The paper presents the trend of applied and approved patents in artificial intelligence and robotics, as well as an example of the use of artificial intelligence in advanced robots to perform certain tasks. Artificial intelligence (AI) is having an increasing impact on robotics, opening up many possibilities.*

Keywords: *artificial intelligence, advanced robots, patents, application, Industry 4.0*

1. Introduction

Artificial intelligence (AI) and robotic technology are currently in the center of interest of researchers throughout the world. Artificial intelligence is broadly defined and deals with intelligent behavior, perception, reasoning, learning as well as navigating and acting in a complex environment. By using artificial intelligence in robotic technology, we aim for all robots involved in all segments

^{*1}Academy of Sciences and Arts of Bosnina, Sarajevo, Bosnia and Herzegovina

²University of Bihać, Bihać, Bosnia and Herzegovina

E-mail: isak1910@hotmail.com

around us to be intelligent, to have perception, to learn, to reason and navigate in a very complex environment, and to make decisions independently. The very implementation of industrial robots in production processes began in the 1960s, when the idea of automating appeared, primarily in the automotive industry. The roots of the formalization of human thinking and reasoning come from ancient Greek times. For example, Aristotle defined three basic laws of thought and syllogism, and later many important minds tried to create formal systems based on a combination of symbols. In 1950, Alain Turing (Turing's test - which requires the ability to process natural language, knowledge representation, reasoning and learning) laid the foundations of the philosophy of artificial intelligence. In 1955 Allen Newell and Herbert A. Simon created the first framework of artificial intelligence called Logic Theorist, while the authors Alfred N. Whitehead and Bertrand Russell demonstrated the use of this framework through 40 hypotheses *Principia Mathematica* [1-3]. At the conference in Dartmouth, a group of researchers defined the field called "artificial intelligence", which marks the beginning of research in this field. As a scientific field, like all other fields, it has been developing until today. However, with the development of the basic technologies of Industry 4.0 such as: Things (IoT), Big Data, Cloud Computing, Smart Sensors, Radio Frequency Identification (RFID), computers with increased processing power encourage its research and growth with a large amount of digitized data. Artificial intelligence is moving from theoretical research to the global market using a large amount of data based on which it conveys a revolutionary effect [4,5]. Its implementation can improve productivity in industry, and can make significant changes in medicine, agriculture, meteorology and other segments of our environment. Even though AI is becoming a part of our lives today, we are still at the very beginning and we must discover what impact it will have on society, culture and business. Artificial intelligence is currently the light of transformation in various industrial branches and sectors, because it changes the way we work, buy, communicate with each other, travel, etc. With the emergence of Big Data technology, i.e., vast amount of information and artificial intelligence, both industrial and service robots are becoming more and more reliable, and their application in expanding the tasks they perform increases with human supervision. In other words, they are introduced to the tasks that were previously performed by workers [6,7]. Artificial intelligence is very interesting for robotic technology because of intelligent automation in production processes, which aims to make production processes intelligent.

2. Trend of Registered and Approved Patents in the World

The global competition of companies worldwide depends on investments in the development and research of new technologies. This is best shown in the trends

of registered and approved technological innovations, i.e., patents submitted in the offices in charge of filing and approving patents. The most important offices for the application and approval of patents in the world are: EPO – European Patent Office, USPTO – United States Patent and Trademark Office, JPO – Japan Patent Office, KIPO – Korean Intellectual Property Office, and CNIPA – China National Intellectual Property Administration. The annual trend of registered patents in the world for the period 2013-2023 in the aforementioned patent offices is shown in Figure 1 [8, 9].

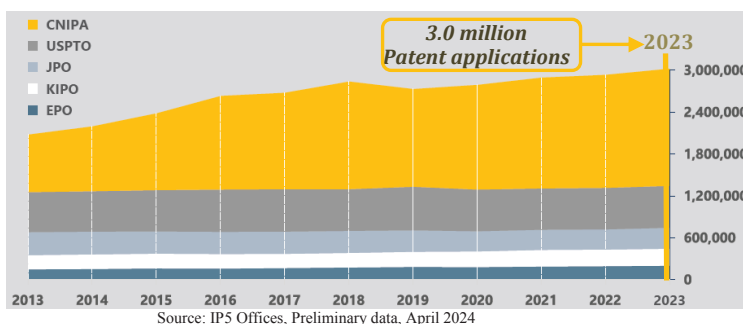


Figure 1. Trend of annually applied patents in the world for the period 2013-2023 in five patent offices in the countries of USA, EU, Japan, Republic of Korea and China [8, 9]

Based on Figure 1 we can conclude that the trend of patent applications in the world in the period 2013-2023 is continuously increasing. The number of applied patents in the world reached a value of about 3 million in 2023, as shown in Figure 1. Given that not all applied patents in patent offices were approved, Figure 2 shows the trend of approved patents in the world.

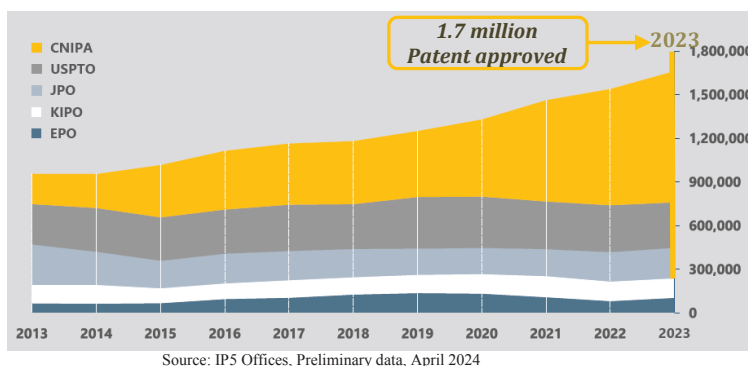
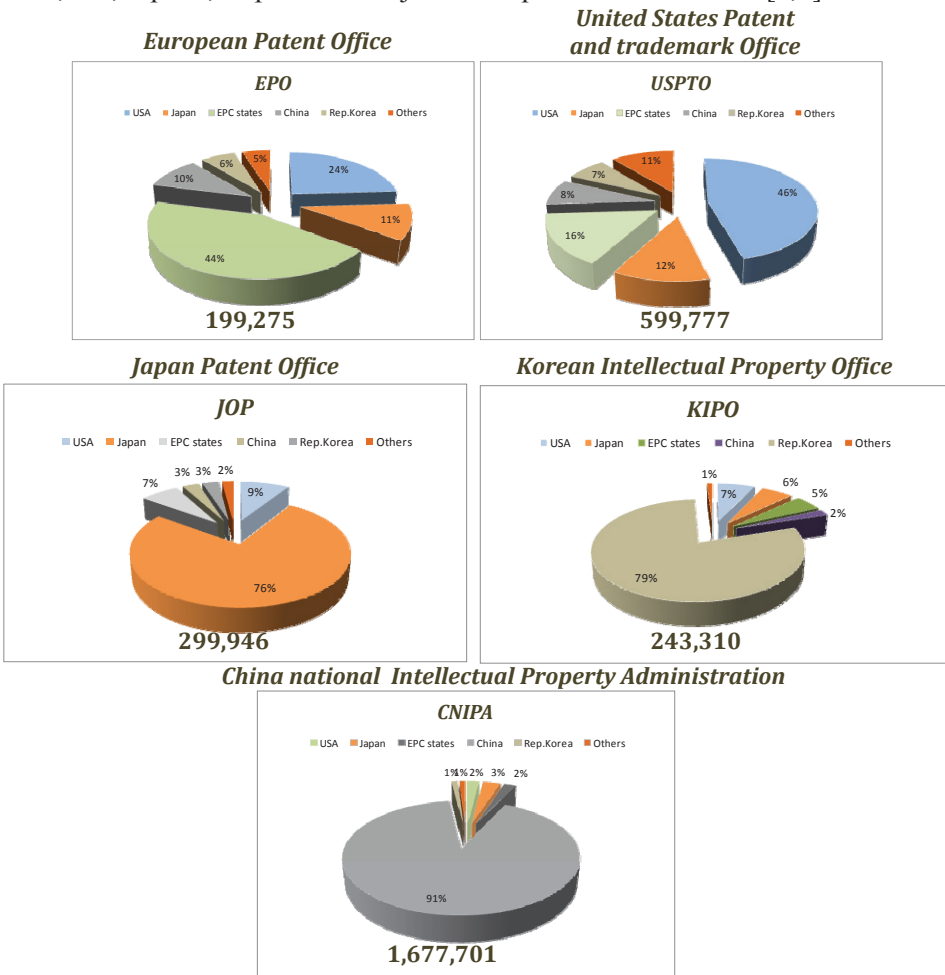


Figure 2. Trend of annually approved patents in the world for the period 2013 - 2023 in five patent offices in the USA, EU, Japan, Republic of Korea and China [8, 9]

Likewise, as shown by the annual trend of approved patents in the world for the period 2013-2023, five patent offices are continuously increasing their work, so that in 2023 the number of approved patents amounted to about 1.7 million. We come to the conclusion that out of 3 million registered patents in the world, around 1.7 million patents have been approved. The number of approved patents compared to those applied for is around 57%. The real state of the number of registered patents in the world will be obtained if we look at the number of registered patents in five patent offices located in the countries: USA, EU, Japan, Republic of Korea and China, as shown in Figure 3 [8,9].

USA, EU, Japanu, Republici Koreji i Kini a prikani su na slici 3. [8,9].



Source: IP5 Offices, Preliminary data, April 2024

Figure 3. Annual applications of patents in the world in 2023 by application of origin in five patent offices in the USA, EU, Japan, Republic of Korea and China [8,9]

The diagrams shown in Figure 3 illustrate the trend of registered patents in the world in five patent offices: EPO – European Patent Office, USPTO – United States Patent and Trademark Office, JOP – Japan Patent Office, KIPO – Korean Intellectual Property Office, and CNIPA – China National Intellectual Property Administration, based on which we can conclude that companies from one country apply for patents in other countries because they have their own companies that work and operate in those countries. If we examine Figure 3, we can see that if we rank the patent application offices by the number of most filed patents, we will get the following arrangement from highest to lowest:

- 1.677.701 registered patents in CNIPA – China National Intellectual Property Administration,
- 599.777 registered patents in the USPTO – United States Patent and Trademark Office,
- 299.946 registered patents in JOP – Japan Patent Office,
- 243.310 registered patents in KIPO – Korean Intellectual Property Office, and
- 199.275 registered patents in: EPO - European Patent Office.

We come to the conclusion that the highest number of registered patents in the world is located on the continent of Asia because it includes three patent application offices, namely CNIPA, JOP and KIPO.

The current state of research, the most represented scientific research in artificial intelligence and robotic technology, as well as the identification of research directions in the future will be obtained based on the biometric analysis of the trends of registered and approved patents in artificial intelligence and robotic technology. In the last ten years, the interest of the academic community, research institutes, and companies that strive to implement these two technologies in their production processes has been growing. The paper provides an analysis of the trend of registered and approved patents in artificial intelligence and robotic technology in the last ten years.

3. Trend of Applied and Approved Patents in Artificial Intelligence

Extensive research is being conducted all over the world when it comes to artificial intelligence. Huge amount of money is being invested in this research with the aim of creating an environment in which productivity can be increased for the purpose of reducing the production time that can be provided by the individual. The development and research of artificial intelligence reveals its application in every field, whether it is the production process in companies, logistics, accounting, assembly companies, agriculture or medicine. There is no segment of human life where AI cannot be used. The advantage of using artificial intelligence is that it can analyze a large amount of information which usually does not fit into any of the ordinary computer, much better than any

human. In recent years, there has been a sudden interest in artificial intelligence research, which is shown by the trend of approved patents in artificial intelligence shown in Figure 4.

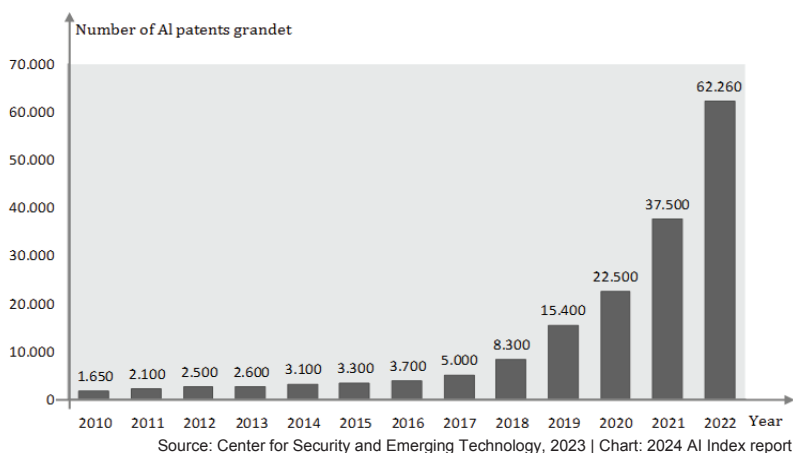


Figure 4. Diagram of the number of approved AI patents in the world for the period 2010-2022 [9-11]

The analysis of the diagram of the number of approved artificial intelligence patents shown in Figure 4 for the period 2010-2022, provides us with the conclusion that the number of approved patents in artificial intelligence has been growing rapidly since 2016. For example, between 2010 and 2015, the growth of registered patents in artificial intelligence was 2 times higher, whereas in the next five years in the period 2015-2020, the growth of registered patents in artificial intelligence was 6.8 times higher. In just one year, in the period 2021-2022, the trend of approved patents was 1.7 times higher. In twelve years, during the period between 2010 and 2022, the number of approved patents increased enormously by 37.7 times. Based on Figure 4, we come to the conclusion the trend of approved patents in artificial intelligence in the world had a sudden growth since 2015. In order to get more complete information about the number of patents in artificial intelligence, Figure 5 shows the trend of approved and unapproved patents in artificial intelligence.

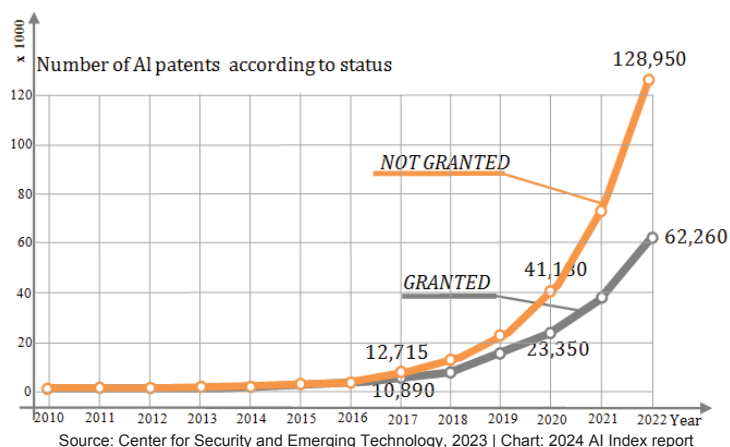


Figure 5. Diagram of the number of unapproved and approved AI patents in the world for the period 2010-2022 [9-11]

The diagram shown in Figure 5 gives us an insight into the trend of approved and unapproved patents in artificial intelligence in the period 2010-2022. As the diagrams in Figure 5 show, the number of approved patents has been increasing since 2016, so that in 2017, 10.890 artificial intelligence patents were approved, whereas 12.715 patents were not approved. In 2020, the number of approved artificial intelligence patents was 23.350, while far more patents in artificial intelligence were not approved, i.e., 41.130. In just one year, the number of patents that were not accepted in 2022 amounted to 128.950 patents, whereas 62.260 patents were approved. It can be seen that the number of unapproved patents is 2 times higher than the number of approved patents in artificial intelligence. We come to the conclusion that since 2016, the trend of the number of applied patents has followed an exponential function, including the number of approved and unapproved patents in artificial intelligence.

The gap between approved and unapproved artificial intelligence patents has been continuously increasing since 2016, as shown in Figure 5. To get a more complete insight in the number of approved and unapproved artificial intelligence patents, Figure 6 shows the trend of approved and unapproved artificial intelligence patents in countries that invest in its development and research.

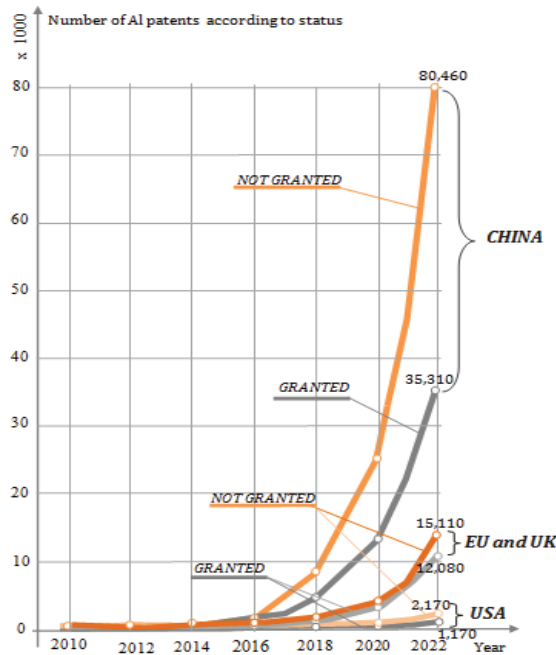


Figure 6. Diagram of the number of unapproved and approved patents in artificial intelligence in China, EU and UK, USA for the period 2010-2022 [9-11]

Figure 6 shows the trends of unapproved and approved artificial intelligence patents in China, European Union, Great Britain and USA in the period 2010-2022. Through the analysis, we can conclude that China ranks first in the world in terms of the number of approved and unapproved patents in artificial intelligence. As of 2015, the number of approved and unapproved patents in artificial intelligence has been continuously increasing in China. However the gap between approved and unapproved patents was also increasing. We can see that in 2022, China had 35.310 approved patents, but also 80.460 unapproved patents. The second place in terms of the number of approved and unapproved patents in artificial intelligence is held by the European Union and Great Britain. These countries have a small difference between approved and unapproved patents in artificial intelligence, so that in 2022 the number of approved patents amounted to 12.080, while there were 15.310 unapproved patents. The third place is occupied by the USA, with a large discrepancy between approved and unapproved patents in artificial intelligence in 2022. They had 1.170 approved patents and twice as many unapproved ones, that is, 2.170. In order to describe the actual situation of the countries that are investing in and developing artificial intelligence, we have shown the percentage of approved patents in relation to the

total number of approved patents in the world. The analysis includes the countries: China, EU, UK, USA, India and the rest of the world in the period from 2010 to 2022, as shown in diagrams in Figure 7.

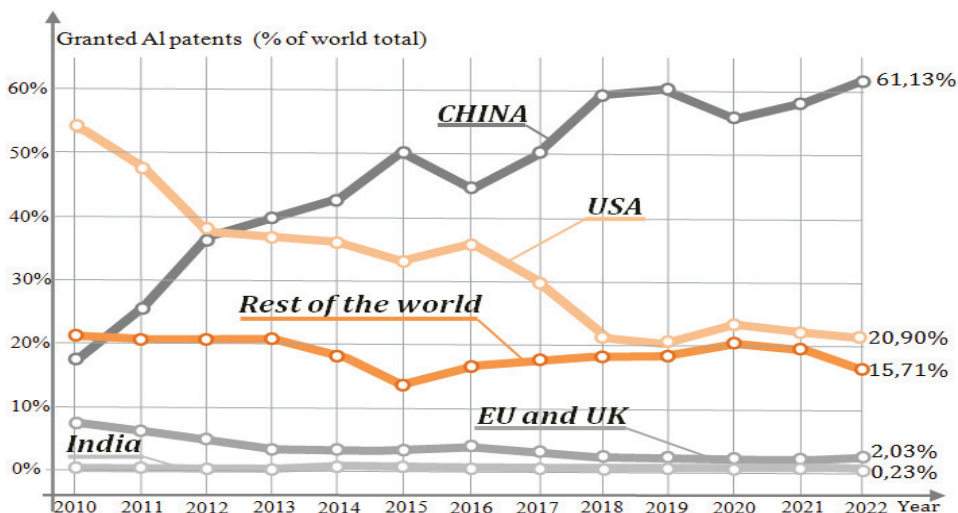
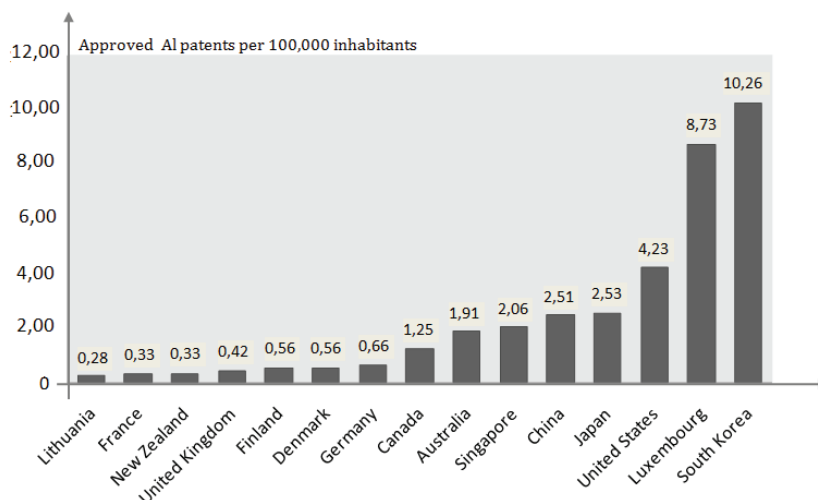


Figure 7. Diagram of approved patents in artificial intelligence in China, USA, EU and UK, India and the rest of the world in the period 2010-2022 [9-11]

By analyzing the diagram in Figure 7, we can conclude that up to 2012, the USA was the world leader whose percentage of approved patents in relation to the total number of approved patents in the world was the highest, so that in 2010 it amounted to about 54 %. In the same year, namely 2012, China had about 18 % of approved patents, followed by EU and UK with about 8.85 % of approved patents, India with about 0.15 % of approved patents, while other countries of the world had about 21 % of approved patents in relation to the total number of patents in artificial intelligence. The trend of approved patents was constantly changing in the period 2010-2022, as shown in Figure 7, so that the trend of approved patents in China was continuously increasing, while in the USA, as well as in the EU and the UK, it was continuously decreasing. As of 2012, China has taken the lead in terms of the percentage of approved patents in artificial intelligence in relation to the total number of approved patents in the world until today. In 2022 China had 61.13 %, followed by the USA with 20.90 %, the EU and the UK with 2.03 %, India with 0.23 % and rest of the world countries 15.71 % of approved patents. The USA, EU and UK recorded a decline in approved patents in the period from 2010 to 2022. We can note that the largest share of the world's approved patents in artificial intelligence originates from Asia, where

China is the leading country. Figure 8 shows which countries in the world are the leaders in the number of patents per capita.



Source: Center for Security and Emerging Technology, 2023 | Chart: 2024 AI Index report

Figure 8. Diagram of the number of approved patents in fifteen leading countries in the world per 100.000 inhabitants in 2022 [9-11]

Based on the diagram shown in Figure 8, we see that out of fifteen countries in the world in 2022, five countries are leading in terms of the number of approved patents per capita, namely the Republic of Korea, Luxemburg, the USA, Japan and China. The first place is held by the Republic of Korea with 10.26 approved patents per 100.000 inhabitants, followed by Luxembourg with 8.73, the USA with 4.23, Japan with 2.53 and China with 2.51 approved artificial intelligence patents per 100.000 inhabitants. The other countries ranked by the number of approved patents per capita are: Singapore, Australia, Canada, Germany, Denmark, Finland, the United Kingdom, New Zealand, France and Lithuania, with the values shown in the diagram. We come to the conclusion that the largest number of approved patents in artificial intelligence in the first five countries is coming from Asia, mostly owing to the Republic of Korea, Japan and China. It is expected that this trend of approved artificial intelligence patents will continue in the coming years, and that China will join the leading country in the world, namely the Republic of Korea. China is implementing its National Program for research, development and implementation of high technologies named “863 Program” which is implementing both specific goals and the long-term goals in order to transform China into a global leader in future production.

4. The Trend of Applied and Approved Patents in Robotics

Robotic technology, one of the basic technologies of Industry 4.0, leads us in the direction of intelligent automation, even though its research and development has potential impacts in other areas, such as healthcare, agriculture to security. Advanced robotics refers to specific design characteristics as well as the ability to:

- be autonomous to work in an unstructured or unsafe environment,
- be mobile and cooperate with other workers in the production process,
- be designed to manipulate or physically interact with their environment,
- be trained to achieve the desired results without the need for reprogramming,
- safely perform tasks in intimate operations with people or in extremely dangerous environments.

In order to provide all these characteristics of advanced robotics, it is necessary to continuously work on research and development of robotic technology, as well as technologies that represent support for robotic technology. In the last fifteen years, the demand for robotic technology has been increasing, as companies are hastily implementing it in their production processes in order to be competitive on the global market. The advancements in development and research in autonomous navigation, smart sensors, computer vision, and machine learning technology help create more mobile robots, which are more agile and collaborative, as their awareness of the environment in which they operate increases their implementation. In order to have an insight in the research and development of robotic technology, we must follow the trend of applied and approved patents in robotic technology. Figure 9 shows the trend of first patent applications in the field of robotics by origin and type of applicant in the period 1960–2011 [10,11].

The analysis of the trend of approved patents since the very beginning of the application of the first robotics patent in the 1960s has shown that the number of applications for robotics patents has been continuously increasing until today. We must note that since 2000, the number of patents began to increase not only in Japan, but also the Republic of Korea and China, which can clearly be seen from the diagram shown in Figure 9.

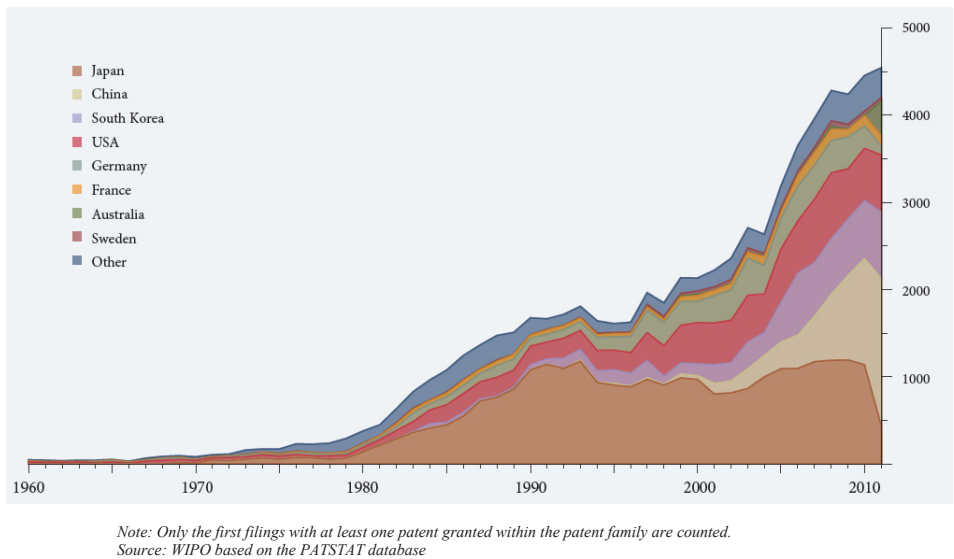


Figure 9. Trend of patent applications in robotic technology in the period 1960-2010 in countries that focus on research and development of robotic technology in the world [10, 11]

The basis for applying patents in robotic technology is the development of the automotive and electrical/electronic industry, because in order for companies to be competitive on the market, they must invest in research and development due to the increase in quality and productivity by implementing robots in their production processes. In recent years, increasing number of Internet companies are entering the field of robotics, employing experts in robotic technology and entering the market with new robotic systems.

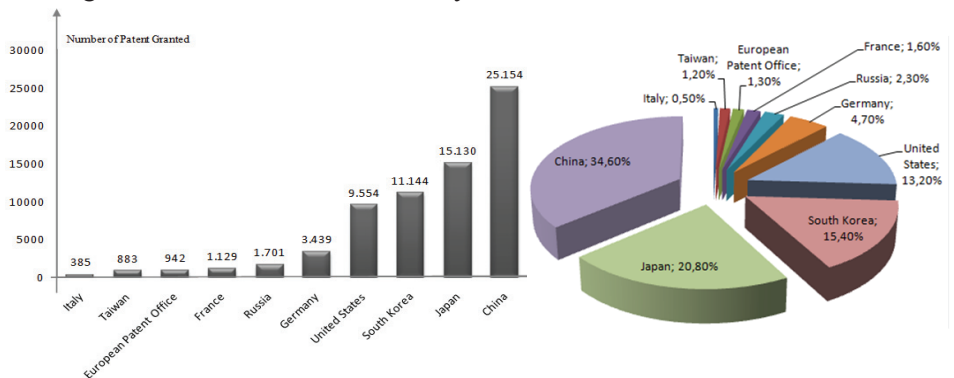


Figure 10. Approved patents in robotic technology in ten countries in the world in the period 2005-2019 [12-15]

Likewise, as of 2010, the percentages and numbers of registered patents change from country to country depending on the national strategy of that country. The real insight into the development and research in robotic technology in the world can be seen in the analysis of patent applications in the top ten countries in the world, as shown in Figure 10 [9,10-12]. When applying for robotics patent, a patent can be an underwater robotics patent, a manufacturing robotics patent, a mobile robotics patent, a military robotics patent, robotics with artificial intelligence, or a patent involving the physical components of the robot and control system characteristics. Figure 10 includes all the approved patents in robotic technology previously listed in the top ten countries in the world.

The top ten countries in the world include those countries that have companies for the development, research and production of robots, and have developed automotive and electrical/electronic industries, since it is known that these two technologies implement more than 50 % of all robots produced in one year. These countries are the following: China, Japan, USA, Italy, Taiwan, France, Russia, Germany, and Republic of Korea. If we examine the number of approved patents in robotic technology for the period 2005-2019, we can see that China holds the first place with 25.154 approved patents, which represents 34.6 % of the total number of approved patents, followed by Japan with 15.130 approved patents or 20.8 %, South Korea with 11.144 approved patents or 15,4 %, United States with 9.554 approved patents or 13.2 % and finally Germany with 3.439 or 4.7 % f approved patents in robotic technology. The following top five countries in terms of the number of approved patents in robotic technology, ordered by size, are the following: Russia (1.701), France (1.129), European Patent Office (943), Taiwan (883) and Italy (385). The analysis of approved patents in robotic technology in 2019 gives a clearer picture of research, development and improvement of advanced technology in countries that develop robotics. The analysis is shown in Figure 11 [12-15].

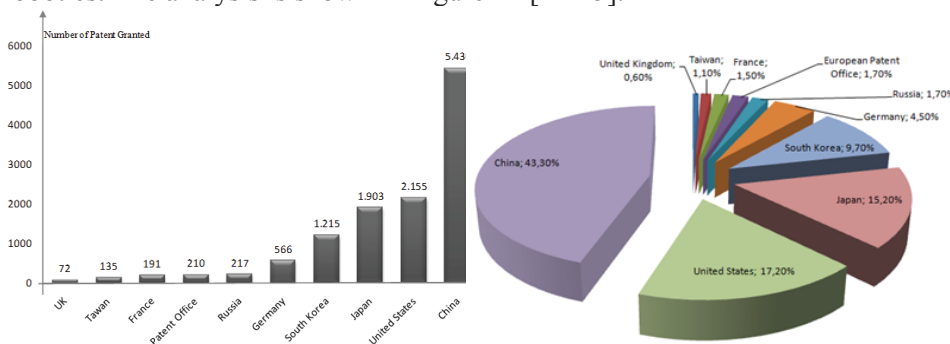


Figure 11. Approved patents in robotic technology in 2019 in ten top countries in the world [12-15]

The leading country in terms of the number of approved patents in 2019 was China with 5.410 patents in robotics, followed by the USA with 2.155 approved patents. Next in the line are: Japan with 1.903, Republic of Korea with 1.215 and Germany with 566 approved patents in robotic technologies. In terms of percentage, the ranking goes as follows: China 43.3%, USA 17.2%, Japan 15.3%, Republic of Korea 9.7% and Germany 4.5% of approved patents in robot technology. The trend of approved patents in robotic technology in these countries in the period 2005-2019 is shown in Figure 12 [12-15,16]. As we have already mentioned by analyzing Figure 9, as of 2000, in addition to Japan from Asia, the increasing trend of approved patents originates from China and the Republic of Korea, and continues to grow until today. China has become the leader when it comes to the number of applied and approved patents in robotics technology in the world and remains so until today. China is also the leader in development and research in robotic technology, as shown in Figure 12 [9-11]. Nowadays, many companies face global competition. However, the solution to competitiveness on the global market is based on new technologies that offer answers. Only those technologies that deal with the growth of production and productivity are able to significantly improve the quality of production and enable the company to achieve an improvement in its competitive advantage, in other words to make the company successful

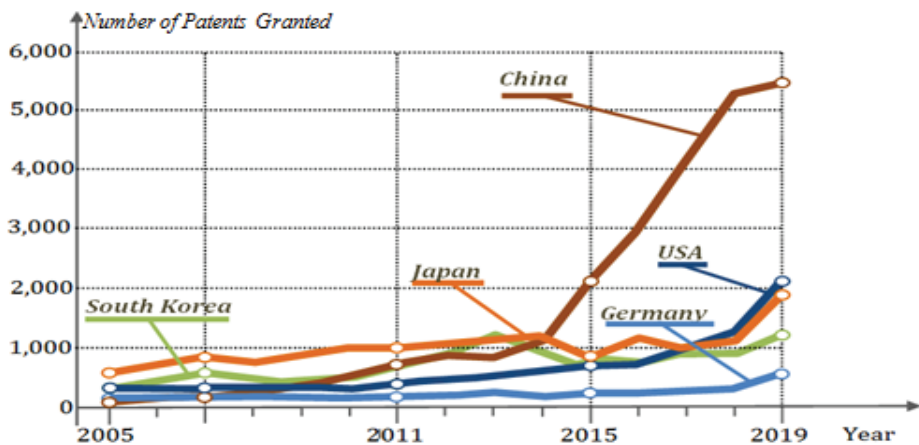


Figure 12. Approved patents in robotic technology in China, Japan, Republic of Korea, USA and Germany in the period 2005-2019

The company can achieve all of the above by implementing Industry 4.0 technologies, which include: Internet of Things (IoT), Big Data, Cloud Computing, smart sensors, Radio Frequency Identification (RFID), 3D printing, advanced security systems, virtual and augmented reality (AR), etc. In addition to the above-mentioned technologies, the basic technology is robotic technology,

without which there is no flexible automation, reduction of production costs, higher quality, increased productivity and other steps, from the delivery of raw materials to the transportation of products to the customer, and recently, the successful implementation of artificial intelligence. By implementing all the mentioned Industry 4.0 technologies, the company should be competitive with other companies on the global market. To analyze the role of advanced robotics in industry, it is necessary to explain the difference between conventional robots, that is, robots of the first generation in relation to advanced robots, that is, robots of the second generation. The first-generation robots are surrounded by fences from workers so as not to injure them, and the best example is industrial robot that serves to perform certain operations. The second-generation robots, or advanced robots, are a combination of powerful hardware and sophisticated programming, using smart sensor technology (including ultrasonic, touch, and light sensors) that interact with their real-world environment to perform real-world tasks.

Advanced robotics is the robotics of the future, because advanced robots are becoming flexible, smart, intelligent, autonomous and efficient. The implementation of advanced robotics in industry will transform production processes. Advanced robots are equipped with new technologies so that they have programmed abilities to make decisions on their own, which allows them to make decisions independently and function and navigate in any new environment. By implementing advanced robots with the basic technologies of Industry 4.0, we transform the rigid automation of production processes into flexible production processes that are simplified, increase productivity, reduce errors, increase product accuracy and ensure the efficiency of the production process. With its implementation in industry, advanced robotics has a number of benefits, some of which are the following:

- increased productivity,
- increased speed,
- quality improvement,
- increased security,
- increased accuracy,
- increased agility.

The implementation of advanced robotics in production processes will enable it to understand the production process faster, become more intelligent, and busier. At the same time, manufacturers can use it in the configuration of new production processes that meet the growing demand for more product variations, customized products and redesigns. Currently, advanced robotics is in the focus of research because its capabilities have not yet been utilized, as its complexity and the depth of its implementation need to be understood. There are many technologies that will contribute to its greater use in all segments of society, and

one such technology is material technology, which has great potential because it is possible to find materials that will create soft grippers that are flexible and have the ability to grip a wider selection of objects. This technology is also known as soft robotics. Soft robotics is in the research trend and smart solutions are expected in the coming years in different technologies such as material technology, so that materials and design can be carefully studied and evaluated. Other technologies that will contribute to the advancement of advanced robotics are robotic hardware-smart sensors, machine learning and quantum computers.

5. Implementation of Artificial Intelligence in Advanced Robotics

Trend of approved patents in artificial intelligence and robotic technology in relation to other technologies in the period 2011-2019 is shown in Figure 13 [13, 14, 16].

By analyzing the diagrams shown in Figure 13, we conclude that the trend of artificial intelligence patents is experiencing significant growth since 2016 until today and is far greater than that of robotics technology. In 2019, it made an increase of about 140% compared to the previous year, while at the same robotic technology had an increase of about 32%.

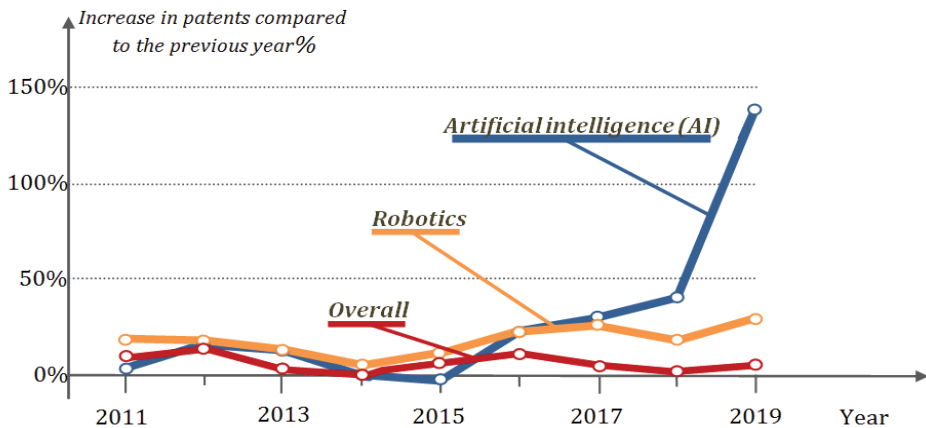


Figure 13. Comparative analysis of the percentage increase in patents from robotic technology and artificial intelligence compared to other technologies in the period 2011-2019

In recent years, artificial intelligence (AI) and computer vision have experienced an increase in patents compared to other technologies. The reason is that this is a newly emerging field of research, even though it had a lower number of patents than robotic technology until 2016. This trend of patent growth will continue in the years to come. The overall representation of approved patents in robotics

technology in the world is obtained by analyzing the patents of five top countries in the world in this technology, as shown in Figure 13 [13-14, 16].

The implementation of advanced robotics and artificial intelligence is the future, because advanced robots with artificial intelligence become flexible, smart, intelligent, autonomous and efficient. The implementation of advanced robotics and artificial intelligence in industry will transform production processes. Using artificial intelligence, advanced robots can have programmed abilities that allow them to make decisions independently because it enables them to function in any type of environment and navigate within that environment. By implementing advanced robots and artificial intelligence with the basic technologies of Industry 4.0, we transform the rigid automation of production processes into flexible production processes that are simplified, increase productivity, reduce errors, increase product accuracy and ensure the efficiency of the production process. As we mentioned, advanced robotics includes the entire spectrum of advanced technologies, all of which contribute to the improvement of industrial processes in all industrial branches. Processing, precision and speed are key characteristics in the field of robotics. High precision sensors, computer vision systems and advanced control algorithms are integral components [17]. Computer-aided design (CAD) models allow robots to be controlled with high precision, whereas force feedback sensors allow them to adjust their actions in real time to ensure optimal material removal rates and achieve results close to specialized CNC machines [18,19,21].

Let us mention one example of the use of advanced robotics and artificial intelligence, which was discussed by the authors [20]. This is the system architecture and communication between multiple devices, as shown in Figure 14. In the processing cell, an industrial robot with six DoF (FAUNC-M20iD/35) is used for the automatic grinding process, and a motor-driven gripper is installed at the end of the robot for accepting workpieces. In their paper, the authors [20,21], show the grinding, deburring and polishing processes that also benefit from robotization, especially through the use of modern end effectors. Force and torque sensors integrated into the end effectors provide feedback to the robot so that it can adapt its force application to material properties and surface conditions.

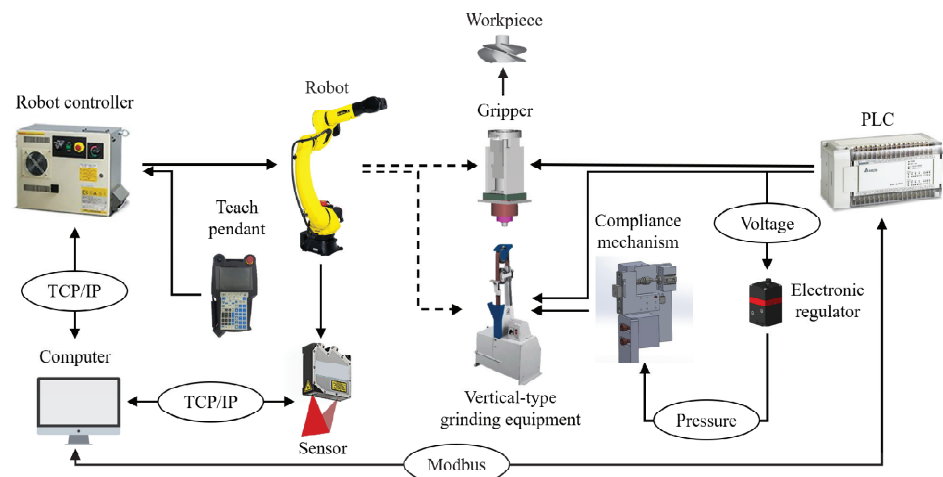


Figure 14. Schematic representation of the architecture and communication of an industrial robot with multiple devices – an example of grinding, deburring and polishing (obtained from [20,21])

Machine vision systems are used in quality control to check surfaces for defects and ensure that the product meets the required specifications. These technologies not only improve the efficiency of the machining process, but also contribute to the overall quality of the manufactured components.

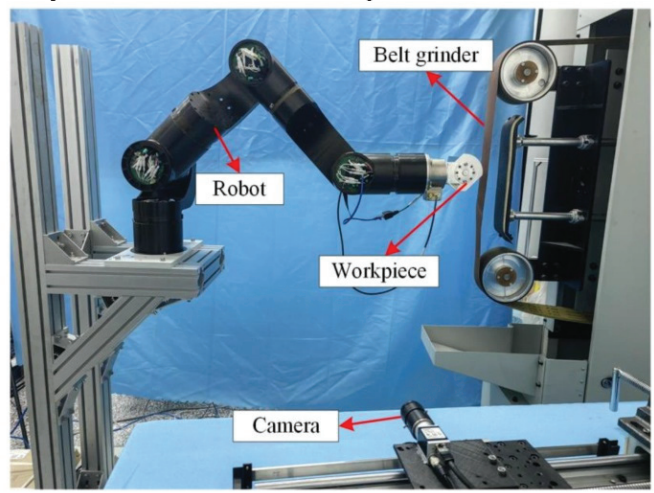


Figure 14. An example of controlling the roughness of the product surface with the use of an advanced robotic system and artificial intelligence during grinding of workpieces (obtained from [22])

When talking about the interaction between a human and a robot, the use of artificial intelligence (AI) offers innovative approaches to improve collaboration. One important capability is intelligent task assignment, where artificial intelligence algorithms evaluate human and robot performance data in real time and optimize productivity by assigning tasks tailored to individual strengths. Additionally, AI-powered robots can continuously learn from human behavior using adaptive learning systems to promote more intuitive and responsive collaboration. AI-based predictive maintenance uses sensor data to detect potential problems, ensure proactive interventions and reduce downtime. The implementation of natural language processing (NLP) enables contextual communication and therefore more seamless interaction between humans and robots. One of the biggest advantages of integrating artificial intelligence in human-robot collaboration is security monitoring. These systems play a key role by detecting threats or deviations from safety protocols and initiating interventions to ensure safe and efficient human-robot collaboration [21,23].

6. Conclusion

We can conclude that in the last ten years the number of registered and approved patents in artificial intelligence and robotics technology has been continuously growing worldwide. However, artificial intelligence has been in the first place among all technologies in recent years. As of 2000, the number of patents began to increase not only in Japan, but also the Republic of Korea and China, which is clearly visible from the diagrams presented in the paper. The basis for applying for patents in robotic technology is the development of the automotive and electrical/electronic industry, because in order for companies to be competitive on the market, they must invest in research and development in robotic technology to achieve an increase in quality and productivity by implementing robots in their production processes. We must single out China as the leading country in the world in terms of registered and approved patents in artificial intelligence and robotic technology. Artificial intelligence is becoming crucial for the development of advanced robots, as it enables advanced robots to adapt to unpredictable situations, to learn from experience and make intelligent decisions. Robots use artificial intelligence to process sensor data, navigate, recognize objects, plan paths and interact with the environment. In short, artificial intelligence enables robots to be smart, and robotics uses artificial intelligence to create autonomous and useful devices. This symbiosis contributes to progress in many industries, from healthcare to manufacturing and transportation. Artificial intelligence (AI) and robotics are two key fields that complement each other. Advanced robotic technology is enhanced by hardware and software innovations based on the basic technologies of Industry 4.0, so that it enables devices, machines and equipment to make independent decisions and

undertake further actions independently. The infrastructure of advanced robotics is based on Cloud Computing, Deep Learning, Big Data and the Internet of Things (IoT). In order to plan and create a path, advanced robotics uses a holistic data model that includes digital representations of the product and the production system along with the life cycle. The system of controls during the performance of the task for efficient management synchronizes the integration of digital technologies and workers in the production process. Advanced robotics increases efficiency in the production process, because solutions are made through self-learning technology and artificial intelligence, and familiar tools such as robotic process automation and intelligent process automation can also be included. The implemented advanced robotic systems in production systems providemobility, perception, integrability and adaptability. There has been a greater implementation of advanced robotic systems in production processes in recent years. Industries that use robotic systems can gain a competitive advantage by increasing productivity and efficiency. The legal and regulatory framework developed around artificial intelligence requires a responsible approach to implementation and emphasizes the need for ethical considerations when integrating advanced technologies. Future directions also include exploring new applications of artificial intelligence in robotics to push the boundaries of what can be achieved beyond human expertise. The development of technologies to improve precision, machine learning algorithms and collaborative robots will contribute to the development of increasingly sophisticated and adaptive robotic systems.

7. References

- [1] J. Manyika, M. Chui, M. Miremadi et al., “*A future that works: AI, automation, employment, and productivity*,” *McKinsey Global Institute Research Technical Report D*, vol. 60, pp. 1– 135, 2017
- [2] Schwab, K.: *The Fourth Industrial Revolution*, World Economic Forum, Geneva, Switzerland, 2016.
- [3] Karabegović, I., Kovačević, A., Banjanović-Mehmedović, L., Dašić, P.: *Integrating Industry 4.0 in Busniess and Manufacturing*, IGI Global, Hershey, PA, USA, 2020. <https://www.igi-global.com/book/handbook-research-integrating-industry-business/237834>
- [4] K. Wang, *Intelligent Predictive Maintenance (IPdM) system – Industry 4.0 scenario WIT Transactions on Engineering Sciences*, Vol 113, © 2016 WIT Press www.witpress.com, ISSN 1743-3533 (on-line); doi:10.2495/IWAMA150301

- [5] Fengwei Yang & Sai Gu, *Industry 4.0, a revolution that requires technology and national strategies*, *Complex & Intelligent Systems* Vol. 7, pp.1311–1325 (2021)
- [6] Klaus-Dieter Thoben, Stefan Wiesner, Thorsten Wuest, *Industrie 4.0” and Smart Manufacturing – A Review of Research Issues and Application Examples*, January 2017 International Journal of Automation Technology 11(1):4-19,DOI: [10.20965/ijat.2017.p0004](https://doi.org/10.20965/ijat.2017.p0004)
- [7] Karabegović, I., Husak, E., Karabegović, E., Mahmić, M. (2023). Robotic Technology as the Basis of Implementation of Industry 4.0 in Production Processes in China. In: Karabegovic, I., Kovačević, A., Mandzuka, S. (eds) *New Technologies, Development and Application VI. NT 2023. Lecture Notes in Networks and Systems*, vol 687. Springer, Cham.
https://doi.org/10.1007/978-3-031-31066-9_1
- [8] KEY IP5 STATISTICAL INDICATORS 2023, All statistics herein are preliminary. The full IP5 Statistics Report 2023 will become available in the last quarter of 2024
<https://link.epo.org/ip5/KOzTLwU>
- [9] Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles C. J. ., Shoham, Y., Wald, R., and Clark, J.: “The AI Index 2024 Annual Report,” AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2024.
<https://aiindex.stanford.edu/report/>
- [10] Keisner A., Raffo J., Wunsch-Vincent S. (2016) Robotics: Breakthrough Technologies, Innovation, Intellectual Property. *Foresight and STI Governance*, vol. 10, no 2, pp. 7–27. DOI: 10.17323/1995-459X.2016.2.7.27
- [11] 2023 Anaqua, Analysis of uspto patenting statistics, 2024 Anaqua, Inc.,
https://www.anaqua.com/resource/anaqua_analysis_of_uspto_patenting_statistics_final-2023/ [Accessed; June. 20, 2024].
- [12] Karabegović, I., Karabegović, E., Mahmić, M., Husak, E. (2020). Dissemination of Patents of the Base Technologies of the Fourth Industrial Revolution - Industry 4.0. In: Karabegović, I. (eds) *New Technologies, Development and Application III. NT 2020. Lecture Notes in Networks and Systems*, vol 128. Springer, Cham.
https://doi.org/10.1007/978-3-030-46817-0_1

- [13] World Intellectual Property Indicators 2022, <https://www.wipo.int/edocs/pubdocs/en/wipo-pub-941-2022-en-world-intellectual-property-indicators-2022.pdf>
- [14] European PatentOffice, Munich, Germany, 2020, pp:7-13; <https://baa.no/en/services/ip-rights/>
- [15] UNCTAD (2018) Technology and innovation report 2021, United Nations conference on Trade and Development; based on data from Maddison Project Database, version 2018, Bolt et al. (2018), Perez (2002), and Schwab (2013).https://unctad.org/system/files/official-document/tir2020overview_en.pdf
- [15] Trademarks and patents in China: The impact of non-market factors on filing trends and IP systems.” January 2021. <https://www.uspto.gov/sites/default/files/documents/USPTOTrademarkPatentsInChina.Pdf>
- [16] Yang G-Z, Bellingham J, Dupont PE, Fischer P, Floridi L, Full R, et al. The grand challenges of ScienceRobotics [Internet]. Vol. 3, Sci.Robot.2018. Available from: <http://robotics.sciencemag.org/>
- [17] Perez, L., Rodriguez, I., Rodriguez, N., Usamentiaga, R., and Garcia, D. F.: Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review, Sensors (Basel), 16. (2016)
- [18] Nagata, F., Hase, T., Haga, Z., Omoto, M., and Watanabe, K.: CAD/CAM-based position/force controller for a mold polishing robot, Mechatronics, 17: 207-216. (2007)
- [19] Karabegović, I., Majtorović, V., Industry 4.0: Digital transformation shapes the future, Society for Robotics, Academy of Sciences and Arts of Bosnia and Herzegovina, 2024. pp. 21-48. www.anubih.ba
- [20] Cheng, Y.-S., Shah, S. H., Yen, S.-H., Ahmad, A. R., and Lin, C.-Y.: Enhancing Robotic-Based Propeller Blade Sharpening Efficiency with a Laser-Vision Sensor and a Force Compliance Mechanism, Sensors, 23: 5320. (2023)
- [21] Gotlih, J., Šket, K., Crnokić, B., Rakić, K., Šaravanja, L., Ficko M., „From the first industrial robots to intelligent systems in robotics“ in (Ed. Karabegović, I., Majtorović, V.: Industry 4.0: Digital transformation shapes the future, Society for Robotics, Academy of Sciences and Arts of Bosnia and Herzegovina, 2024, pp.59-84.
- [22] Zhang, G., Liu, C., Min, K., Liu, H., and Ni, F.: A GAN-BPNN-Based Surface Roughness Measurement Method for Robotic Grinding, Machines, 10: 1026. (2022)
- [23] Mohammadi Amin, F., Rezayati, M., van de Venn, H. W., and Karimpour, H.: A Mixed-Perception Approach for Safe Human–Robot Collaboration in Industrial Automation, Sensors, 20: 6347. (2020)

The Role of Blockchain Technology in the Transformation of Digital Marketing

Savo Stupar^{*1}, Elvir Šahić¹, Maida Cico², Mirha Bičo Ćar¹

Abstract: *With the advent of the Internet and the accelerated development of information and communication technologies (Cloud Computing, Big Data, etc.), marketing has changed forever. However, it will soon be going through another revolution, largely due to the potential of Blockchain technology. In order for the exchange (transaction) to be carried out to the satisfaction of both entities, there must be trust between them. Unless the entities trust each other, they usually find a solution by hiring a third entity, called an intermediary. At its core, Blockchain enables transactions between two parties without the need for third party verification. If a transaction means any business transaction of transferring ownership of goods or money (securities) from one entity to another, then Blockchain is an alternative way of exchanging and recording transactions by which it is carried out to the satisfaction of both entities, without intermediaries and without the need to trust each other. One of the more significant aspects of applying Blockchain technology is that it gives consumers back control of their personal information, eliminating the possibility for companies to take data from customers without offering them compensation. The aim of this paper is to highlight the potential of Blockchain technology, which puts customers in an even more favorable position, but also enables merchants to access non-intermediary customers such as Google, Facebook, etc. Another aspect of the application of Blockchain technology will be discussed in the paper, which concerns the possibility that customers have all relevant and true (i.e. non-changeable information due to the security provided by Blockchain technology) information about the origin of the products they purchase.*

Keywords: *Transaction, BitCoin, Blockchain Technology, Digital Marketing, Blockchain Marketing, Blockchain technology, Product origin*

1. Introduction

Good connoisseurs of trends in the development of new information technologies believe that the emergence of Blockchain technology represents a revolutionary technology that, due to its characteristics, contains a huge potential for changes in almost all areas of human activity, which can be manifested in improving the efficiency of performing a large number of human activities, in

^{*1}University of Sarajevo, School of Economics and Business, Sarajevo

²International High School "Bloom"

E-mail: savo.stupar@efsa.unsa.ba, elvir.sahic@efsa.unsa.ba, maida.cico@bloom.edu.ba, mirha.car@efsa.unsa.ba

increasing or even complete transparency of these activities, eliminating the need for intermediaries (banks, state, municipality, court, prosecutor's office, lawyers, notaries, etc.) when performing a significant number of human activities, reducing the possibility of fraud and increasing the level of security in performing various types of transactions, whether they relate to the creation of cryptocurrencies and their exchange (the first application of Blockchain technology) or to the exchange of securities, documents, real estate, gems or various other goods, in trusting the accuracy and correctness of various types of records due to the impossibility of changing once stored and from a large number of equal partners, verified transactions.

Any exchange of goods (goods, services or things and money, as their measure of value) is based on transactions. Generally speaking, a transaction can be defined as the transfer of ownership of goods (things, money, real estate, jewels, securities, etc.) from one subject to another, whereby one subject (the one who initiates the transaction) loses the right of ownership of the goods, and the one who buys the goods acquires the right of ownership over them. Different names for the transaction are buying or selling, payment or spending, providing or using (services, for example), owing or claiming (on an account, for example), increasing or decreasing, receiving or giving, etc. Almost all services today are based on that concept. In order for the exchange of transactions to be carried out to the satisfaction of both entities, there must be trust between them, or if the entities do not trust each other, they usually find a solution by engaging a third entity, i.e. an intermediary, to whom, according to mutual agreement, (one or both) pay for the mediation service. Banks are the best example of this. But it is not only banks and organizations that make money from mediation services. Very often, the state, courts, lawyers, notaries, prosecutors' offices, real estate agencies, educational institutions, marketing agencies and various other institutions appear as intermediaries that manage the records of various transactions. The ultimate goal of carrying out transactions to the satisfaction of both participants in the exchange is not only the elimination of intermediaries and the avoidance of commission payments, but also the increase in security and speed of transactions, as well as the improvement and liberalization of transaction business, and the elimination of the possibility of abuse and fraud. The technology that has a huge potential to support the achievement of these goals, as well as to reduce business costs and thereby increase profits, is Blockchain technology, which is therefore considered a revolutionary technology, that is, the technology of the future.

Blockchain technology was created as an amalgam of several different, but equally revolutionary technological achievements in the field of mathematics (power of large numbers), cryptography (hash functions, which will be discussed more later), P2P computer networks (network of distributed computers or Peer to Peer network) and improvement of computer hardware characteristics, such as

processor capacity and speed, capacity and speed of access to external memory media Peer-to-Peer network (or partner network) is a distributed network made up of connected independent computers. The main characteristic of these networks is that they are constructed in a way that allows establishing a connection and realizing the exchange and joint processing of information with all other computers in the system without giving privileges to any computer in the network, that is, without a central authority. The first application of this technology refers to the support for the creation and exchange of the first cryptocurrency - Bitcoin. [1]

2. Bitcoin – First Application of Blockchain Technology

If Bitcoin did not appear, there would be no Blockchain technology. Likewise, if there was no Blockchain technology, there would be no Bitcoin. Before a brief introduction to Bitcoin, it is necessary to define some elements of the categorical (conceptual) apparatus, which are used for its explanation. Namely, the terms virtual and digital money are often used as synonyms, which is completely wrong, so it is necessary to define the essential differences between them, or more precisely the differences between virtual money and the money we use every day. Digital money is money that is transferred electronically. More precisely, it is money, the amount of which (some integer or decimal number) is stored in the computer memory in the form of a binary representation. It is the money of a user on a current account stored in the bank's computer memory, to which credit, debit cards and other cards can be attached. It is also money in the PayPal account. In this context, digital money is also virtual money, but digital money does not have to be virtual money. There are different types of digital money and different types of virtual money. Compared to traditional money, virtual money does not have all the attributes of real money. The term virtual can be defined as something that exists but is not in physical reality. Before the advent of cryptocurrencies like Bitcoin, virtual money mostly referred to money in popular MMO (Massively Multiplayer Online) games like World of Warcraft and EVE online. Bearing in mind that these computer games serve several hundreds of thousands of players who trade with each other in "raw materials and final products" in that virtual world, it very often happens that designers (game developers) ask for expert advice from economists in order to design the economic aspect of games. Virtual money in computer games is centralized, so designers (developers) of computer games can control supply and demand in the form of monetary rewards and penalties. It is important to emphasize that computer games have developed markets, which are exclusively maintained by players, and that they respond to the laws of supply and demand. The virtual markets of computer games have spawned new markets in the real world, where players can buy in-game money or goods in exchange for real money. That's

why this practice is banned in most MMO computer games, so administrators look for and punish players who engage in it by blocking their access. Virtual money suddenly gained popularity with its appearance in the form of Bitcoin in 2008, and then it was noticed by central banks in the transitional period between 2012 and 2013. The concept of virtual money is still insufficiently researched, especially from the aspect of the rapidly growing number and total value of transactions, the number of users and the extent of the confusion that cryptocurrencies caused in the regulatory systems of banks and states.

The current concept of "virtual social community" was created as a product of the rapid development of information and communication technology, globalization and the economy of virtual crypto money. Bitcoin represents the first case of the appearance of a cryptocurrency, so a more detailed look at it in the form of a "case study" from the aspect of Blockchain technology, which enabled its creation and functioning, seems necessary and useful for a better understanding of all other applications of this technology, and especially its application in the field of digital marketing.

Due to the potential of modern computers to very quickly process a huge number of transactions via a computer network, as well as due to their characteristic of possessing a "perfect and powerful memory", the idea of a decentralized currency system did not represent any novelty considering that much earlier, there were a large number of proponents of the concept of anonymous digital money.

The concept of anonymous digital money, which is considered the first wave of digital money, begins in 1981 with the work of cryptographer David Chaum on the topic: "e-mails without a trace, sender addresses and digital pseudonyms"[1]. He founded the International Association for Cryptologic Research (IACR) in 1982, and published his best-known paper entitled "Blind signatures for Untracable Payments"[2] in which he created a technical model for future digital currencies. He believed that the world needed a digital pseudo-money that would replace physical (paper) bills and metal coins (coins) and that would enable private and secure transfer of money "from hand to hand" by its design. His views on money and privacy were the inspiration for the creation of the ideology of the crypto-activist group "Cypherpunks", which was founded in 1992 and which left an indelible mark in the world of cryptography.

Back in 1991, Phil Zimmerman coined the term "Pretty Good Privacy" or abbreviated PGP, and made a free public key encryption program. For the first time, the PGP application gave every interested individual access to the enormous possibilities of cryptography, as Zimmerman printed the source code of his program in a book, which was then publicly distributed[3].

One of the main founders of the „Cypherpunks“ group, Eric Hughes, wrote the Cypherpunks Manifesto and mission statement in 1993, which read: *"Cypherpunks are dedicated to building anonymous systems...Privacy is*

*essential to a free society in the modern age...We cannot expect governments, corporations and other faceless organizations to provide it to us... We must defend our privacy if we intend to have it. We defend our privacy with cryptography, with systems for forwarding anonymous emails, **digital signatures and electronic money**."*[4] The first wave of digital money in the form of an **E-cash** system as a finished product of the **DigiCash** company, was based on blind digital signatures to provide users with anonymity. Representatives of **the second wave of digital money** are **PayPal** and **E-Gold**. PayPal acts as an intermediary between the merchant and the customer, and does not reveal the customer's personal information during the transaction, because the money is actually transferred from PayPal's account. Of course, Paypal takes a fee for this mediation. E-Gold represents the first hybrid of a digital currency and a PayPal-like protocol. At the same time, E-gold is the first popular digital currency that uses gold and precious metals as collateral. Users would send money via check or card or gold or silver, and the **E-Gold** company would then allocate the currency to the user's account, as a unit of measure for the sent counter value. The **E-Gold** company would hold the physical gold, while users with e-gold currency could perform transactions worldwide.

The third wave of digital money in the form of the Bitcoin protocol is linked to Satoshi Nakamoto and the appearance of the Bitcoin cryptocurrency. When Satoshi Nakamoto, whose identity still remains a mystery, published the document "Bitcoin: A Peer to Peer Electronic Cash System" [5] in 2008, he designed and proposed a "true version of electronic money through a Peer-to-Peer network" called Bitcoin, for the first time Blockchain technology appeared in public. In January 2009, Bitcoin is offered "open source"² to the public. Blockchain concept answered the question of digital trust by recording important information in the public space without allowing the information to be changed or deleted. The basic characteristics of Blockchain are: transparency, time-stamping and decentralization.

Nakamoto defined the technology that makes Bitcoin work as: "An electronic payment system that is based on cryptographic proof, instead of belief, allowing two willing parties to make direct transactions without the need for a third, independent party." [5]

To describe the meaning of the term Bitcoin, various terms are used, such as virtual and digital money, then virtual, digital, electronic, synthetic and cryptocurrency. As with most relatively recent terms, there is no single definition. The American FinCEN and the European Central Bank have designated Bitcoin as a virtual currency. The People's Bank of China has classified Bitcoin as something that is "not originally a currency, but an investment". A German court characterized Bitcoin as a unit of measure. The

²Open source software with a license to modify, modify and improve.

Finnish government, as well as Wall-Street, daily newspapers, have classified Bitcoin as a commodity. Bitcoin.org, the wiki portal for Bitcoin, gave the following definition: "Bitcoin is a payment method based on the concept of digital cryptocurrency, which functions without any central authority or third party as a creditor." [6]

As can be seen, the definitions differ significantly with regard to the aspect of observation. However, what is certainly an indisputable fact is that Bitcoin is a new technology or, more precisely, a protocol, and that any other function that Bitcoin performs derives from its technical characteristics as a protocol. Generally speaking, a protocol is an agreed procedure or set of agreed instructions to be followed in a given situation.

This cryptocurrency is not based on a gold base, is not tied to any country of origin and is not backed by any country, nor the Central Bank of a country, nor the Central Bank of a union of states. Thus, Bitcoin is the first cryptocurrency, i.e. the first form of virtual, digital currency that functions on the basis of cryptographic algorithms. Its basic characteristic is that there is no central authority, nor any central institution like the Central Bank, nor any other intermediary that manages it. There is the Bitcoin protocol, which uses Blockchain technology as support to achieve a fully decentralized system.

How does Bitcoin work? That is a question that is not very easy to answer. That answer is not so important for the functioning of Bitcoin itself as it is important for the explanation of the Blockchain technology, which makes it possible, which will be discussed in the next chapter. More important than anything for the functioning of Bitcoin is that a new user can start using Bitcoin without understanding the technical details of its functioning, in the same way that anyone can use the program without being its author or programmer. It is only necessary to install the Bitcoin Wallet application on a computer or mobile phone, which will then generate the user's Bitcoin address, that is, its account. Transferring Bitcoin from one address to another, or in banking terms "from one account to another", is in practice similar to sending and receiving e-mails. Just as in the case of e-mail, the sending user needs to know the address of the user to whom he is sending (the address of the recipient) a document or message, that is, a certain amount of Bitcoin, so he (the recipient) needs to know the address of the user who is sending it to him (the sender), if he wants to send something.

New cryptocurrency units are produced [7] ("printed") or, in banking terms, "emitted" by mining, as a reward for solving a mathematical task, set by the Bitcoin protocol, which verifies the newly created block and writes it into the Bitcoin Blockchain file, which will be discussed in more detail later in the text. Users can acquire cryptocurrency by purchasing a certain amount of cryptocurrency, and by selling goods or services. On average, a new block is generated on the Bitcoin Blockchain file every 10 minutes, so this is also the average time for transaction confirmation. After generating a new block, the

miner receives a reward (incentive) which currently amounts to 3,125 BTC (from 6,25 BTC per mined block to 3,125 BTC and which is halved every 4 years (BTC is the abbreviation for the Bitcoin cryptocurrency unit) as well as fees for transactions written in the new block. Mining is the only way in which new BTCs are "issued". The Bitcoin protocol defines the maximum number of Bitcoin cryptocurrency units, so it is one of the the reason why we can say that Bitcoin is in fact a deflationary currency. This is another difference compared to classic currencies, which central banks can issue as needed and in this way cause inflation. Until today it is "mined" or "issued" 19.746.787 BTC or about 94% of the projected sum of BTCs (21.000.000, and the current value (29.8.2024) of 1 BTC is 107.103,89 BAM.[8]

The more BTC there are, the smaller and smaller the reward, thus further preventing inflation. When miners have created 21 million bitcoins, this reward will be zero, and miners will only earn from transactions.

When we talk about Bitcoin transactions, we mean the transfer of money, or the value it represents, from one digital wallet to another, which is then registered as one digital record in the BitcoinBlockchain file (or the public ledger of all transactions), which together with other records (transactions) from all over the world, previously written in one block of transactions, is a block of still unconfirmed transactions. When all transactions generated within an average of 10 minutes from the moment of writing the last verified block into the BitcoinBlockchain file are collected, all miners are given the task to find the corresponding number, i.e. solve the corresponding mathematical problem. The first to do so ("wins the lottery") writes a new block of unconfirmed transactions into the BitcoinBlockchain file according to the system of chained blocks, so that from that moment all transactions in that block become valid and public. The size of that file today, 29.8.2024, has risen to about 596,51 GB[9].

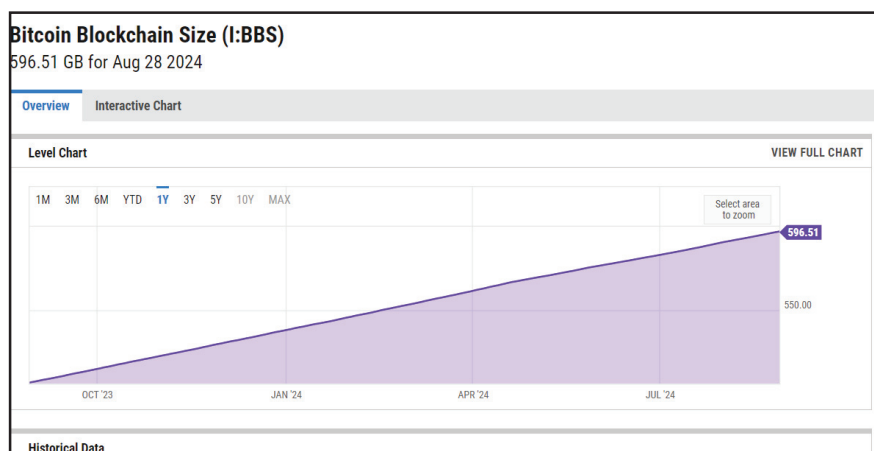


Figure 1. Bitcoin Blockchain file size[9]

A Bitcoin wallet contains a piece of data that is hidden from the public and is called a private key or password (the access code to the safe where the money is stored or the access code to the wallet). It is used to sign transactions confirming the mathematical proof that connects the key and the owner of the wallet. Once a transaction has been initiated and signed, it cannot be changed any more, which is exactly why the signature plays an important role. Transactions on the Blockchain network are public, and you can see the entire history (archive) of all processed transactions from the first to the last.

In practice, if entity A wants to send a certain amount of Bitcoin to entity B, it must have a digital wallet that contains a private key, which allows the creation of a cryptographic (digital) signature. Entity A enters the amount of Bitcoin that it wants to send to entity B. Entity B gives the public key (digital wallet address) to entity A to transfer the specified amount.

To be able to transact a certain amount of Bitcoin from one digital wallet to another, three things are required:

1. Wallet address or Public Key,
2. Privatekey or code access to the wallet and
3. Cryptographic (digital) signature.

When we describe the Bitcoin protocol, we often attribute two qualities to it: 1. that it is completely transparent (with a public balance book) and 2. that it is anonymous. If we focus on one item of the public balance book and the minimum necessary attributes for it to be recorded, we will notice that data such as name and surname, and other personal data are not necessary for the functioning of the balance book in terms of functionality, of course not taking into account KYC (Know Your Client) protocols, i.e. "get to know your client", which the bank must implement. The account number uniquely identifies the user, as is the case with the Unique Identification Number of the citizen³ on the identification documents of the citizens of Bosnia and Herzegovina. That's why, when creating an "account" with the Bitcoin protocol, unlike creating a bank account, the Bitcoin protocol never requires the user to enter personal data. So, since this is a Bitcoin protocol, and not a bank, it would be more correct to use the term address, rather than account, as a label for the location where the number, which represents the amount of Bitcoin, is stored.

Every transaction in the Bitcoin protocol is public, and everyone has access to the public balance book, which they can see on the Internet at any time, as well as the history (archive or traffic) of all transactions made at any time from every active account.

Since the addresses of the participants in the transaction are not linked to their personal data, it is very difficult (practically impossible, but theoretically

³This number uniquely identifies every citizen of Bosnia and Herzegovina.

possible) to figure out which address belongs to whom. This further means that Bitcoin is not anonymous, but a pseudo-anonymous system. However, we should not forget that everyone in the Bitcoin protocol can have an almost unlimited number of addresses.

The address is defined as a bank account. A Bitcoin address is an identifier 26-34 alphanumeric characters long, starting with the number 1 or 3, that represents a possible destination to pay Bitcoins using a cryptocurrency exchange account or through a digital wallet. It is important to point out that one user, as in the case of a bank account, can have and use several addresses. The example of Bitcoin address: 1KFHE7w8BhaENAswwryaocDb6qcT6DbYY. These addresses are created using cryptographic algorithms, more specifically SHA-256⁴ for generating private keys and RIPE-MD160⁵ for generating addresses based on SHA-256 operation.

The private key is a unique and secret access code, which provides the right to transfer Bitcoins from the wallet using a cryptographic (digital) signature. When a new Bitcoin address is created, it comes with a private key that is mathematically linked to that account number. Bitcoin private keys usually contain 51 characters and start with the number 5. These private keys are memorized (stored) on a personal computer. In the case of using a software or web wallet, they are stored on the server.

A cryptographic (digital) signature is a mathematical method (algorithm), which is used to verify the origin and determine the integrity of information, and thus enables the owner to prove their ownership of a certain address, or wallet. Digital signatures have two essential characteristics:

- They ensure the integrity of the transaction (document), that is, they confirm that the transaction (document) has not been changed in the intermediate phase between reading (reviewing) and signing, and in the case of a contract that the articles and conditions have not been changed without the consent of both parties.
- The impossibility of forging a signature, which means that the signature uniquely identifies the signatory, and implies that it is not possible to avoid responsibility for the signed transaction (document), claiming otherwise.

At the moment when the Bitcoin protocol (software) signs the transaction with the corresponding private key, everyone in the Bitcoin network is enabled to display the (digital) signature that corresponds to the executed transaction and which is generated by a cryptographic algorithm, but the private key, which protects the account, is impossible to reach.

⁴The name of the algorithm for generating a digest (hash) from an input of arbitrary length, which will represent the private key.

⁵The name of the algorithm for generating a digest (hash), from an input of arbitrary length, which will represent a Bitcoin address.

A **cryptographic hash function** takes input data, i.e. a message or any string of characters of arbitrary length and converts it into a string of fixed length, i.e. an "encrypted" message better known as hash-code, hash-result, hash-value or simply hash⁶. This procedure is called hashing and is usually done using the Sha256 algorithm⁷. It is almost impossible to discover the original data from such an encrypted message. With the slightest changes to the input data to the hash function, the hash value changes completely.

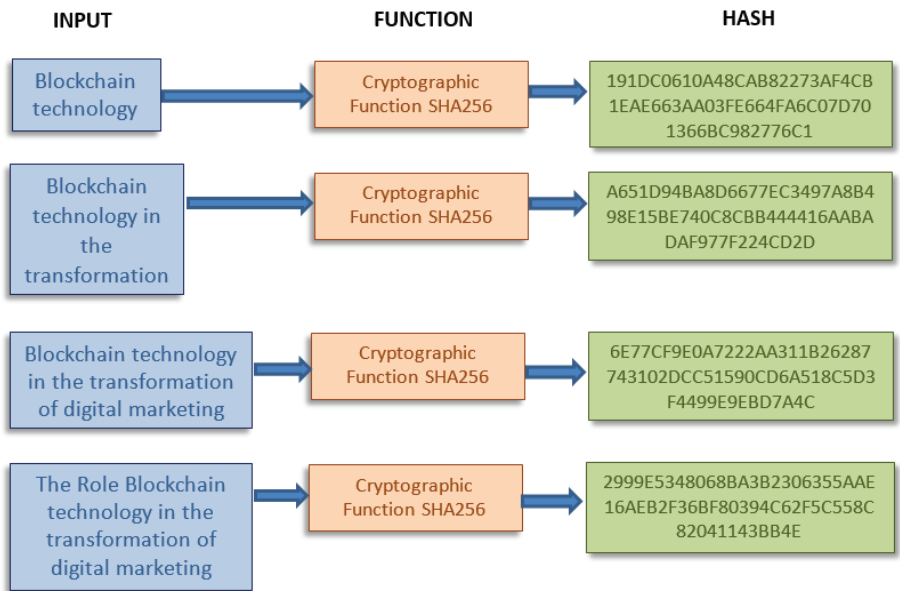


Figure 2. Calculating Hash Function Values for Different Inputs

2. What Is Blockchain Technology?

Most people still equate Bitcoin with Blockchain technology, so even though they are related, they are still two different concepts. The first practical realization of Blockchain technology was Bitcoin. We can say it differently. In order to realize the idea of anonymous digital money in the form of a cryptocurrency, it was necessary to find a new technology that would enable the decentralization of the management of that currency and the best possible

⁶A hash function is any function that can be used to map (convert) input digital data of arbitrary size to an array of digital data of fixed size (hash value). A good hash function is one in which small differences in the input data result in very large differences in the output data. Hashing is the process of obtaining a hash value.

⁷Hashing is the process of transforming an input of arbitrary length and obtaining a hash value of a fixed length, which in the case of applying the SHA-2 algorithm (Security Hashing Algorithm version 2) is 256 bits or 32 bytes or 64 hexadecimal characters.

protection of transactions. This is how Blockchain technology was born, which was quickly realized that it can be used for many other things, not only for the functioning of cryptocurrencies.

One of the basic ideas on which the concept of this technology is based is taken from accounting [7], and refers to the recording and archiving of all business transactions that have occurred since the establishment of a company. According to the principles of double-entry bookkeeping, every business event in the company is recorded so that the corresponding account is debited (increased), and another account or more are approved (reduced) in total for the same amount of the transaction or vice versa. Up-to-date records of all changes in accounting accounts (transactions) and the balance of these accounts are kept in a record called the general ledger or balance book.

What in the accounting records of the company is the main ledger with all transactions (changes) in all accounts of the general ledger since the beginning of the year, in the Blockchain records is the main Blockchain file, with all transactions since the beginning of record keeping. Records of transactions, i.e. changes in the accounts of the company's general ledger are kept in one place (the company's computer) and there is no verification of transactions by third parties, and the Blockchain record of addresses (user accounts) together with the transactions related to them, is kept in a large number places, that is, on powerful computers - servers around the world, which participate in the verification of each transaction and receive an appropriate reward for it. In the first case, when it comes to records maintained by the company, it is the main balance book, and in the second case, when it comes to Blockchain records, it is a public (global) balance book.

In the first case, since it is about central records or records in one place (computer), it is possible to change (maliciously or not) the content of transactions in a relatively easy way, whether it is done by company employees or someone outside the company who has access to this data via the internet. In the second case, it can be claimed with a very high degree of probability that this record is practically immutable, because it is kept in the same form on a large number of computers around the world, so in case of manipulation (hacking) attempts, it should be changed on all these computers, which makes it practically impossible. This is the second basic idea on which Blockchain technology rests, which could not be realized before the advent of the Internet and Peer to Peer computer networks (decentralized networks). The realization of this idea was also contributed to by the rapid development of hardware through an increase in the speed of the processor, an increase in the capacity and speed of access to external memory media, as well as the lower prices of the processor, as well as the main and external memory elements.

The third basic, and possibly the most important idea on which Blockchain technology is based, is the creation of additional security for

uninterrupted transactions, which is reflected in the impossibility of changing the content of once recorded transactions in the Blockchain file. The realization of this idea is done by connecting the blocks of transactions into a chain of transactions (Blockchain) in such a way that in each block of transactions, which is currently entered into the Blockchain, the content code (summary or hash) of the previously entered block is also entered, so if someone tries to change at least one character in any transaction block, the content code (hash) of that transaction block is also changed so that in that case it does not match the content code (hash) of the previously entered transaction block, and it should also be changed and so on until of the last written block. In other words, if someone wants to change the content of a transaction without authorization, he should change the hashes of all blocks from the block he wants to change to the last written block in the Blockchain file, and then distribute the changed file to all servers (nodes) around the world, which seems practically impossible.

Blockchain[10] represents a shared data structure, i.e. a list of information shared (distributed) between all nodes (nodes or servers) that are in the system (network). Given that the database is not stored in one place or on one server, it is decentralized. As the council has said more than once, decentralization, i.e. eliminating the need for the existence of a central authority, is one of the most important ideas of Blockchain technology. Instead of everyone who is engaged in some business and who generates some transactions in connection with it, having their own separate transaction record books, Blockchain technology offers them a unique and global (comprehensive) balance book, which contains all transactions since the beginning of record keeping, is public and is owned by all. It was created by the implementation of a Peer-to-Peer (P2P) network and a network of distributed servers (nodes), which mark transactions with a time stamp.

A Blockchain file consists of blocks, and each block can record a certain number of transactions, and a certain amount of data in a transaction. When a block is filled with transactions, a new one is created, thus creating an unbroken chain of interconnected blocks. When someone initiates a transaction from one address (account) to another address (account), it is time-stamped and recorded by every participant in the system), i.e. owned, because the balance of the account is public, synchronized with other system participants and transaction records, and the rules are defined at the beginning and implemented through program code. In this way, double spending problem⁸ was solved without third parties, i.e. an intermediary trusted by both parties.

Blockchain is based on a distributed database, which eliminates the need for a separate entity (bank, state, municipality, court, lawyer, notary,

⁸The possibility of spending the same (virtual) money effectively at least twice, because it is recorded in digital form.

intermediary marketing agency, etc.) that will manage transactions as an intermediary and ensure the security of their performance. Thus, one central register of transactions with all transactions since the beginning of record keeping in the centralized banking system is replaced by copies of Blockchain files on a large number of powerful computers (servers), which will be located in different locations, which are called nodes. The owner of these powerful servers can be anyone, that is, anyone who applies for such a job and provides the availability and possibility of updating the records (Blockchain file) to all participants in the transaction at any time of the day. Due to the fact that this job requires significant investments in equipment, as well as the payment of electricity, computer administration and maintenance costs, it is logical that none of the participants in the transaction will apply for such a job just like that. In order for someone to engage in these jobs, there must be an appropriate reward for entering and doing such work. The basic task of "miners" is to confirm the transactions of the participants in the transaction and to ensure their validity.

Systems that use Blockchain technology belong to peer systems (Peer-to-Peer) and are called decentralized or distributed systems. In this way, the exchange of data through the computer network is enabled, whereby the nodes receive information from each other instead of from one central computer. With respect to the tasks that equal partners can perform, there are four types of partners. These are: a simple partner or a partner called a "simple wallet", a "router" partner, which performs network routing, a "miner" partner, which performs mining, i.e. finding the code that closes the block transaction and maintenance of the entire Blockchain. In a private Blockchain system, each partner manages all four tasks.

3. Application of Blockchain Technology in Digital Marketing (Blockchain Marketing)

An example of a P2P network is represented by torrents such as the BitTorrent⁹ client, on which model Bitcoin was constructed. The first software application that used a P2P network was called Napster. It was launched in 1999, and was used for sharing primarily audio and video, but also other content. Napster made it possible to search for content, and often illegally "download" of mentioned content. Napster is an application, which was still partially centralized even though it relied on a distributed P2P network. Let's assume that a Napster client is installed on each of the several thousand computers that make up the network.

⁹BitTorrent is just one of the tools in a series that helps to share data more easily and quickly, unlike the conventional ways of sharing data via a centralized server, via email or File Transaction Protocol (FTP). While, for example, email and FTP rely on a central hosting server and sending data to connected computers, BitTorrent relies on a network of users where everyone sends data to everyone (peer-to-peer network).

When someone (peer1) using one of those computers wants to "download" some video content in the form of an mp4 file, it is necessary to connect to the Napster server. This server performs the role of an intermediary and its sole purpose is to index the contents of other computers in the network and inform the owner of the computer (peer1) about other computers in the network that are online and have the requested content on their hard drives. The job of this server is not to store the requested content (file), but to provide a service of referencing that content in the network and help in establishing a connection between these points (computers) in the network.

In this sense, the server can be described as a central indexing server. This is the reason why P2P networks are often in the "grey zone" of the law, precisely because they have no direct contact with potential illegal files, nor a way (technique) to control the freedom of users to "download" the content they want. This partial centralization in the form of a central server for indexing was a weak link in the Napster system, so it is abandoned, and the BitTorrent protocol, designed by Bram Cohen, enters the scene.

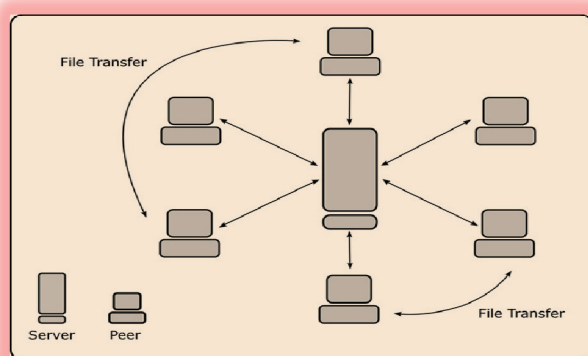


Figure 3: Napster server architecture configuration [11]

When it comes to Blockchain digital marketing, the most important are three changes, that is, differences in relation to not only traditional, but also digital marketing, which could revolutionize marketing as we know it so far. Namely, it is about the potential inherent in Blockchain technology and which enable the removal of the digital intermediary, increasing transparency and giving consumers back control over the sharing of their own data. [11].

As we have already emphasized in the paper, the Blockchain did experience its first application by enabling the functioning of the first cryptocurrency Bitcoin, so this is the main reason for the initial connection of this technology to Bitcoin, but it is clear to everyone today that the Blockchain is

much more than Bitcoin itself and that it has a huge the potential of application to all areas of human activity, not only cryptocurrencies.

One of the main areas where the application of Blockchain can make radical changes is digital marketing. One of the most important aspects of Blockchain technology is enabling decentralized, i.e. direct communication between different parties in such a way that every detail of that communication is verified and documented.[11]

The first change brought by Blockchain Marketing compared to traditional, i.e. digital, marketing is the ability to completely eliminate the intermediary in digital marketing. Digital marketing uses „intermediaries“ which control the space between the advertiser and the user. Take, for example, Search Engine Marketing (SEM)¹⁰. One of the reasons a website chooses to display Google banner ads on its site, as opposed to organic finder advertisers, is that Google is a trusted source. If the company was „reviewed“ Google Display Network, then it's likely to be a credible business that won't damage the host site's brand. Google also performs transaction processing so that the website owner is paid fairly for the clicks generated on the advertiser's ad. In this case, Google is basically an intermediary between the advertiser and the website owner. For this reason, they reduce profits.

Now imagine SEM through the Blockchain. Website owners would not have to go through the Google Display Network to find advertisers, as each „user“ had already been confirmed and authenticated. The advertiser would know they are paying for real clicks and the website owner can trust that the amount they are being paid is fair. There would be no need for Google (or Facebook or any other intermediary). It all comes down to reducing additional costs and increasing the profit margin.

The second change [11]brought by Blockchain Marketing compared to traditional, i.e. digital marketing is the increase in transparency between advertisers and users, which builds and increases trust between them. One of the frequent problems faced by many companies, mostly large ones, is the problem of consumer trust. Customers are suspicious of everything they buy, from the origin of the food they buy to the factory conditions in which the clothes they buy are made. Companies gain customer trust with the help of transparency, which Blockchain marketing can provide in a way that some other tactics and techniques cannot. An additional advantage and value of Blockchain marketing is the level to which everything is documented and verified. For example, with the help of Blockchain marketing, a customer could find out the “behind the scenes” view of the seller's supply chain, i.e. get accurate information about how

¹⁰Search Engine Marketing (SEM) is the process of generating traffic on a website by purchasing advertisements on internet search engines. For many users, the journey from obtaining information about a product to purchasing a product begins on Internet browsers. They get the answer to every question by searching through an internet browser.

the product is made. Transparency allows the consumer to find out whether the company is telling the truth when it publicly presents its products or not, whether it is trustworthy or not. In one case study, Walmart hired IBM to create a project that aims to make their supply chain process as transparent as possible. By digitally searching the chain blocks that store information about the origin of their products, consumers have been given the proof they need to trust what Walmart is selling them.

The third change brought by Blockchain Marketing [11] in relation to traditional, i.e. digital marketing is enabling the consumer (user/buyer) to control the sharing of their own data. Blockchain Marketing has the potential to find a way in the future to balance advertising with consumer identity issues, i.e. how to not violate consumer privacy through advertising. "There are already several services that offer the user full control over their identity and transaction history, such as uPort, MetaMask and Keybase. This means that marketers will have to earn the customer's permission in a way that is completely different than before: Blockchain allows consumers to pay attention to their contact information. In the future, the advertiser may end up paying the user to consume their advertising material." [11]

In addition to these three changes for which Blockchain Marketing has the potential, some other advantages that it has compared to traditional marketing should be mentioned. They are [12]:

1. **Blockchain has the potential to drive public responsibility**, i.e. socially responsible business. Transparency and documentation stored in Blockchain files can be used to create digitized contracts that can be viewed by the public and used to pressure companies to be socially responsible.
2. **Blockchain has the potential to offer advantages for easier and better branding**. If the company uses innovative technologies, such as Bitcoin-friendly payment methods and/or the use of Blockchain to formalize digital negotiations between companies, it will be seen as an advanced and robust organization. In other words, Blockchain Marketing increases transparency and trust in the brand. "For example, if a business used Blockchain to track the delivery of a purchase and allow both parties (both buyer and seller) to see where the package is on its journey, this increases consumer confidence in the business to deliver the item. This represents a leading USP¹¹ for marketing companies for promotion and a means of attracting more market share" [12].

¹¹It is a specific advantage that the seller has compared to the competition, which tells the buyer how the seller differs from the competition and why he should choose him.

3. Blockchain technology opens international markets. The first way Blockchain can benefit digital marketers takes us back to its most famous use - cryptocurrency. If marketing efforts can include and promote the use of cryptocurrencies to purchase products and services using online Bitcoin wallets, it opens up an international market. This happens because cryptocurrency payments can be made across borders without additional fees. This will motivate more people to buy from overseas companies and this is something that retailers can use to their advantage. [12]
4. Blockchain helps in handling Big Data[12]The potential relationship between Big Data and Blockchain has not been fully explored. It is predicted that Big Data could analyze the Bitcoin Blockchain to detect trends and help understand price fluctuations, as well as predict the future value of Bitcoin. Another benefit though, is that large amounts of data can be stored on Blockchain files to keep the records secure and unaltered, preventing them from being lost and helping to improve the accuracy of big data analysis between large teams. There is no reason why this cannot be the case with marketing departments as well. A large marketing team may want to analyze consumer trends from a dataset, and by keeping that data within an internal Blockchain, team members can work more efficiently without multiple versions of the same dataset. This will overall increase the accuracy of the final result and prevent data loss.

The greatest impact of Blockchain technology is reflected in the area of leasing advertising space on digital channels. To a lesser extent, Blockchain technology affects the value of data that customers/clients/users continuously leave on the Internet, and in connection with that, the way of communicating through digital channels, i.e. the creativity of the message. Not so long ago, and even today, one of the biggest challenges in the field of digital advertising was security, i.e. lack of trust, and the impossibility of checking the accuracy of the information received. This is most pronounced with programmable advertising, i.e. renting digital advertising space with the help of innovative technologies and advanced algorithms.[13]

Programmable platforms are mediators between advertisers and media that, with the help of advanced technology, provide the advertiser with the highest return on investment. The programmable platform provides the possibility for advertisers to find targeted groups of users in the right place at the right time on online channels and to communicate the right message to them. The key problem that arises here is non-transparency, which results in the

expansion of the space for fraud. "One of the most famous and biggest cases of fraud in digital advertising is the so-called "Methbot Fraud". Namely, Russian hackers created a network of fake portals with a name similar to CNN or FOX. They made them so well that they managed to deceive the ad networks that, due to wrong recognition, placed ads on them from some of the world's biggest advertisers. At the same time, they developed a bot that could simulate human behavior and click on those ads. In this way, they earned over 180 million dollars. Thus, the advertisers paid 180 million dollars for clicks on ads that no one ever saw, and the portals were left without 180 million dollars in income because these ads were not shown at all".[13]

The creators of Blockchain solutions made sure that such things would not happen in the future. A new layer has been added to the intermediary, that is, the programmable platform, which enables a more significant level of transparency, and eliminates the space for malfeasance and fraud. Every click, every payment and display is recorded in a Blockchain file, so that record cannot be deleted or changed. This record is available and clearly visible to everyone at any time. One of the better examples of Blockchain platforms that try to take the programmable mechanism to the next level with the help of Blockchain technology is AdBank network. [13]

4. Conclusion

Due to the potential and benefits provided by the application of Blockchain technology, which are reflected in the provision of transparency, the availability of accurate information and the verifiability of the accuracy of this information, the practical impossibility of erasing and changing information once recorded, the elimination of almost all fraud and malfeasance, the elimination of all intermediaries in any processes, it is expedient and advisable to use Blockchain technology in marketing. One of the most significant benefits of using Blockchain technology is to give consumers back control over their personal data, eliminating the possibility of companies taking unauthorized data from customers (in a way that seems to be taken for granted) without offering them any compensation for it. As emphasized in the abstract of the work, the goal of this work was to point out the incredible, that is, the revolutionary potential of Blockchain technology in almost all areas, including in the field of Marketing. What could be concluded when it comes to the application of Blockchain technology in Marketing or Blockchain Marketing is that it puts customers in a significantly more favorable position, primarily because they have the option of checking the accuracy of the information that advertisers place on them, both in relation to the origin of the product, as well as with the entire product supply chain, which prevents abuses and eliminates fraud. The following benefits provided to customers by Blockchain Marketing are the

possibility to pay for products with cryptocurrencies, then the possibility that before purchasing the product, customers have all the relevant and true information about the origin of the product they are buying, and the change in the paradigm of ownership, i.e. control of their personal data, which is returned to them, and also they can earn extra money from it. As for the benefits for merchants, Blockchain Marketing provides them with direct access to customers without intermediaries such as Google, Facebook, Amazon, and in this way ensures a better quality of communication with customers, as well as a reduction or complete elimination of advertising costs through intermediaries, etc.

As a consequence of the application of Blockchain technology in marketing, the concept of data ownership is changing. Facebook changed the well-known slogan "data is the new gold" to "we enable you to use the software platform for your needs, and in return we get your data that we will use for our needs". [13]Blockchain technology offers the potential to change the paradigm of data ownership in this area as well. So there are already media and platforms that are ready to pay for the time that customers spent reading ads, as well as the data that customers gave them.

4. References

- [1] Chaum, D., *Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms*. California University, Department of Computer Science, 1981.
- [2] Chaum, D. (1982). *Blind Signatures for Untraceable Payments*. California University, Department of Computer Science, 1982.
- [3] Zimmermann, Ph., *PGP Source Code and Internals*, ISBN 0-262-24039-4. MIT Press, 1995.
- [4] Hughes, E., *A Cypherpunk's Manifesto*, Nakamoto Institute, 1993. <https://cdn.nakamotoinstitute.org/docs/cypherpunk-manifesto.txt> [Accessed: 13. 01. 2019]
- [5] Nakamoto, S., *Bitcoin: A Peer-to-Peer Electronic Cash System*, Bitcoin.org, 2008. <https://bitcoin.org/bitcoin.pdf> [Accessed: 13. 01. 2019]
- [6] Stupar, S., Bičo Ćar, M. Šahić, E. "Challenges of Applying Blockchain Technology". In: *New Technologies, Development and Application II*. NT 2019. Lecture Notes in Networks and Systems, vol 76. edited by Isak Karabegović. Springer, Cham. 2019.
- [7] Stupar, S., "Blockchain - Tehnologija bliske budućnosti?" u Zbornik radova Jahorina Business Forum - VIII naučna konferencija, 339-348, Jahorina, mart 2019.
- [8] Kripto.ba, *Bitcoin*, 2019. <https://kripto.ba/bitcoin/>, [Accessed: 29. 08. 2024]
- [9] Ycharts, *BitcoinBlockchain Size: Daily report and Interactiv Chart*, 2024. Ycharts,

- https://ycharts.com/indicators/bitcoin_blockchain_size[Accesed: 29. 08. 2024]
- [10] Gupta, M., *Blockchain for Dummies*, IBM, John Wiley & Sons, Inc., Hoboken, New York, 2017.
- [11] Siu, E., *Chapter 2: What is Blockchain Digital Marketing*; Single Grain, s.a., <https://www.singlegrain.com/blockchain/blockchain-digital-marketing-101/> [Accesed: 29. 08. 2024]
- [12] Gem, F., *The impact of blockchain in digital marketing*, Media Update, 2020. <https://www.mediaupdate.co.za/marketing/147788/the-impact-of-blockchain-in-digital-marketing>[Accesed: 29. 08. 2024]
- [13] Drmač, D., *Kakoćeblockchainuticatinadigitalni marketing*; Marketing mreža, s.a.<https://marketingmreza.rs/kako-ce-blockchain-uticati-na-digitalni-marketing/> [Accesed: 29. 08.2024]

Implementation of Artificial Intelligence, Smart Sensors, Robots and Digital Transformation in Food and Agricultural Production

Amra Bratovic^{*1}

Abstract: *Today, we use artificial intelligence every day without even being aware of it to a large extent. This paper systematically presents the development of artificial intelligence and its latest use together with sensors and robots in food and agricultural production. The application of such modern methods of production enables the monitoring of raw materials from the moment of production to the moment of consumption by the end user. Precise mapping and monitoring of crops is possible by using GPS receiver, as well as optimizing management through soil sampling, smart irrigation and fertilizer application. On the one hand, we initiate increased yield of agricultural products, and on the other hand, extended life of food, their health safety and reduced waste generation, which directly affects the economy and the environment. The paper presents concrete examples of the production of meals according to the user's wishes, as well as the development of the first energy drink using artificial intelligence.*

Keywords: *artificial intelligence, sensors, robots, food production, agricultural production.*

1. Introduction

The development of artificial intelligence began back in the mid-20th century, when many scientists and researchers began developing different approaches to create machines that could perform tasks typically associated with intelligent human behavior, such as image recognition, natural language processing, and problem solving. Since then it has taken place in several waves. In the 1970s it was slowed down by the limited computing power available and the lack of suitable algorithms. In the 1980s, it came back into focus thanks to the development of hardware that enabled faster and faster processing of ever larger amounts of data, and the development of neural networks and other machine learning techniques. In the last few years, the development of artificial intelligence has taken a significant leap thanks to the development of deep learning and the huge amounts of data that have become available in digital databases and on the Internet, and thanks to the fact that technology giants such as Google and Microsoft have begun to invest huge funds in artificial

^{*1}Faculty of Technology, University of Tuzla
E-mail: amra.bratovic@untz.ba

intelligence. We use artificial intelligence almost every day, for example to search for certain terms on the Internet, to play games, edit photos, drive using a navigation application (Google Maps), to shop on the Internet or to use a banking application. We use also artificial intelligence in numerous applications and tools such as automatic email classification, Internet search engines, smartphone cameras for face recognition and identification of people in photos, systems for finding new movies and series that we might like on streaming platforms like Netflix and HBO, virtual assistants like Siri, Google Assistant and Amazon Alexa etc.

Artificial intelligence is ubiquitous around us and greatly helps modern man and makes his life easier. All these possibilities arise from the ability of artificial intelligence to imitate human traits such as reasoning, learning, planning and creativity. The computer receives data that has already been prepared or collected using its own sensors, processes it and gives answers. Artificial intelligence systems can adjust their behavior to a certain extent by analyzing previous situations and operating independently.

The market has expanded globally and become accessible to all, thanks to digitization and online business, resulting in intense competition. To stay ahead of the competition in global business, continuous advancement of current technology and the implementation of new artificial intelligence-based technologies are essential. Once considered a topic of science fiction, artificial intelligence (AI) has now emerged as a highly valuable asset in supermarkets and the food industry.

Since the inception of the Internet, AI has played a crucial role in gathering, categorizing, and distinguishing important information from irrelevant data in the vast sea of available information. In the upcoming section, we will examine in detail every aspect of how artificial intelligence is effective in both manufacturing and retail industries.

Artificial intelligence represents great potential for improving the food production and agriculture sectors and provides opportunities for improving effectiveness and encouraging sustainability. However, despite all the advantages, there are fears of inequality and job losses in rural areas, and there are also issues of high costs of implementing AI technologies. The future of food production and agriculture is most likely to be shaped by AI and machine learning technologies with possible applications in various fields – from agriculture to supply chain management [1]. In order for modern systems to be truly sustainable, numerous factors must be taken into account, including the development and application of AI technologies, the policies and regulations governing their use, as well as the way in which society adapts to the changes that artificial intelligence brings. Despite existing challenges such as data protection, privacy, high costs and ethical dilemmas, the future of applying the principles of artificial intelligence in agriculture looks promising [2].

1.1. Monitoring the Health of Crops

Monitoring the health of crops and controlling pests used to be simpler when farmers produced enough for their own consumption. With the shift to industrial agriculture, farmers now manage large areas of farmland, making it daunting to monitor the health of every hectare. However, the combination of AI and Internet-enabled devices can streamline this process. By placing sensors on either plants or soil, farmers can remotely track the well-being of their crops. When coupled with an artificial intelligence or machine learning system, this approach has the potential to generate data-driven insights that enhance plant breeding and overall yields.

1.2. Packaging Improvement and Extending Shelf Life in Supermarkets

Convenient food will become more and more necessary, as will the requirement for packaging free of contaminants [3]. Between 2022 and 2026, smart storage, which encompasses automation and industrial IoT, is projected to expand at a CAGR of 55.2% [4]. Automation of any or all of the packing process can boost output and efficiency while enhancing food safety and lowering the possibility of contamination. Food processing plants generate a huge amount of data. Although sell-by or expiration dates may appear arbitrary to the typical consumer, plant operators can use data collected over months or years to better understand product shelf life thanks to the application of AI systems.

1.3. Increasing Food Safety with Artificial Intelligence Applications

Every time a corporation has to recall products because of contamination, consumers have questions about food safety. To keep food edible, equipment must be cleaned and maintained. Many of these jobs can be accomplished by implementing an artificial intelligence system, increasing industrial production and efficiency without sacrificing safety [5].

In order to monitor the levels of water quality, it can also be added to water sources in food and beverage manufacturing plants. This will enable food processors to guarantee that the water used for sanitizing equipment is free of contaminants. Beverage producers can also use this data to keep an eye on levels of sulfate and chloride, which can alter the flavor of wine, beer, and other produced beverages.

As artificial intelligence enters the food industry, people are beginning to wonder if food or drink created by artificial intelligence is safe and if artificial intelligence helps improve food quality? Food safety is of great importance for both consumers and the food industry. It is very important that the food we consume is safe to eat, and the application of artificial intelligence in the food industry is widely recognized as a promising solution related to this topic. Today, artificial intelligence is an increasingly important tool for improving

transparency, verification, validation, traceability and development, communication and security throughout the supply chain and quality control. For example, only a few decades ago it was man himself who controlled some important characteristics of food quality, but today in the modern age that process has been improved through automated food processing systems controlled by artificial intelligence. Using sensors and cameras, these systems can detect and correct problems such as temperature, liquid levels, gas flow, humidity and even control food hygiene.

1.4. Establishing a Transparent Supply Chain

Transparency in the supply chain is crucial to the operation of any modern company. Customers like to do business with firms who share their commitment to minimizing their carbon footprint and environmental impact. They are constantly searching for methods to do this. Businesses who decide not to provide all relevant information about their supply chain frequently discover that they are losing clients quickly and have to find new tactics to gain them back.

Transparency in the supply chain is also crucial for tracking connections in the event of contamination or the need for resistance. Artificial intelligence, according to research, can assist in precisely predicting crop yields, enabling farmers to provide increased transparency while establishing reasonable expectations for other businesses downstream in the supply chain.

1.5. Reducing Food Waste in The Food Industry

Food waste is a major global issue. In 2021, it was 9.5 million tons in Great Britain. In the US, food waste amounts to around 108 billion pounds each year [4]. By altering the algorithms used to identify food and determine whether it should be sold, artificial intelligence systems can contribute to a decrease in these numbers. Many of these seemingly "ugly" fruit and vegetable products are actually completely edible, but their imperfect appearance causes them to be thrown in the garbage. Reducing food waste and keeping otherwise edible items out of landfills can be accomplished by altering these algorithms to take into account other factors in addition to obvious flaws.

1.6. Tracking the Origin of Food

In addition, artificial intelligence can be used to trace the origin of food products. Using artificial intelligence technology and algorithms, food manufacturers can track food products from farm to table, providing consumers with greater transparency and accountability in the food supply chain. Also, food manufacturers can predict the quality of their products throughout the entire production process, from raw materials to finished products. This technology can detect potential problems or defects in time, allowing manufacturers to take

corrective action and prevent quality problems before they occur. RFID tags and QR codes are commonly used today to improve the level of identification and traceability throughout the supply chain. NFC and Bluetooth Low-Energy (BLE) technologies are so simple that consumers can use them on their mobile phones [6]. Artificial intelligence can analyze data from various sources, such as supplier records, production records and customer feedback to identify potential food safety risks and take proactive measures to prevent contamination or other safety issues. Artificial intelligence in predicting quality is revolutionizing the food industry. Using the power of artificial intelligence, food manufacturers can predict and maintain the quality of their products, reduce waste, optimize production processes and improve food safety. This technology has the potential to change the way we produce and consume food, ensuring that consumers receive safe and high-quality products. As artificial intelligence advances, it will be interesting to see how it continues to shape the future of the food industry [7].

1.7. Improving the Food Industry

There are so numerous applications of artificial intelligence within the food industry that it would be troublesome to list them all here. Artificial intelligence can offer assistance to reshape the segment from the ground up, making a difference to food production security, decreasing waste and moving forward with straightforwardness, all with one program.

Artificial intelligence comprehensively helps companies to understand what their customers want, facilitates and monitors trade processes and also performs monotonous tasks, while leaving room for employees to develop further through acquiring new knowledge, learning, developing creativity and the like, because this is very important in order to be able to later help in the programming and development of systems based on artificial intelligence.

Thanks to speeding up trade processes, reducing the time and costs needed to carry them out, and ultimately increasing the efficiency of business, there are more and more people who understand the importance of using artificial intelligence systems in order to keep up with the times and apply it in their business processes. The possibilities and advantages that artificial intelligence provides in international trade are very large, and it is believed that very soon they will become almost infinite [8].

2. Application Of AI in the Food and Agricultural Production

2.1. Meals Powered by AI

Podravka, one of the leading Croatian multinational companies, was the first in the industry in the region to launch an innovative digital product driven by artificial intelligence (AI), based on advanced LLM technology on the so-called GPT-3.5 Turbo OpenAI model. It is about *SuperfoodChef-AI* by Coolinarika, a

personal AI assistant for a nutritionally balanced diet developed to create an even better user experience on the largest regional culinary platform - Coolinarika, by integrating GPT technology (Fig. 1). In addition to the website, SuperfoodChef-AI is also present on the Coolinarika mobile app available on Google Play or the Apple Store. The first Croatian AI assistant in the kitchen was created in a partnership between Podravka and the leading regional IT company Infobip and the international digital agency C3 Croatia - 01 Content & Technology, and is currently in the beta phase of development, that is, it is still under development and being upgraded.



Figure 1. SuperfoodChef-AI by Coolinarika

Podravka is among the leading brands in the world that have integrated generative AI technology on their digital channels. SuperfoodChef-AI is designed to show users the importance of a varied and nutritionally rich diet in a simple and interesting way, through an interesting conversation. The AI guide helps users with culinary and nutritional advice and recommendations of tasty and healthy recipes, with an emphasis on the so-called superfoods. For example, they will answer questions about the nutritional value of beans and bananas or suggest a recipe rich in proteins. With this use of artificial intelligence, Podravka confirms its focus on innovation and the use of new, modern technologies, not only in production processes, but also in communication with consumers. At the time it was created, Coolinarika was an innovative and unique step forward for the food industry in direct communication with consumers, and today it is one of the most visited portals in Croatia, with almost 70,000 users visiting it every day.

The development of its own AI tool is another step in Podravka's development and improvement of user satisfaction and its relationship with consumers. They believe that this virtual assistant will contribute to the popularization of healthy eating and the promotion of healthy eating habits among visitors to Coolinarika. They continue to work on developments that will further improve user satisfaction, increase their engagement and further improve their user

experience. They are continuously working on the development of AI solutions that enable personalization of communication and provide users with innovative ways to interact with brands. For now, SuperfoodChef-AI has information on 30 foods and a database of about a hundred recipes, but it learns quickly and develops, enriches and supplements the knowledge base every day. The more it is used, the more successful it will become as an assistant in the daily cooking of nutritionally balanced meals [9].

2.2. Drinks Powered by AI

The first energy drink was completely developed by artificial intelligence in cooperation with the leading Hungarian energy drink manufacturer. Moreover, the artificial intelligence did not just come up with a new product, but tested and tasted the three flavors it proposed and, based on its own testing, chose the right one. After all this, the questions arise: How is artificial intelligence capable of tasting, how can it replace human perception, taste and smell? And how does he predict? Artificial intelligence is able to process an incredible amount of information from the web and add new data at a speed almost incomprehensible to humans. However, translating the human sense of taste into an artificial algorithm seems so unrealistic that we investigated how such a process takes place.

Namely, several companies around the world specialize in the digitization of tastes and smells. These companies use specially developed equipment, such as the so-called e-mouth and e-nose. For example, one company in New York has been collecting and processing primary samples for 10 years, digitizing ten years of sensor data and processing the largest and most diverse set of food and beverage sensor data on the market. This digital database is constantly expanding and updating, enabling accurate monitoring of changes and trends in perception and preferences. This is where artificial intelligence comes in, which can reliably predict from digital data entry how different consumers will react to a new product - without tasting it live. Hell was digitized by a New York based company and after analyzing a lot of data and statistics, the artificial intelligence used predictive intelligence to choose which flavor would be best for their new product. Artificial intelligence, in fact, can use the collected data to understand how different demographic groups perceive tastes, what they like and what they don't like. Based on this, artificial intelligence algorithms can predict how each demographic group (based on age, gender, socioeconomic status and place of residence) will like the taste, smell and texture of a new product. This recent innovation has resulted in, for the first time in the world, artificial intelligence being used in such a complex product development process in the energy drinks sector. The resulting product, Hell AI, one of Europe's most modern beverage production lines is already available to consumers (Fig. 2).

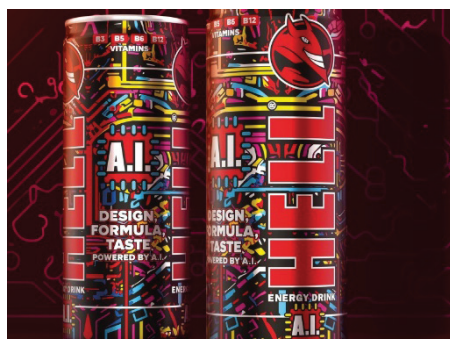


Figure 2. "Hell A.I." energy drink produced by the Hungarian company Hell energy

It is obvious that artificial intelligence has enormous potential in the field of product development, breaking down previously unimaginable boundaries and bringing new opportunities to the food industry. The ability of artificial intelligence to sort and analyze the vast amount of data available on the web in minutes or seconds can shorten the product development cycle from one to two years – or even half a year, at best – to one month. There are currently more than 100 artificial intelligences in the world - all of them are excellent in different areas and their combined action guarantees a high level of innovation. Therefore, one should definitely pay attention to the culinary revolution, which seems to be full of surprises that artificial intelligence can bring to the food industry in the near future. The possibilities offered by artificial intelligence are endless, and we are only just beginning to discover what it is capable of [10].

Use of technology Many companies today are turning to the use of artificial intelligence because it can be of great help in various industries. The capabilities of systems based on artificial intelligence are such that they process large amounts of data with the help of algorithms, make decisions in a short time, provide automated data analysis, and thus shorten many processes and save time. A recent example of the use of artificial intelligence comes from the Hungarian company Hell Energy which used artificial intelligence to develop every element of a new product from its taste, carefully selected ingredients, design and even marketing elements and guess what - the new product was even tried by artificial intelligence! So, artificial intelligence has already proven its value in various industries, from healthcare to finance, and now it is slowly entering the food industry as well.

2.3. Artificial Intelligence in Agriculture

Artificial intelligence (AI - Artificial Intelligence) in agriculture, by applying different methods of machine learning, algorithms and sensors, increases

efficiency, total yields and reduces costs in food production. Over time, agriculture has adapted to new technologies and scientific discoveries, with the aim of increasing productivity and optimizing food production. In the last few years, one of the most influential technological trends transforming agriculture is artificial intelligence (Fig. 3). Artificial intelligence (AI) refers to the ability of computer systems to perform tasks that typically require human intelligence, such as image recognition, natural language interpretation, decision making, and problem solving. AI in agriculture has started to gain importance in view of the growing global challenges such as climate change, loss of biodiversity and increased demand for food due to the growth of world population. The application of AI in agriculture is only in its initial phase and its impact is expected to only increase in the future.



Figure 3. Application of drones in precision agriculture

Precision agriculture represents a synergy between advanced technologies in the era of global information and established agricultural practices, and implies an integrated management approach that tries to adapt to the needs of crops on individual parts of the agricultural area. Precision agriculture distances itself from traditional practices through the application of advanced technologies such as GPS receivers that enable precise mapping and tracking of crops, thereby optimizing management and, for example, soil sampling methods and fertilizer application, for example, in adjusted and variable rations. Remote sensing methods are used to assess crop health and detect plant stress. The amount and speed of information that can be obtained through modern information systems change the traditional approach through the level of management that is specific to smaller units within the board or field. Key tools in precision agriculture include GIS systems, which are used to collect, visualize and analyze data, providing better insight for management decision-making. Efficient information management is essential for the application of precision agriculture techniques

and requires clearly defined goals and input parameters for decision-making. Precision agriculture enables farmers to use resources more efficiently, such as fertilizers, pesticides, biostimulators, feed, irrigation, etc., thereby achieving a higher yield and quality of crops with a reduced rate of environmental pollution with more efficient and rational agrotechnical procedures and reducing the impact of agricultural production on the soil, air and water [11]. Agricultural robots are automated machines or robotic systems capable of performing tasks in an agricultural environment. The use of robots in agriculture offers a number of advantages including an economically viable way to keep up with growing demand through cheaper production (Fig. 4). Agricultural robots, such as those intended for harvesting or packaging products, are becoming more and more present in the agricultural sector. Autonomous machinery and drones are especially used on large agricultural areas. As technology and information systems advance, so do the number of tasks these robots can perform, including harvesting, sowing and even livestock tasks like shearing sheep and milking cows.



Figure 4. Collecting information from the field

2.3.1. Application of robots in agriculture

Predictive analytics tools use artificial intelligence, machine learning and advanced algorithms to predict future outcomes in agricultural production. These tools enable the analysis of large amounts of agricultural, biological, climatic, hydrological and economic data to predict crop outcomes, management changes, and other key issues in agricultural systems. Through crop modeling, the application of machine learning techniques and data analysis, precise forecasting and better decision-making is possible. With advances in analytical techniques, precision agriculture is being raised to a new level. One of the key advantages of using the mentioned models is the possibility of anticipating problems before

they appear. The key for the farmer is to anticipate and prevent problems, such as crop diseases or unfavorable growing conditions [12]. This proactive approach not only minimizes risks, but also saves time, money and resources, instead of dealing with the consequences after a problem occurs. This strategy, in addition to preserving profitability, additionally ensures the sustainability of agricultural operations. The advantages of using such systems are numerous and include: selecting the best crop for certain climatic conditions, optimizing management procedures, improving crop productivity and making decisions in real time that are sustainable and economically profitable in the long run. In addition, the algorithms used to predict future outcomes can play a key role in reducing negative climate change, allowing farmers to better understand the impact of their production on the environment (Fig. 5).



Figure 5. Application of robots in agriculture

3. Conclusion

In this work, we have seen that the application of AI as a tool has a positive effect on increasing productivity and product quality. Although it cannot be said with certainty, but after technological innovations, i.e. the introduction of AI in food and agricultural production, new jobs will be created that will require new competencies of workers. Digitization, standardization and automation are all trends that no industrial sector can avoid. However, strengthening education and vocational training is extremely important for the development of new skills and occupations so that people do not become unemployed, because otherwise people would become incapable of participating in the development of society and technology.

4. References

- [1] Bratovic, A. (2022) Smart and sustainable food processing sector by integration of the concepts of the technologies behind Industry 4.0, Conference: International scientific conference Application of Industry 4.0 An opportunity for a new step forward in all industrial branches, 146-163. DOI: 10.5644/PI2022.202.28
- [2] Brandić, I. (2023) Kako umjetna inteligencija mijenja način proizvodnje hrane?, Poljoprivredna revolucija (gospodarski list) <https://gospodarski.hr/rubrike/nove-tehnologije/kako-umjetna-inteligencija-mijenja-nacin-proizvodnje-hrane/>
- [3] Bratovic, A. (2021). Physical – Chemical, Mechanical and Antimicrobial Properties of Bio-Nanocomposite Films and Edible Coatings. *International Journal for Research in Applied Sciences and Biotechnology*, 8(5), 151–161. Retrieved from <https://www.ijrasb.com/index.php/ijrasb/article/view/236>
- [4] [4] Steve Wynne-Jones, How AI Technology Can Make Food and Beverage Better, 20.07.2022. <https://www.esmmagazine.com/technology/what-unique-applications-of-ai-can-make-food-and-beverage-better-analysis-180918> accessed on 23.06.2024.
- [5] Bratovic, A. (2023) Najnovija dostignuća u primjeni senzora, umjetne inteligencije i digitalne transformacije u prehrambenoj industriji, Časopis za savremena privredna kretanja, 1(1):76-84.
- [6] Bratovic, A. (2023) Sophisticated Transformation of Chemical, Pharmaceutical and Food Industry by Implementation of Digitalization and Automatization, in *Proceedings - International Scientific Conference Basic Technologies and Models for Implementation of Industry 4.0*, 21, 241-253. DOI: 10.5644/PI2023.209.11
- [7] Umjetnainteligencijakaosvevažnijalat u prehrambenojindustriji, Čapljinski portal, <https://caportal.net/2023/11/08/umjetna-inteligencija-kao-sve-vazniji-alat-u-prehrambenoj-industriji/>
- [8] Mesic, I. (2022), Učinkovitost umjetneinte ligencije u prehrambenoj industriji, In Store, <https://instore.ba/ucinkovitost-umjetne-inteligencije-u-prehrambenoj-industriji/>
- [9] Spiljak, L. (2023) Podravka razvija prvog AI asistenta u prehrambenoj industriji u regiji
- [10] <https://www.tportal.hr/tehnolo/clanak/podravka-razvija-prvog-ai-asistenta-u-prehrambenoj-industriji-u-regiji-20231012>
- [11] Kako umjetna inteligencija može kušati i mirisati napitke?, Progressive, <https://progressive.com.hr/?p=28529>
- [12] Bratovic, A. (2022) Positive Aspects Of Nanotechnology On Agricultural Sustainable Development: Application Of Nanoparticles And Fibers For

- Increasing Agricultural Yield, *International Journal of Agriculture and Environmental Research*, 8(6): 780 – 798. doi: [10.22004/ag.econ.333370](https://doi.org/10.22004/ag.econ.333370)
- [13] Bratovic, A. (2024) Role of Nanocomposites Using Graphene-Based Materials for food/Toxin-Sensing Applications in Agriculture, In Book *Graphene-Based Nanomaterials*, eBook ISBN 9781003300540. <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003300540-10/role-nanocomposites-using-graphene-based-materials-food-toxin-sensing-applications-agriculture-amra-bratovic>

A Comprehensive view of the Application of AI in Recruitment and Selection

Edisa Dreko¹, Isak Karabegović², Žaklina Teofilović³

Abstract: *Using AI through industries and business processes is increasingly becoming the subject of theorists and practitioners. In the HRM process, the use of AI gives companies numerous advantages in employee performance, and processes, but also presents them with organizational, financial, technical, legal, and personnel challenges. This paper explores the application of AI systems in recruitment and selection through gamification strategies, people analytics, talent intelligence, AI platforms, video interviews, and conversational AI. It provides an overview of the benefits and challenges associated with their implementation. Additionally, the paper delves into ethical considerations and legislation, focusing on the EU Act, domestic laws, and ISO AI standards. The primary goal of this paper is to provide a comprehensive understanding of AI's role in HR processes and the complexities of implementing AI solutions in recruitment and selection.*

Keywords: *AI application, recruitment, selection, ethical consideration, legal regulation, benefits, challenges*

1. Introduction

The implementation of Industry 4.0 technologies in human resource management processes is leading to significant changes in the way work is approached. Changes related to the way of work include strategic changes in the organization of the human resources management function, which refer to changes in the way of hiring, evaluating and retaining existing employees, creating an innovative and digital organizational culture. These changes also refer to the application of new concepts in HR such as gamification, people analytics and managing a multigenerational workforce and employee experience. When considering changes in human resource management [1], it is crucial to prioritize the digitization of HR activities and provide retraining and additional training for HR employees, particularly in digital and analytical skills. HR professionals should be deeply involved in all aspects of HR digitalization, ensuring alignment with the company's goals. HR business partners need to

¹University of Niš, Faculty of Economics, Serbia

²Academy of Sciences and Arts of Bosnia and Herzegovina, B&H

³University of Belgrade, Faculty of Economics, Serbia

E-mail: edisanp@gmail.com, isak1910@hotmail.com, Zaklina.Teofilovic@mtuserbia.com

work closely with business leaders to align HR strategies with business goals, requiring digital skills in Industry 4.0. Operational managers should also have more decision-making authority, which will help decentralize decision-making. They can utilize digital tools for activities such as recruitment and performance management, in order to enhance efficiency and decision-making.

Meanwhile, HRM is responsible for ensuring that the workforce is equipped to adopt these technologies in different business processes through training, talent development, recruitment, and retention practices.

AI in HR processes can be used in different sub-processes of the HR process (Figure 1).

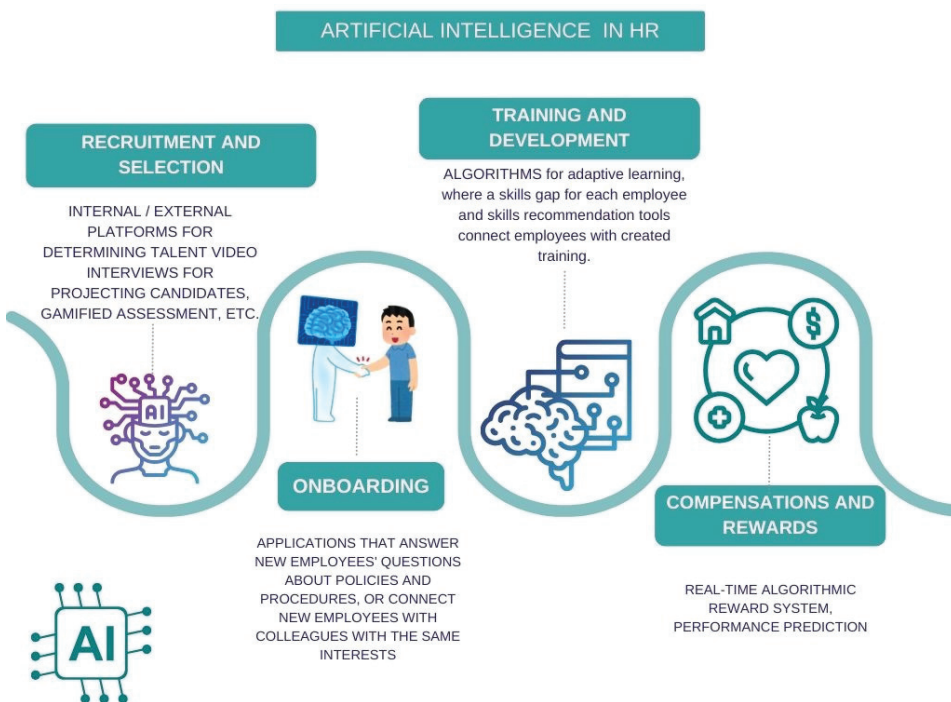


Figure 1. Artificial intelligence in HR processes

The applications of AI in HR processes outlined in Figure 1. represent some of the most prevalent uses. It is essential to understand that these applications are continually expanding in various ways. The advancement of both general and generative artificial intelligence, along with its integration with other Industry 4.0 technologies, broadens the range of its potential applications.

Research [2], shows some interesting trends in examples of AI in companies in the sense that senior executives believe (81%) that digital transformation, the needs of employees for flexible working conditions, and the lack of talent will

make the process of talent planning significantly more challenging in the coming period. Also, most of them (78%) believe that their organizations may not be able to do the processes of retraining employees quickly enough to the requirements of technology application in the next three years. That makes the process of talent management especially difficult. In contrast to this research, when it comes to generative AI, Deloitte's research [3, p.19-20], shows that companies are aware that the application of generative artificial intelligence will require major changes in talent management strategies, and 47% of respondents stated that they are investing high efforts in educating their employees on the application of generative AI within their positions. These efforts are related to the recruitment and engagement of technical talents to launch generative AI initiatives, the education of the existing workforce and the benefits and risks of applying generative AI, as well as reskilling of those employees who are most amenable to the use of generative AI. The same survey reveals [3, p.13] low adoption of generative AI in HR, with only 23% of respondents having adopted gen AI in HR, compared to 41% in marketing and sales, product development, and production.

2. AI in Recruitment and Selection

Artificial intelligence is leveraged in human resources through different applications and platforms for attracting and recruiting candidates, as well as talent intelligence [4], gamification and people analytics [5, 6, 39].

Chatbots

The simplest and most widely used AI application in recruitment and selection is *Chatbots* a communication AI tool to provide information to candidates regarding the position they are interested in and information about the company, policies, and procedures. Companies use this form of AI because, in addition to answering candidates' questions in real time, the chatbot can simultaneously process various inquiries, schedule interviews, and provide feedback to candidates regarding their applications. The automation of routine tasks in recruiting improves efficiency and enhances the candidate experience [7]. The user experience depends on the design and functionality of the AI tool itself, as well as its ability to personalize responses and the human element that is of particular importance for the experience of Generation Z as users of this AI tool [8]. AI chatbots are unable to make decisions, but they can assist recruiters. Implementing a conversational AI system [9](commonly known as a chatbot) to handle tasks such as recording, reviewing, and scheduling interviews can result in significant time and cost savings in the hiring process. Additionally, it can contribute to a higher retention rate of new employees. Automation in these

activities improves the human resources sector by leaving time and space for more creative activities and provides users with a better user experience.

People analytics

People analytics can be classified into different levels of maturity: descriptive, predictive, prescriptive, and autonomous analytics. The key distinction among these maturity levels lies in how data is utilized. Descriptive analytics is used to analyze and understand performance, predictive analytics is used to forecast future performance based on identified patterns, and prescriptive analytics is used to make suggestions for decisions based on the results of predictive analytics. Descriptive analytics, for instance, can help in identifying the most effective source of candidate recruitment, factors influencing a candidate's success in a specific position, and the efficiency of the recruitment process, including any bottlenecks. Empirical studies [10] indicate that predictive analytics for forecasting candidate achievement before CV screening, based on historical information about employees, their performance, and attributes, can significantly reduce recruitment expenses. This is because it helps exclude candidates who lack key characteristics for a certain position even before the CV screening stage. The use of prescriptive analytics through the use of statistical algorithms and machine learning models can significantly improve the prediction of the need for employment, and based on the analysis of historical data, discover patterns that with a high probability predict the success of candidates for a given position before employment [11] and recommend and personalize the strategies of attraction, recruitment, selection and retention. Autonomous human analytics "autonomously makes and executes decisions based on a self-learning algorithm, without human insight"[12]. When it comes to autonomy in decision-making, descriptive and predictive analytics applications can be used as decision support. Prescriptive analytics, although highly influential in decision-making with a high probability that the proposed actions will be adopted, still leaves the autonomy to the people. On the other hand, in autonomous analytics, a large part of the process is automated without human interaction.

Although the benefits of using people analytics are numerous, there are also certain disadvantages of using people analytics, such as [12]:

1. excessive reliance on analytics can negatively affect managerial abilities (cognitive abilities of critical thinking, independent interpretation of situations and reasoning) reducing their decision-making autonomy,
2. negative impact on employees (through the mechanization of thinking, reduction of cooperative decision-making, reduction of direct communication),

3. insufficient transparency and understanding of the functioning of AI algorithms and lack of clear responsibility,
4. reduction of innovation, flexibility and the need to adapt to changes due to excessive dependence on the established path of historical orientation,
5. bad decisions due to the illusion of algorithm control and data quality,
6. taking predictions as self-fulfilling prophecies.

Talent Intelligence

In addition to people analytics, which takes into account internal data, a special use of AI in recruitment and selection is talent analytics, which combines an organization's internal data with external data from the market based on a skills-based approach, to assess the needs of companies in the field of labor and the business development goals. Its benefits in recruitment are more efficient workforce planning, more agile access to talent, and reduction of recruitment costs. One of the most well-known AI platforms for TI is the Eightfold Talent Intelligence Platform based on deep-learning AI models, which aims to find qualified internal and external candidates and fill the gap in talent availability, with unbiased selection with a personalized experience. Many companies use this AI platform[13] to find, select, develop, and retain talent.

Gamification

Gamification is „the application of game elements in a non-game context“ [14], which aims to achieve benefits related to the performance indicators of the processes in which the game is applied. Thus, within the human resources framework, gamification is most often used in recruitment and selection, onboarding, training and development, and to ensure the well-being of employees [1]. In the selection process, gamified assessments are interactive, objective, and inclusive assessments of the candidate's skills and behavioral characteristics. The design of games helps ensure interactivity and increases candidates' motivation to participate. The gamified assessment measures various behavioral, numerical, motor, and cognitive abilities without candidates being mindful of what is being evaluated. This provides employers with unbiased and objective insights into individual candidates. Unlike conventional assessments, where candidates can easily tailor their answers to fit socially acceptable behavior, gamification avoids this bias by not revealing what is being measured. Thus, using game design elements related to: "self-presentation, social interaction, immersive and engagement, self-challenges and challenges to others, choices and unpredictability" [5], companies can gain insights into the candidate's abilities for specific positions. Some research shows that the choices of game elements and game design are closely related to the results of the game

in the learning process [6]. The AI-based algorithms behind these games facilitate the selection of suitable candidates based on pre-defined selection criteria.

AI video interview

Candidate selection based on AI video interviews and gamified assessments provided by various providers (Spark Hire, Hire Vue, Pymetrics, ZenHire, and others) is increasingly used by companies of various sizes. These providers combine insights from psychology and behavioral economics with technological know-how from ML, neural networks, and deep learning to create their AI tools. In the selection, predictive analytics is used, where, based on the predictive assessment of talents, a comparison can be made with the job profile as a set of competencies that positively affect job performance[15].

Some of these AI systems are focused on specific areas of business (Business Process Outsourcing, Call Centers), enabling the *automation of workflows* from position advertising, CV review, AI interviews, interview scheduling, *personalized feedback to candidates* with suggested opportunities for improvement (through a personalized video by HR which has a positive effect on the employer brand and a high Net Promoter Score where employees become promoters of the company), *customized testing with the assessment of skills and abilities* characteristic of the respective roles, as well as a *significant level of real-time analytics for various employee performance metrics or recruitment sources, etc.* [16]. Apart from AI systems oriented towards specific niches in recruitment and selection, there are others more generalized concerning the niche criterion. These AI systems assess candidates' skills, abilities, and attitudes, can measure attributes for different jobs, and analyze the most important ones for a given job. AI system can evaluate job attributes related to work style and personality type (through reliability, initiative, flexibility), domain knowledge and skills, candidate's ability to work with information and data (cognitive abilities and problem-solving skills), ability to cooperate with people (through the skills of emotional intelligence, communication, and negotiation) [17]. That assessment of the candidate can be done in the form of a game an AI-supported interview or a combination thereof. Gamification in selection occurs in psychometric games lasting no more than 15 minutes per game, which measure cognitive abilities, emotional intelligence, and personality characteristics[17]. AI Video interviews are used to assess the skills, abilities, and attitudes as well as personality traits of candidates, based on AI speech analysis [17].

Certain AI systems in terms of transparency provide users and the public with information (Explainability statement) by which they are described in more detail [18]:

1. The purpose of AI systems such as video interviews used as decision support tools in employment, is to help recruiters effectively and efficiently find qualified candidates. This is achieved through the use of an AI algorithm that aims to avoid or minimize human biases in the assessment process.
2. The way of using the meaning of AI techniques and factors - The way to ensure this is to control and test the biases of only the algorithm in each of the phases of the algorithm's life cycle following legal procedures, so if possible biases are noticed, then it is possible to minimize those data points. An explanation of how the tool ensures consistency in interviewing (by avoiding recruiter bias), equal opportunities (all candidates have the opportunity to interview at a time and place that suits them), and how it manages the candidate experience (through the possibility of re-answering if the candidate is not satisfied with their previous speech, an evaluation report as feedback). Companies can modify the AI solution in part of the data set used and their estimated impact on the assessment results.
3. Design of an AI solution that includes: *converting words into text* using the ML model, *understanding individual words and complete sentences* using the NLP model, *evaluating and scoring each candidate* using a regression ML model that evaluates the candidate's answers according to the competencies that the company wants to examine in the evaluation.
4. Such explanatory statements include information on how and when the AI system is tested, how the risks are managed, how AI suppliers are chosen to create solutions and how to cooperate with third parties, how system changes are made, and how and who performs system audits.

It's important to understand that these statements on explainability highlight the need for AI system providers, users, and other stakeholders to comprehend how the AI system functions so they can provide input on whether to accept or reject the system. Trust in AI systems is low because it's difficult to explain how they work and predict their strengths and weaknesses. Legal regulations require everyone involved in creating and using AI systems to align their practices with legal requirements. As a result, we can expect to see more similar documents (such as strategies, statements, rulebooks, and procedures) in the future.

2.1. Benefits and Challenges

It is crucial to thoroughly evaluate the role of artificial intelligence in the recruitment process to gain a comprehensive understanding of its functionality,

including the advantages, challenges, and potential risks for the organization, its employees, and other stakeholders.

Table 1. The benefits and challenges of AI in recruitment and selection

BENEFITS	
➤	More effective candidate assessment through improved candidate matching [19]
➤	Greater recruitment efficiency [20]
➤	Automation of parts of the recruitment process (CV review) [9]
➤	Low employment costs [22]
➤	Reduced the bureaucracy of HR and time and cost in the process of selecting candidates [22]
➤	Minimized human bias [23]
➤	Improved user experience through more dynamic and early interaction with hiring managers[24]
➤	Hiring based on skills and abilities, not based on a resume [24]
➤	Higher employee satisfaction and better performance with better NPS (Net Promoter Score) [16]
➤	Increased diversity and equality in employment
➤	Strengthened the employer brand
CHALLENGES	
➤	Data privacy and security [21, 19]
➤	Fairness and unbiased AI algorithms [21]
➤	Algorithmic equity in uniform environments, such as talent acquisition in IT where an AI system is deployed [23]
➤	Upholding non-discrimination and human rights principles [19]
➤	Data quality and validity
➤	Transparency of the algorithm
➤	High costs of development and utilization
➤	Organizational and technological readiness for adopting AI solutions
➤	Candidates' willingness to use (regard for data privacy and trust in the data management system).

It is crucial to prioritize the integration of AI systems into legal procedures. However, due to the complexity involved in developing and using AI systems, including technical, organizational, and employee competency factors complying with legal regulations may take longer and be more costly. This could significantly slow the progress and adoption of AI systems.

2.2. Ethical Considerations and Legal Regulation

When considering the ethical use of AI and adherence to legal regulations in the hiring process, the main challenges lie in safeguarding human rights and complying with the legal standards of the markets where the AI tool is employed. As per a study by researchers [25], the primary ethical concerns raised by both businesses and job seekers regarding the use of AI in recruitment include:

1. AI algorithm biases that can compromise the principles of *non-discrimination* [23]. The AI system should promote fairness and eliminate any kind of bias, but if the data used to build the algorithm contains biases or reflects socioeconomic inequalities, then the AI system results in an unfair outcome [26] when the data set on which the AI model is trained is dominated by one certain race or sexual orientation [27];
2. *Preservation of human autonomy* in decision-making, because automated decision-making is also problematic from a legal and practical point of view.
3. The degree of *validity* of predictions based on data (because, unlike humans, AI cannot recognize complex emotions or assess the impact of certain candidate abilities such as charisma on job performance, does not understand the candidate's value system, etc.);
4. Adequate *data management*, especially in the domain of *privacy and security*;
5. *Transparency of AI solutions in decision-making* because an explainable AI system enables users to trust and minimize harmful decisions [26].

To detect AI algorithm bias problems (bias in a non-representative data set used for training AI solutions or machine learning models), companies use tools for detecting and minimizing bias [23] such as What-If Tool, Fairlearn, AI Fairness 360 (AIF360), FactSheets for AI, Audit-AI algorithm bias detection tool. These tools can be in different forms from open source applications where according to multiple fairness metrics the ML model is tested using the visualization of the analyzed ML model and in the form of a What-if tool to examine the model in different scenarios of data sets and subsets [28] or audit tools ML models and in all - "detection and minimization of discriminatory patterns in data training and machine learning predictions for socially sensitive decisions" [29].

In the European Union countries, regulations regarding artificial intelligence (AI) are based on the Regulation on Artificial Intelligence (AI Act) as well as legal procedures governing labor relations and data privacy. In Serbia, there is no specific legal regulation concerning the use of AI in employment. However,

indirect regulation occurs through compliance with laws such as the Personal Data Protection Act, the Labor Act, and the Anti-Discrimination Act.

Regarding the Personal Data Protection Act, special attention in AI solutions should be focused on the basic principles of personal data protection [30, p.55] (legality, fairness, transparency, limitations concerning the purpose of processing, minimization of personal data, completeness and accuracy, limited storage of personal data, integrity and confidentiality, operator's responsibility for actions related to personal data) and individual rights related to the protection of personal data [30, p.90] (right to information, right to access, the right to correction, the right to be forgotten, the right to limit processing, the right to transfer, the right to object and rights related to automatic data processing). In the realm of employment solutions, the integration of artificial intelligence (AI) in handling personal data must adhere to data protection principles to ensure lawfulness, fairness, and safeguarding of individual rights. This necessitates a legal foundation for the collection and processing of data, typically exemplified by the candidate's consent for data collection or the requirement to fulfill a specific agreement. The AI solution should adopt an impartial, non-discriminatory approach, devoid of bias towards candidates during their tasks. In terms of transparency, candidates should receive clear information regarding the entity collecting the data and their methods for processing, utilization, and storage of the data. According to this Law, data processing operations may be undertaken only for *specific, explicit, and justified purposes* [30, p.66], e.g. for recruitment selection, and may not be used for other purposes without the additional consent of the candidate. The *principle of minimization of personal data* in the employment context refers to the use of the smallest possible amount of data about the candidate, mainly the information related to the assessment of the candidate's suitability for the required position during recruitment. The *principle of data storage limitation* refers to the period in which the collected and processed data about the candidate is stored. When the purpose for which the data was collected and processed is fulfilled, the need to save the data disappears, that is, the basis of the processing is lost and the data must be deleted. Integrity and confidentiality in data processing to prevent abuses are achieved mostly through adequate personnel and technical and organizational security measures. The aim is to prevent losses, damages, or other errors that may arise from data processing. The *principle of responsibility* refers to the obligation of the person who collects, processes, and stores data to do so following legal requirements. This responsibility is expressed through internal procedures such as "personal data protection strategy, creation of internal procedures related to the method of data processing, mapping of activities related to personal data, determination of the person authorized for data protection or other persons who are responsible for personal data, - ensuring appropriate training of employees in connection with the protection of personal

data, rulebook related to efficient management of personal data breaches, impact assessments in special cases of processing, control of implemented measures" [30, p.75]. Apart from the Law on the Protection of Personal Data, AI applications should be harmonized with the regulations from the Law on Prohibition of Discrimination [31] based on gender, ethnic and racial affiliation, religion, disability, age, and sexual orientation in the field of work, which guarantees equal employment opportunities and equal rights (Article 16) such as "the right to work, to free choice of employment, to advancement in the service, to professional training and professional rehabilitation, to equal remuneration for work of equal value, to fair and satisfactory working conditions, to rest, on education and joining a trade union, as well as on unemployment protection".

The EU AI Act classifies the application of AI in employment (recruitment and selection) and employee management (working conditions, assignment of tasks, authorizations and responsibilities, evaluation and promotion of employees, termination of contractual relations) as high-risk systems [32], because these systems may have significant impact on the career, financial, physical and psychological well-being and human rights of employees. Algorithms can lead to issues regarding the protection of candidates and employees from discrimination, as well as data privacy.

AI systems classified as high-risk must adhere to specific requirements, including[33]:

- Establishing, documenting, and maintaining a comprehensive risk management system throughout the life cycle of the AI system
- Managing data to ensure the provision of high-quality datasets for developing, evaluating, and testing AI solutions
- Developing and maintaining thorough technical documentation outlining system design, advantages, limitations, and compliance with the requirements stipulated in the AI Act
- Automatically recording the traceability of the functioning of the high-risk AI solution throughout its entire life cycle
- Ensuring transparency of user instructions in digital format, providing comprehensive, accurate, clear, and user-friendly information
- Providing human supervision through an appropriate human-machine interface to mitigate risks related to data, health, or user safety and ensuring robustness, accuracy, and cybersecurity throughout the life cycle of AI solutions.

The EU AI Act also prescribes the obligations of users and suppliers of high-risk AI systems [34]. In the hiring process, most companies are like users, as developing their own AI solution is a financially, organizationally, and personnel-demanding project. It's their responsibility to ensure they have the necessary organizational and technical requirements for the correct use of high-risk AI systems according to the provider's instructions. They should assign the

supervision of the system to a knowledgeable and skilled employee, providing them with technical support. Continuous monitoring of AI systems for potential risks and timely notification of suppliers in case of any risks occurring is crucial. Additionally, automatic logs generated by high-risk AI systems should be kept for at least six months, unless otherwise specified in the legal regulations on personal data protection. If, on the other hand, a certain company decides to develop its own AI solution in the field of employment that is marked as high-risk, or if it significantly modifies someone's solution that is already in use on the market, then it will be considered a provider of a high-risk AI system [35] and will have the obligation of adequate risk management, documentation through the maintenance of an automatic diary, the implementation of a QMS system, undertaking corrective actions and cooperation with regulatory bodies, training the system on quality data that excludes biases and encourages fairness in decision-making. When developing high-risk AI systems, organizations must have a QMS meeting EU Act requirements for policies, procedures, and instructions [36] which refers to:

- strategies whose main goal is to comply with regulations through procedures for managing modifications of AI systems and procedures for assessing compliance with regulatory requirements,
- procedures concerning the projection control and verification of AI system design,
- procedures related to quality management in terms of development quality control and quality assurance of AI systems,
- testing, testing, and validation procedures of AI systems, before during development, and after development,
- technical documentation and specification as well as applicable standards
- procedures related to data management (collection, tagging, storage, analysis, filtering, mining, aggregation, retention)
- procedures related to the risk management system (establishment, implementation, documentation, and maintenance) as a process of continuous improvement through the life cycle of the AI system. This system should cover potential risks to fundamental rights, health, and safety. These risks should be identified, analyzed, and defined and adopted measures for their management.
- procedures for monitoring AI systems on the market and reporting incidents,
- procedures for managing communication with all interested stakeholders
- procedures for clearly establishing authority and responsibility for employees regarding the AI system
- procedures for keeping records of relevant information and documentation

Companies that develop, provide AI products and services, modify and use AI have at their disposal various guides for the responsible creation and use of AI systems, and among the most important are the standards of the International Organization for Standardization [37] ISO/IEC 42001:2023, ISO/IEC 23894:2023, ISO/IEC 23053:2022, ISO/IEC 27701:2022. AI Management System Standard ISO/IEC 42001:2023 contains requirements and guidelines for the establishment, implementation, maintenance, and continuous improvement of AI systems. The goal is to ensure the responsible development and use of AI systems from the point of view of ethical norms, transparency, and continuous improvement. It is intended for organizations that not only develop an AI system but also use it, regardless of the size and form of the organization. The requirements of the standard include special requirements in terms of risk management in autonomous decision-making, data analysis, and ML, as well as AI system training. ISO/IEC 23894:2023 refers to the risk management of AI systems and is particularly important as a support for meeting the requirements of the EU Act because it provides specific instructions on how to identify risks and how to effectively manage them. This is a guide based on which organizations can develop their practices and procedures for identifying, assessing, and mitigating the risks of AI systems, to ensure that the system is fair and non-discriminatory, secure from the point of view of data security and privacy, that it is robust and does not produce negative effects, and of course transparent in terms of decision making. The ISO/IEC 23053:2022 standard is a framework for understanding AI systems based on machine learning by explaining how to solve problems by building, testing, modifying, and evaluating ML models, using a variety of data (related to training, validation, testing, and use) and software tools and techniques (for data preparation, the definition of algorithm category, optimization methods, and ML model evaluation criteria) [38]. ISO/IEC 27001:2022 refers to the management of information, privacy and cyber security and their alignment with legal frameworks.

When talking about the ethical use of AI systems, it is important to emphasize that it is achieved through education about ethical concerns, cases of ethically problematic AI systems, legal frameworks, and requirements regarding the ethical repercussions of AI systems, as well as through the design and implementation of strategies, procedures, and rules. development or use of the AI system, audits, and reviews for each phase of the life cycle of the AI system.

3. Conclusion

It is important to understand that AI requires significant organizational changes in processes. This necessitates a highly qualified workforce and additional financial investments. The benefits of utilizing artificial intelligence in hiring and choosing candidates are apparent in improved organizational, financial, market, and operational outcomes. Along with the benefits, it's important to consider the obstacles and the legal requirements that companies must meet if they opt to adopt AI solutions or create their own. Approaching the sustainable use of AI in business processes involves a mix of legal regulation, technology, and ethical standards. With the complexity of AI systems, there is a growing need to ensure transparency, validity, fairness, decision-making autonomy, and accountability. This necessitates the association and inclusion of many stakeholders in the AI system development process, which may complicate the fulfillment of EU AI Act standards and slow down the development of these systems. It is also necessary to assess the long-term effects of AI systems on employees and the labor market. Companies require a competent workforce with cross-functional knowledge to audit compliance with the requirements of AI system providers, suppliers, and users. Enhancing employees' skills in this area will be a top priority in the upcoming period. Therefore, when it comes to applying AI systems, it is essential to take a comprehensive approach to understanding both the possibilities and limitations of AI to make informed decisions.

4. References

- [1] Drekočić E, Teofilović Ž. (2024). A chapter Upravljanje ljudskim resursima kao kritični faktor digitalne transformacije preduzeća, in a book "Industrija 4.0: digitalna transformacija oblikuje budućnost" (Karabegović I. i Majstorović V.), Društvo za robotiku u Bosni i Hercegovini and Akademija nauka i umjetnosti Bosne i Hercegovine, ISBN 978-9958-9262-8-0, ISBN 978-9926-410-94-0, pp 371-394
- [2] World Employment Confederation (2023). Agile talent in the age of AI. <https://insights.wecglobal.org/the-work-we-want/home/agile-talent-in-the-age-of-ai> [Accessed:17.07.2024]
- [3] Now decides next: Insights from the leading edge of generative AI adoption Deloitte's State of Generative AI in the Enterprise Quarter one report. (2024). Deloitte Development LLC <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-state-of-gen-ai-report.pdf> [Accessed:17.07.2024]

- [4] Eightfold.:Unlocking the power of Talent Intelligence: A definitive guide.(2024). <https://www.youtube.com/watch?v=jiXj6YB8lhE&t=1159s>
- [5] Bina, S., Mullins, J., & Petter, S. (2021). Examining game-thinking in human resources recruitment and selection: A literature review and research agenda. <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/3fa22381-291c-48ca-9ec9-1554622d5e7e/content>
- [6] Kulkarni, P., Gokhale, P., Satish, Y. M., & Tigadi, B. (2022). An empirical study on the impact of learning theory on gamification-based training programs. *Organization Management Journal*, 19(5), 170-188, p.183.
- [7] Koivunen, S., Ala-Luopa, S., Olsson, T., & Haapakorpi, A. (2022). The March of Chatbots into Recruitment: Recruiters' Experiences, Expectations, and Design Opportunities. *Computer Supported Cooperative Work (CSCW)*. <https://doi.org/10.1007/s10606-022-09429-4>
- [8] Kumar, S., Chaurasia, M., & Singh, P. Ai based chatbots in recruitment: Generation Z users experience. *Manager - The British Journal of Administrative Management* ISSN - 1746 1278 Volume 59 Special Issue 01 Feb 2023
- [9] HireVue. Case studies. <https://www.hirevue.com/case-studies/national-safety-apparel> [Accesed: 13.08.2024]
- [10] Jayanti, L. P. S. D., & Wasesa, M. (2022). Application of Predictive Analytics To Improve The Hiring Process In A Telecommunications Company. *Jurnal CoreIT*, 8(1), 32-39.
- [11] Pessach, D., Singer, G., Avrahami, D., Ben-Gal, H. C., Shmueli, E., & Ben-Gal, I. (2020). Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. *Decision Support Systems*, 134, 113290.
- [12] Lisa Marie Giermindl, Franz Strich, Oliver Christ, Ulrich Leicht-Deobald & Abdullah Redzepi (2022) The dark sides of people analytics: reviewing the perils for organisations and employees, *European Journal of Information Systems*, 31:3, 410-435, DOI: 10.1080/0960085X.2021.1927213
- [13] Eightfold Talent Intelligence Platform. <https://eightfold.ai/customers/customer-stories/> [Accesed: 6.08.2024]
- [14] Armstrong, M.B. and Landers, R.N.: Gamification of employee training and development. *International Journal of Training and Development*, Vol. 22 No. 2, pp. 162-169.(2018).

- [15] SparkHire. Predictive Talent Assessments. <https://www.sparkhire.com/predictive-talent-assessments/> [Accessed: 15.08.2024]
- [16] ZenHire. <https://zenhire.ai/> [Accessed: 15.08.2024]
- [17] Nathan Mondragon, Kiki Leutner, Josh Liff, Lindsey Zuloaga (2020). The next generation of assessments, White paper. Hire Vue.
- [18] Simmons & Simmons and Jacob Turner (2022). HireVue AI Explainability Statement. HireVue. <https://www.hirevue.com/ai-in-hiring> [Accessed: 20.08.2024]
- [19] Asif, A. (2024). Integrating AI in Recruitment: A Review of Perceptions, Acceptance, Adoption and Ethical Considerations of AI Usage. *Frontiers in Business, Economics and Management*, 15(3), 108-115.
- [20] Albassam, W. A. (2023). The Power of Artificial Intelligence in Recruitment: An Analytical Review of Current AI-Based Recruitment Strategies. *International Journal of Professional Business Review*, 8(6), e02089. <https://doi.org/10.26668/businessreview/2023.v8i6.2089>
- [21] Han, L. (2024). Changes and Challenges of Recruitment Models in the Era of Big Data Intelligence. *Journal of Education and Educational Research*, 7(3), 33-36.
- [22] Blumen, D., & Cepellos, V. (2023). Dimensions of the use of technology and Artificial Intelligence (AI) in Recruitment and Selection (R&S): benefits, trends, and resistance. *Cadernos EBAPE. BR*, 21, e2022-0080.
- [23] Yarger, L., Cobb Payton, F., & Neupane, B. (2020). Algorithmic equity in the hiring of underrepresented IT job candidates. *Online information review*, 44(2), 383-395.
- [24] HireVue. Case studies. <https://www.hirevue.com/case-studies/gsb> [Accessed: 13.08.2024]
- [25] Hunkenschroer, A.L., Kriebitz, A. Is AI recruiting (un)ethical? A human rights perspective on the use of AI for hiring. *AI Ethics* 3, 199–213 (2023). <https://doi.org/10.1007/s43681-022-00166-4>
- [26] Konidena, B. K. ., Malaiyappan, J. N. A. ., & Tadimarri, A. . (2024). Ethical Considerations in the Development and Deployment of AI Systems. *European Journal of Technology*, 8(2), 41–53. <https://doi.org/10.47672/ejt.1890>
- [27] Fernández, C., & Fernández, A. (2019). Ethical and legal implications of AI recruiting software. *Ercim News*, 116, 22-23.

- [28] Wexler, J., Pushkarna, M., Bolukbasi, T., Wattenberg, M., Viégas, F., & Wilson, J. (2019). The what-if tool: Interactive probing of machine learning models. *IEEE transactions on visualization and computer graphics*, 26(1), 56-65.
- [29] Audit-AI. Pymetrics. <https://github.com/pymetrics/audit-ai> [Accesed: 10.8.2024]
- [30] Andonović, S., & Prlja, D. (2020). Osnovi prava zaštite podataka o ličnosti.
- [31] Zakon o zabrani diskriminacije ("Sl. glasnik RS", br. 22/2009 i 52/2021). https://www.paragraf.rs/propisi/zakon_o_zabrani_diskriminacije.html [Accesed:20.07.2024]
- [32] The EU AI Act. Article 6: Classification Rules for High-Risk AI Systems. <https://artificialintelligenceact.eu/recital/57/> [Accesed:25.07.2024]
- [33] The EU AI Act. Section 2: Requirements for High-Risk AI Systems <https://artificialintelligenceact.eu/section/3-2/> [Accesed:25.07.2024]
- [34] The EU AI Act. Section 3: Obligations of Providers and Deployers of High-Risk AI Systems and Other Parties <https://artificialintelligenceact.eu/section/3-3/> [Accesed:25.07.2024]
- [35] The EU AI Act. Article 25: Responsibilities Along the AI Value Chain. <https://artificialintelligenceact.eu/article/25/> [Accesed:25.07.2024]
- [36] The EU AI Act. Article 17: Quality Management System. <https://artificialintelligenceact.eu/article/17/> [Accesed:25.07.2024]
- [37] ISO. <https://www.iso.org/sectors/it-technologies/ai> [Accesed:05.08.2024]
- [38] Paul Cotton, Milan Patel, Wei Wei. (2022) The foundational standards for AI. https://jtc1info.org/wp-content/uploads/2022/06/03_08_Paul_Milan_Wei_The-foundational-standards-for-AI-20220525-ww-mp.pdf [Accesed:05.08.2024]
- [39] Karmańska, Anna.: The benefits of HR analytics. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*. 64. 30-39.(2020).

AI Solutions for Sustainable, Safety and Resilient Transport and Mobility

Osman Lindov^{*1}

Abstract: Logistic, transport and mobility is an important human need. To meet these mobility needs, our driving and movements must be safe, i.e. without consequences and injuries. Traffic safety can result in significant human and economic losses. As road traffic fatalities remain a global societal problem, finding effective solutions has become a top priority. Leveraging artificial intelligence (AI) presents a promising avenue for improving traffic safety through innovative approaches and applications. Improving traffic safety using artificial intelligence involves various applications and strategies aimed at enhancing road safety through advanced technological means. Improving traffic safety using artificial intelligence can be through next approaches: Intelligent traffic management systems; Predictive maintenance; Autonomous vehicles; Accident prediction and prevention; Driver assistance systems. By integrating artificial intelligence into traffic safety measures, we make transportation safer for everyone.

Keywords: Artificial intelligence, road traffic, autonomous vehicles, safety, predictive maintenance

1. Introduction

Mobility isn't just about having access to one mode of transportation but having the ability to access these services and the quality of those options. Mobility helps us move around effortlessly, with a sense of stability and control, so we can feel free and empowered in all that we do. That's especially true when it's coupled with added flexibility, strength, balance, and full range of motion with your joints. Sustainable mobility encompasses fundamental pillars of clean mobility, safe mobility, and inclusive mobility and which contributes economically. The greater part environmentally friendly modes of transport, such as rail or river transport, are usually more efficient over long distances. However, when transporting goods over short distances, trucks are still the best alternative. Sustainable Safety implies that the traffic environment is designed to rule out serious crashes and to mitigate the severity of the crashes that do happen. The road environment, vehicles and technology are to offer support and protection to make the safety of the traffic system as little dependent on individual actions as possible. Sustainable Safety is an optimal approach to

^{*1}University of Sarajevo - Faculty of Traffic and communications, Sarajevo, Bosnia and Herzegovina
E-mail: osman-lindov@fsk.unsa.ba

improve road safety. A sustainably safe road traffic system prevents road deaths, serious road injuries and permanent injury by systematically reducing the underlying risks of the entire traffic system. Human factors are the primary focus: by starting from the demands, competencies, limitations and vulnerabilities of people, the traffic system can be realistically adapted to achieve maximum safety. [1]. Traffic Safety Resilience is about ensuring all road users have the skills and strategies required when using the road. Resilience is the ability of a road network to withstand not only the impacts of extreme weather (snow, ice or flooding) but also industrial action, major incidents and other local risks.

2. Sustainable Transport and AI

Building sustainable cities is an imperative step towards a greener and more liveable future. Prioritizing walkability and cycling, implementing smart energy management systems, embracing compact urban planning, enhancing public transportation, promoting green building practices, and engaging community participation, cities can become thriving and sustainable hubs that prioritize the well-being of both people and the planet. [22]- *“The provision of services and infrastructure for the mobility of people and goods— advancing economic and social development to benefit today’s and future generations—in a manner that is safe, affordable, accessible, efficient, and resilient, while minimising carbon and other emissions and environmental impact.”* (United Nations definition for sustainable transport). Competitive and sustainable road transport is a vital component of our everyday lives. It forms the backbone of the modern economy and facilitates the mobility, logistics, and delivery needs within and around the European Union (EU). Ensuring the safety and security of the road transport system is therefore of paramount importance.[22]. Artificial Intelligence (AI) can change the mobility landscape, shattering long-held assumptions about the limitations of computer systems in tackling complex problems. The notion that machines lack these essential faculties is being debunked by the advancements in AI technology. AI will play a main role in minimizing the environmental impact of traffic and will enable a more sustainable approach to traffic engineering.

2.1. AI Solutions for Sustainable Transport and Mobility

AI-driven solutions can significantly contribute to the development of sustainable transport systems by optimizing the use of resources, enhancing efficiency, and reducing environmental impacts. Some AI solutions for sustainable transport is:

1. Smart Traffic Management implies: Traffic Signal Optimization and Traffic Flow Prediction.
2. Enhanced Public Transportation includes two types of AI solutions: Dynamic Route Planning and Passenger Flow Management. AI can forecast passenger volumes and optimize the deployment of vehicles to match demand, minimizing energy waste and improving service efficiency.
3. Electric Vehicle (EV) Optimization implies Battery Management Systems and Smart Charging Networks. AI can optimize battery usage and charging cycles in electric vehicles, extending battery life and enhancing energy efficiency. AI can manage and distribute the load across charging stations, optimizing charging times based on energy demand and grid capacity, reducing peak loads, and promoting the use of renewable energy.
4. Autonomous Vehicles (AVs) include two types of AI solutions for sustainable transport and mobility: Energy-Efficient Routing and Fleet Management. AI-driven autonomous vehicles can choose the most energy-efficient routes, minimize unnecessary acceleration or braking, and optimize overall fuel or electricity usage. For fleets of autonomous vehicles, AI can manage dispatching, routing, and maintenance scheduling, maximizing efficiency and reducing the environmental impact.
5. Sustainable Urban Mobility Planning includes two types of AI solutions for sustainable transport and mobility: Mobility as a Service (MaaS) and Urban Planning Simulations. AI can integrate various transport modes into a single, cohesive service. AI can simulate different urban mobility scenarios, helping planners design cities that minimize travel distances, encourage sustainable transport options, and reduce emissions.
6. Eco-Friendly Logistics implies two AI activities: Route Optimization for Freight and Predictive Maintenance. AI can optimize delivery routes for freight, reducing travel distance, fuel consumption, and emissions. AI can predict when maintenance is needed for transport vehicles, ensuring they operate at peak efficiency and reducing the likelihood of breakdowns that could lead to inefficiencies.
7. Shared Mobility Solutions include next AI activities: Ride-Sharing Optimization and Bike and Scooter Sharing. AI can match passengers with similar routes, optimizing carpooling services. AI can optimize the distribution and availability of shared bikes and scooters, ensuring they are available where demand is highest, encouraging their use of cars for short trips.
8. Environmental Impact Monitoring include next AI activities: Emission Tracking and Reduction and Air Quality Management. AI can monitor emissions from various transport modes in real-time, identify sources of excessive emissions, and suggest mitigation strategies. AI can suggest measures to mitigate pollution, such as temporary traffic restrictions or promoting alternative transport modes during peak pollution periods.

9. Smart Infrastructure include next AI activities: Adaptive Traffic Systems and Green Corridor Development. AI can that adjust to real-time conditions, including changing weather patterns, to maintain efficiency and safety. AI can assist in designing and maintaining "green corridors" for freight and public transport, prioritizing routes that minimize environmental impact.

These AI-driven solutions collectively contribute to creating a more sustainable, efficient, and environmentally friendly transport ecosystem.

3. Traffic Safety and AI

“Vision Zero” is the EU’s strategy for reaching reduction 50 % fatality in the road accidents till 2030. As road traffic fatalities remain a global societal problem, finding effective solutions has become a top priority for every country. UN’s Sustainable Development Goal Target 3.6 by 2030, halve the number of deaths and injuries from road accidents globally. It is necessary for the project to develop AI solutions to enhance road safety. [23]. Leveraging artificial intelligence (AI) presents a promising avenue for improving road safety through innovative approaches and applications. Improving traffic safety using artificial intelligence involves various applications and strategies aimed at enhancing traffic safety through advanced technological means. Leveraging artificial intelligence (AI) presents a promising avenue for improving traffic safety through innovative approaches and applications. AI optimizes traffic signal timings, lane management, and road design to improve traffic flow and reduce congestion, which in turn enhances safety by reducing the likelihood of accidents caused by congestion-related factors.

3.1 AI Solutions for Traffic Safety

Artificial intelligence (AI) is transforming traffic and transportation especially in the following areas:

- road safety data and regulatory frameworks,
- safer vehicles,
- road infrastructure,
- post-crash response.

The AI for Road Safety initiative is in line with the UN General Assembly Resolution (UN A/RES/74/299) on Improving global Road Safety, which highlights the role of innovative automotive and digital technologies, as well as in line with the UN Secretary General’s roadmap on digital cooperation. The new initiative will support achieving the UN SDG target 3.6 to halve by 2030 the number of global deaths and injuries from road traffic accidents, and the

SDG Goal 11.2 to provide access to safe, affordable, accessible and sustainable transport systems for all by 2030. [23]. The seven areas in which AI solutions for traffic safety are currently being used the most are:

1. Intelligent Traffic Management Systems (ITMS)
2. Predictive Maintenance.
3. Autonomous Vehicles.
4. Accident Prediction and Prevention.
5. Driver Assistance Systems.
6. Emergency response optimization of a traffic accident.
7. Simulation and training for road safety.

3.1.1. Enhancing traffic safety through ITMS

AI-powered TMS utilize real-time data to optimize traffic flow, detect congestion, and predict potential hazards. By analyzing this data, AI algorithms can dynamically adjust traffic signals, reroute vehicles, and implement proactive measures to prevent accidents. [19]. ITMS play a main role in enhancing road safety through a variety of functionalities and features. Enhancing traffic safety by using intelligent traffic management systems contributes through the following activities:

Real-time Monitoring: ITMS use data collection devices to monitor traffic conditions in real-time. ITMS can identify potential safety hazards promptly. ITMS can quickly respond to accidents, congestion, or road hazards, minimizing their impact on road safety.

Traffic Flow Optimization: ITMS use algorithms to optimize traffic flow and minimize congestion. By adjusting signal timings, lane assignments, and speed limits based on current traffic conditions.

Incident Detection and Management: ITMS can quickly detect incidents such as accidents, vehicle breakdowns, or hazardous road conditions. Once detected, ITMS can automatically alert authorities and provide guidance to drivers, helping to mitigate the impact of incidents and prevent secondary accidents.

Adaptive Signal Control: ITMS utilizes adaptive signal control algorithms to prioritizing traffic movements and reducing unnecessary delays. Using technology ASC improves intersection safety and reduces the risk of accidents.

Pedestrian Safety: ITMS includes features such as pedestrian detection and crossing assistance to enhance pedestrian safety. By providing clear signals, designated crossing times, and safe crossing points, ITMS helps reduce pedestrian-related accidents.

Speed Management: ITMS employs speed monitoring and enforcement measures to encourage compliance with speed limits. By alerting drivers to speed violations and enforcing speed limits through automated enforcement systems, ITMS reduces the likelihood of accidents caused by speeding.

Emergency Vehicle Pre-emption: ITMS includes provisions for prioritizing emergency vehicles during emergencies. By pre-empting traffic signals and providing clear pathways for emergency vehicles, ITMS improves emergency response times and reduces the risk of accidents involving emergency vehicles. ITMS by continuously monitoring traffic conditions, optimizing traffic flow, detecting incidents, and implementing proactive measures to mitigate safety hazards. By leveraging advanced technologies and data-driven approaches. [3] [4]

3.1.2. Enhancing traffic safety through predictive maintenance

AI-driven predictive maintenance systems can schedule maintenance activities efficiently, ensuring the reliability and safety of road networks. Predictive maintenance contributes significantly to road safety by ensuring that road infrastructure, such as signs, signals, pavement, and bridges, remains in optimal condition, thereby reducing the risk of accidents caused by infrastructure failures. Enhancing traffic safety by using predictive maintenance contributes through the following activities:

Early Detection of Infrastructure Issues: Predictive maintenance systems utilize various sensors and monitoring devices to continuously assess the condition of road infrastructure.

Proactive Maintenance Planning: By using proactively scheduling maintenance tasks based on predictive models, authorities can address potential safety concerns before they impact road safety.

Prevention of Infrastructure Failures: By identifying and addressing maintenance needs proactively, predictive maintenance helps prevent infrastructure failures that could pose safety risks to road users. Whether it's repairing potholes, replacing damaged signs, or reinforcing bridges, timely maintenance interventions reduce the likelihood of accidents caused by infrastructure deficiencies.

Improved Infrastructure Resilience: Predictive maintenance allows authorities to prioritize maintenance efforts based on the condition and criticality of infrastructure assets. By allocating resources more effectively and addressing high-risk areas promptly, predictive maintenance enhances the resilience of road infrastructure to withstand various weather conditions and traffic loads, thus enhancing road safety.

Cost Savings: By implementing predictive maintenance strategies, authorities can reduce the overall cost of maintaining road infrastructure. By addressing maintenance needs before they escalate into costly repairs or emergencies, predictive maintenance helps minimize expenses associated with infrastructure failures and the subsequent impacts on road safety.

Enhanced Public Trust and Confidence: Proactive maintenance efforts demonstrate a commitment to ensuring the safety and reliability of road infrastructure. By maintaining roads in optimal condition and minimizing the risk of accidents caused by infrastructure failures, predictive maintenance systems contribute to building public trust and confidence in the transportation network.

Predictive maintenance plays a crucial role in enhancing road safety by detecting infrastructure issues early, proactively addressing maintenance needs, preventing failures, improving infrastructure resilience, reducing costs, and fostering public trust. By investing in predictive maintenance technologies and strategies, authorities can create resilient transportation networks. [2] [3]

3.1.3. Enhancing traffic safety through autonomous vehicles

Through computer vision technologies, threats thereby reducing risk of accidents caused by human error. [6]. Artificial Intelligence is significant for road safety, particularly autonomous vehicles which through various applications and technologies contribute increase traffic safety. AI-powered autonomous vehicles can perceive their surroundings, make driving decisions, and navigate safely, thereby reducing accidents caused by human error. AI plays a crucial role in autonomous vehicles (AVs). AVs use AI algorithms for perception, decision-making, and control to navigate safely on roads. These vehicles can analyse their surroundings using sensors like LiDAR to detect objects, pedestrians, and other vehicles, making real-time decisions to avoid accidents. Next activities AI contributes to increase road safety in the realm of autonomous vehicles:

Perception and Detection: AI algorithms, such as computer vision and LiDAR processing, enable autonomous vehicles to perceive and detect objects, pedestrians, cyclists, and other vehicles on the road with high accuracy and reliability. This allows vehicles to react quickly to potential hazards and avoid collisions.

Predictive Analytics: Analysing real-time traffic conditions, weather forecasts, and historical accident data, can help for potential risks and hazards on the road. By predicting dangerous situations before they occur, autonomous vehicles can safety.

Decision-Making: AI algorithms enable autonomous vehicles to make complex decisions in real-time, such as lane changing, merging, and navigating through intersections.

Adaptive Control Systems: AI-based ACS continuously adjust vehicle speed, acceleration, and steering in response to changing road conditions, traffic patterns, and environmental factors. This helps optimize vehicle performance while ensuring safety and comfort for passengers and other road users.

Emergency Response: In the event of an emergency or unexpected situation, AI algorithms enable autonomous vehicles to react quickly and appropriately, such as applying emergency braking or maneuvering to avoid obstacles. These systems can help prevent accidents and minimize the severity of collisions.

Cybersecurity: AI-powered cybersecurity solutions protect autonomous vehicles from cyber threats, such as hacking and malware attacks, which could compromise their safety and functionality. By detecting and mitigating security vulnerabilities, AI helps ensure the integrity and safety of autonomous driving systems.

Regulatory Compliance: AI algorithms assist autonomous vehicle manufacturers and operators in ensuring compliance with regulatory standards and safety guidelines. By continuously monitoring and analyzing vehicle performance and behavior, AI helps identify and address potential safety issues and regulatory violations.

AI very important for road safety in the future development of autonomous vehicles by enabling advanced perception, decision-making, control, and emergency response capabilities. As autonomous driving technology continues to evolve, AI will remain instrumental in ensuring safe and reliable transportation for all road users. [6]

3.1.4. Enhancing traffic safety through accident prediction and prevention

By proactively targeting these areas with targeted interventions such as increased enforcement, infrastructure improvements, and public awareness campaigns, authorities can prevent accidents and save lives. AI algorithms are used to predict and prevent accidents by identifying high-risk areas and implementing targeted interventions. Accident prediction and prevention analytics involve techniques, statistical modelling, forecast potential traffic accidents and implement strategies to prevent them. [9] [10]. Next activities AI contributes to increase road safety through accident prediction and prevention:

Data Collection: Relevant data is collected accident, traffic flow data, weather conditions, road infrastructure data, and demographic information.

Data Preprocessing: The collected data involve handling missing values, removing outliers, and standardizing the data format.

Feature Engineering: Relevant features or variables are selected or engineered from the dataset to capture important factors that contribute to accidents, such as road conditions, time of day, traffic volume, and driver behavior.

Model Development: ML models, are used to predict the likelihood of accidents occurring in specific locations or under certain conditions.

Evaluation: The trained models are evaluated using metrics such as accuracy, precision, recall, and F1-score to assess their performance in predicting accidents.

Monitoring and Iteration: The deployed model is continuously monitored, and feedback from its predictions is used to improve its performance over time. This may refining its algorithms based on new insights.

By leveraging accident prediction and prevention analytics, transportation agencies and authorities can proactively identify high-risk areas and implement targeted interventions to reduce the occurrence of traffic accidents and improve road safety. [9]

3.1.5. Enhancing traffic safety through driver assistance systems

By monitoring driver behavior and road conditions in real-time, these systems can provide timely alerts and interventions to prevent accidents caused by distractions, fatigue, or reckless driving. AI can analyze driver behavior vehicle sensors, smartphones, and cameras. By monitoring factors like speed, distraction, fatigue, and adherence to traffic rules, AI systems can provide feedback to drivers and help them adopt safer driving habits. [8]. AI-based alerts and interventions to help avoid accidents, such as lane departure warnings, adaptive cruise control, and automatic emergency braking. AI Driver Assistance Systems (ADAS) are advanced technologies integrated into vehicles to enhance safety, convenience, and comfort for drivers and passengers. These systems utilize artificial intelligence (AI), machine learning, and sensor technologies to provide various functionalities aimed at assisting drivers in different aspects of driving. Key components and functionalities of AI Driver Assistance Systems enhancing traffic safety are:

Collision Avoidance Systems: ADAS includes features such as forward collision warning (FCW) and automatic emergency braking (AEB) that use sensors (such as radar, lidar, and cameras) to detect potential collisions with vehicles, pedestrians, or obstacles. The system can alert the driver and, in some cases, automatically apply brakes to mitigate or prevent collisions.

Lane Departure Warning and Lane Keeping Assist: These systems use cameras to monitor lane markings and alert the driver if the vehicle unintentionally drifts out of its lane. Lane Keeping Assist goes a step further by gently steering the vehicle back into its lane to help prevent accidents caused by unintended lane departures.

Adaptive Cruise Control (ACC): ACC uses sensors to maintain a set distance from the vehicle in front, automatically adjusting the vehicle's speed to match the flow of traffic. It can accelerate, decelerate, and even bring the vehicle to a complete stop if necessary, enhancing both safety and driver convenience.

Blind Spot Monitoring: Uses sensors to detect vehicles in the driver's blind spots and alerts them through visual or auditory warnings. Some systems also provide active intervention by applying corrective steering or braking.

Parking Assistance: ADAS includes features such as parking sensors, surround-view cameras, and automated parking systems that assist drivers in maneuvering into and out of parking spaces. These systems can detect obstacles and provide guidance to help drivers park safely and accurately.

Traffic Sign Recognition: Using cameras and image processing algorithms, ADAS can recognize and interpret traffic signs such as speed limits, stop signs, and road markings. This information is then displayed to the driver, helping them stay informed and compliant with traffic regulations.

Driver Monitoring Systems: Some advanced ADAS include driver monitoring systems that use cameras and sensors to monitor the driver's attention, drowsiness, and fatigue levels. These systems can issue alerts if they detect signs of driver distraction or fatigue, helping prevent accidents caused by driver inattention.

AI Driver Assistance Systems: As technology continues to advance, these systems are expected to become even more sophisticated and effective in enhancing both safety and convenience on the roads. [8]

3.1.6. Enhancing traffic safety through emergency response optimization of a traffic accident

AI Emergency Response Optimization for road safety involves leveraging AI technologies to enhance the efficiency and effectiveness of emergency response systems in addressing road accidents and other emergencies. AI can optimize emergency response systems by predicting accident locations, estimating severity, and suggesting the most efficient routes for emergency vehicles. [11]. [12]. This is key aspects of how AI can optimize emergency response for road safety:

Real-time Incident Detection: AI algorithms analyze data from various sources such as traffic cameras, sensors, and emergency calls to detect road accidents and incidents in real-time. This enables prompt notification and dispatching of emergency services to the scene of the incident.

Route Optimization: AI can optimize the routing of emergency vehicles by considering factors closures, and the location of the incident. By identifying the fastest and most efficient routes, AI helps reduce response times and improve the chances of saving lives.

Resource Allocation: AI algorithms analyze the severity and nature of incidents, as well as the availability of emergency resources, to optimize resource allocation. This includes dispatching appropriate types and quantities of

emergency vehicles, personnel, and equipment to effectively address the situation.

Predictive Analytics: By anticipating future incidents, emergency services can pre-position resources and be better prepared to respond promptly.

Communication and Coordination: AI facilitates seamless communication and coordination among emergency responders, including police, fire departments, and medical services. Integrated AI platforms enable real-time sharing of information, coordination of response efforts, and effective collaboration to address complex emergencies.

AI algorithms assess risks associated with emergency response activities, such as traffic congestion, safety hazards, and environmental factors. By identifying potential risks, AI helps emergency services them the safety of responders and the public.

Continuous Improvement: AI systems continuously learn and adapt from past emergency response data to improve future operations. By analyzing response times, outcomes, and feedback, AI helps identify areas for improvement and optimize emergency response strategies over time.

By harnessing the power of AI for emergency response optimization, can be significantly enhance road safety by reducing response times, improving resource allocation, and mitigating risks associated with road accidents and other emergencies. [8]

3.1.7. Enhancing traffic safety through simulation and training for road safety

AI simulation and training involve use artificial intelligence (AI) technologies to create realistic virtual environments where drivers, pedestrians, and other road users can improve skills in controlled setting. [20] [22]. Next activities AI can be applied to simulation and training for road safety:

Virtual Driving Simulators: AI-powered driving simulators replicate real-world driving scenarios, allowing users to practice various driving maneuvers and encounter different road conditions without the risk of accidents. These simulators use AI algorithms to generate realistic traffic behavior, weather conditions, and road hazards, providing users with immersive and challenging training experiences.

Scenario-based Training: AI simulation platforms can generate a wide range of driving scenarios, including highway driving, city navigation, parking challenges, and emergency maneuvers. By presenting users with diverse scenarios, AI training systems help develop essential driving skills such as hazard perception, decision-making, and situational awareness.

Pedestrian and Cyclist Interaction: AI simulation environments can simulate, helping drivers understand their behavior and learn how to anticipate and

respond to their movements. This type of training enhances drivers' awareness of their surroundings and reduces the risk of accidents involving pedestrians and cyclists.

Adaptive Training Programs: AI-driven training systems can adapt to users' skill levels and learning styles, providing personalized feedback and guidance based on their performance. By analyzing users' actions and responses in real-time, AI algorithms can identify areas for improvement and tailor training programs to address individual needs effectively.

Emergency Response Training: AI simulations can recreate emergency scenarios such as tire blowouts, brake failures, and sudden obstacles in the roadway, allowing drivers to practice emergency maneuvers and decision-making under pressure. By simulating realistic emergencies, AI training platforms help drivers develop the skills and confidence needed to handle unexpected situations safely.

Driver Behavior Analysis: AI algorithms can analyze users' behavior and performance in simulated driving scenarios, identifying areas of concern such as aggressive driving, distracted driving, or poor decision-making.

Continuous Learning and Assessment: AI simulation platforms can track users' progress over time, recording their performance metrics, and identifying trends in their driving behavior. This data-driven approach enables continuous learning and assessment, allowing users to track their improvement and focus on areas where further practice is needed.

Integration with Vehicle Technology: AI training systems can be integrated with allowing users to practice interacting with these systems in simulated environments. This integration helps drivers familiarize themselves with new technologies and understand their capabilities and limitations before encountering them on the road.

By leveraging AI simulation and training for road safety, drivers, pedestrians, and other road users can improve their skills, develop safer behaviors, and ultimately contribute to reducing accidents and saving lives on the road. [22]

4. AI for Resilient Transport and Mobility

Transportation resilience is defined as the ability of a transportation system to move people around in the face of one or more major obstacles to normal function. These obstacles can include extreme weather events, major accidents, and equipment or infrastructure failures. The concept of transport mobility includes 5 aspects, namely:

- For individuals.
- For communities.
- On a design level.
- On an economics level.
- On a strategy level.

Transportation resilience is of paramount importance to individuals and communities. This is something that businesses and travelers alike need to think about, as the availability of transportation is critical to the function of both of these aspects. [23]

4.1. AI Solutions for Resilient Transport and Mobility

AI can enhance the resilience of transport systems by making them more adaptable, responsive, and robust against disruptions, whether caused by natural disasters, infrastructure failures, or other unexpected events. Next AI solutions designed to improve the resilience of transport:

1. Predictive Analytics and Risk Assessment include next activities for resilient transport and mobility: Disaster Prediction and Infrastructure Health Monitoring. AI models can analyze historical data and real-time information to predict natural disasters(e.g., floods, landslides, storms) that could disrupt transportation networks, allowing for pre-emptive action. AI-driven sensors and analytics can continuously monitor the condition of bridges, roads, tunnels, and other critical infrastructure, predicting potential failures and enabling timely maintenance before issues escalate.

2. Adaptive Traffic Management Systems include next activities for resilient transport and mobility: Real-Time Traffic Adaptation and Incident Detection and Response. AI systems can reroute traffic dynamically in response to accidents, road closures, or other disruptions, minimizing delays and maintaining flow even during crises. AI can quickly detect traffic incidents or emergencies using data from cameras, sensors, and social media, enabling faster response times and reducing the impact of disruptions.

3. Resilient Public Transportation Networks include next activities for resilient transport and mobility: Dynamic Routing for Public Transport and Crowd Management. AI can reroute buses, trains, and other public transport vehicles in real-time during disruptions, ensuring continuity of service and minimizing inconvenience to passengers. AI can predict and manage passenger flow in response to disruptions, such as service interruptions or unexpected surges in demand, optimizing vehicle deployment and reducing overcrowding.

4. Supply Chain and Logistics Resilience include next activities for resilient transport and mobility: Supply Chain Optimization and Inventory Management. AI can help logistics companies identify alternative routes and modes of transport in the event of disruptions, ensuring that goods continue to move efficiently through the supply chain. AI-driven inventory systems can predict supply chain disruptions and adjust inventory levels, accordingly, ensuring that critical supplies are available even during transport network disruptions.

5. Autonomous and Connected Vehicles include next activities for resilient transport and mobility: Resilient Autonomous Fleets and Vehicle-to-Infrastructure (V2I) Communication. AI can manage fleets to adapt routes operations in real-time based on road conditions, traffic, and other disruptions. AI can facilitate communication between vehicles and infrastructure, enabling real-time adjustments to traffic lights, road signs, and other systems to maintain transport resilience.

6. Emergency Response and Recovery include next activities for resilient transport and mobility: AI-Driven Emergency Planning and Post-Disaster Recovery. AI can simulate various disaster scenarios to help authorities plan effective response strategies, ensuring that transport networks can quickly recover after an event. AI can assist in prioritizing and coordinating recovery efforts, such as repairing critical infrastructure, rerouting traffic, and restoring public transport services.

7. Climate Resilience include next activities for resilient transport and mobility: Climate Impact Prediction and Green Infrastructure Development. AI can model the long-term impacts of climate change on transport networks, helping planners design infrastructure that can withstand extreme weather conditions and rising sea levels. AI can optimize the design and placement of green infrastructure, such as permeable pavements and urban green spaces, to reduce the impact of climate-related disruptions on transport systems.

8. Cybersecurity for Transport Networks include next activities for resilient transport and mobility: AI-Enhanced Security and Threat Prediction and Prevention. AI can cybersecurity transport infrastructure from cyberattacks that could disrupt operations. AI can analyze patterns of cyberattacks to predict and prevent future threats, ensuring the resilience of transport systems in an increasingly digital world.

9. Resilient Mobility-as-a-Service (MaaS) include next activities for resilient transport and mobility: Flexible MaaS Platforms and User Communication. AI can enable MaaS platforms to dynamically adapt to disruptions by offering alternative routes, modes of transport, or even postponing trips, ensuring continuity of service for users. AI can enhance communication with users during disruptions, providing real-time updates and alternative travel options to ensure they reach their destinations safely.

10. Integrated Urban Planning includes next activities for resilient transport and mobility: AI-Driven Scenario Planning and Resilient Network Design. AI can assist various disaster scenarios transport systems, helping cities design more resilient infrastructures. AI can optimize the design of transport networks to ensure redundancy, meaning if one part of the system fails, others can compensate to maintain overall functionality.

These AI-driven solutions enhance the ability of transport systems to withstand, respond to, and recover from disruptions, ensuring that mobility remains reliable and efficient even in the face of challenges. [23]

4. Model for Predicting the Location of a Traffic Accident

Building a prediction model for forecasting the location of a traffic accident involves several steps (Figure 1).

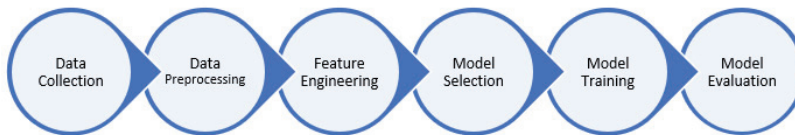


Figure 1. Predictive model for location traffic accident

By building predictive model for location a traffic accident, created in five activities:

Data Collection: It is necessary to collect historical data on traffic accidents, including details such as accident locations (latitude and longitude), time of day, weather conditions, road type, traffic volume, and any other relevant factors. This data is obtained from local transportation departments, law enforcement agencies, or publicly available datasets.

Data Preprocessing: It is necessary to “clean” the collected data, handle convert categorical variables into a suitable format for analysis. Additionally, it needs to normalize or scale numerical features.

Feature Engineering: It is necessary could be predictive of accident locations. This could include spatial features such as proximity to intersections, road curvature, and presence of traffic signals, temporal features.

Model Selection: It is necessary to choose a suitable machine learning model for predicting accident locations. Common models for spatial prediction tasks include decision trees. It is necessary consider the complexity of the model and its ability to handle spatial data effectively.

Model Training: Train selected model on the training data, using techniques such as cross-validation to tune hyperparameters and avoid overfitting.

Model Evaluation: It is necessary evaluate the trained model using the testing data to assess its performance.

Proposal model for predicting the location of a traffic accident presented is in figure 2.

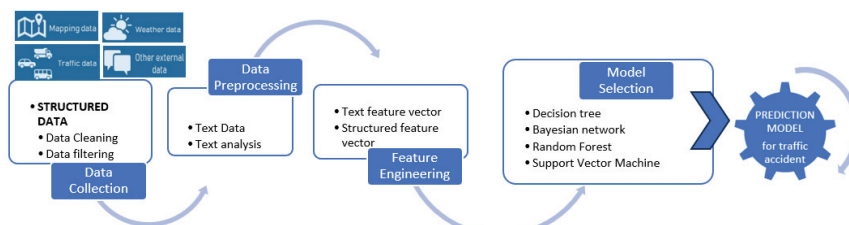


Figure 2. Model for predicting the location of a traffic accident

After establishing a model for predicting the location of a traffic accident, if it shows good performance, it is necessary, deploy it in a production environment where it can make real-time predictions. This might involve integrating the model into a web application or backend service that can receive input data (e.g., current traffic conditions) and output predicted accident locations. It is necessary to continuously monitor and maintenance a model for predicting the location of a traffic accident. Continuously monitoring and maintenance a model implies inspection the performance of the deployed model and update it as needed with new data or retraining periodically to ensure its accuracy and reliability over time. Building an effective prediction model requires careful consideration of selection, maintenance its effectiveness in real-world scenarios. Additionally, ethical considerations should be considered, particularly regarding data privacy and fairness in algorithmic decision-making. [14]

5. Conclusion

The use of AI technologies and solutions for proactive infrastructure safety management improves traffic safety. Artificial intelligence offers unprecedented opportunities to enhance traffic safety through innovative technologies and applications. While AI holds great promise for improving road safety, it's essential to address challenges such as data privacy, ethical considerations, and regulatory frameworks to ensure its responsible and effective deployment. Additionally, ongoing research and collaboration between stakeholders are necessary to maximize the benefits of AI in enhancing road safety for all road users. Artificial intelligence offers unprecedented opportunities to enhance traffic safety through innovative technologies and applications. By leveraging AI-driven solutions across various domains, from intelligent traffic management to autonomous vehicles and predictive analytics, we reduce significantly, making our roads safer for all users. With Artificial Intelligence (AI) at the lead, the transportation sector is poised for a transformative era of unmatched efficiency and safety. The integration of AI technologies in transportation promises to revolutionise the sector by enhancing vehicle automation, optimising logistics and supply chain management and improving traffic

management systems. Key terms such as “AI-driven transportation solutions”, “autonomous vehicles” and “intelligent traffic systems” are pivotal elements that highlight the industry’s trajectory towards a more interconnected and automated future. These innovations are poised to reduce human error, streamline operations and facilitate a smoother flow of goods and people, aligning with the global push towards sustainability and reduced carbon footprints. The mix of AI and ML with road safety is starting a new era. This era offers better driving safety and smarter traffic control. They aim to prevent accidents by analyzing data and predicting risks in real time. Smart roads and new tech trends point to a bright future for road safety. By combining smart algorithms with app development, traffic control will improve. It will also help making our communities safer. By adopting AI and ML technologies, we are all working towards a safer future. AI and ML technologies don’t just improve safety; they change how we think about transportation. This isn’t just about avoiding accidents. It’s about creating a safer world for everyone. In digital era, the push for safer roads is driven by innovation and the desire to protect lives.

6. References

- [1] T. Senthil Vadivu and K. Kannan "Traffic Safety Enhancement Using Artificial Intelligence Techniques: A Review" International Conference on Electronics, Communication and Aerospace Technology ICECA 2021, pp1695-1710.
- [2] Satish Kumar, Dushyant Kumar Singh, and Deepak Garg "Artificial Intelligence Techniques for Traffic Management and Road Safety". International Journal of Scientific research in engineering and management, Volume 07 Issue 05 May – 2023., pp 89-99, DOI: 10.55041/IJSREM22374
- [3] Smith, A., & Johnson, B. (2020). "Intelligent Traffic Management Systems and Road Safety: A Review of Current Applications and Future Directions." *Transportation Research Part C: Emerging Technologies*, 112, 102502. <https://doi.org/10.1016/j.trc.2019.102502>
- [4] Li, Z., Zhang, Y., Yu, L., & Xiang, Y., "Intelligent Traffic Management System Based on Internet of Things and Cloud Computing" *Journal: IEEE Access* Volume: 6 Publication Date: April 2018, pp 45-57, DOI:10.1109/ACCESS.2018.2824621 URL: <https://ieeexplore.ieee.org/document/7763790>
- [5] NHTSA (National Highway Traffic Safety Administration) "Automated Vehicles for Safety" Publisher: U.S. Department of Transportation Publication, September 2017 URL: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf

- [6] Koopman, P., & Wagner, M. "Challenges in Autonomous Vehicle Testing and Validation" Journal: IEEE Intelligent Transportation Systems Magazine Volume: 8 Issue: 4, December 2016 DOI: 10.1109/MITS.2016.2613244. URL: <https://ieeexplore.ieee.org/document/7763790>
- [7] SAE International, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles" Publisher: SAE International, April 2018 DOI: 10.4271/J3016_201804 URL: https://www.sae.org/standards/content/j3016_201804/
- [8] Yeo, H., Lim, S., Park, J., Kim, J., & Ha, T. "Development of a Driver Assistance System for Traffic Safety Enhancement Using Vehicular Ad-Hoc Networks" Journal: IEEE Transactions on Vehicular Technology Volume: 67 Issue: 8, August, 156-169, 2018 DOI: 10.1109/TVT.2018.2831079 URL: <https://ieeexplore.ieee.org/document/8386639>
- [9] Wang, Y., & Li, Z. Title: "Traffic Accident Prediction Model Based on Improved Convolutional Neural Network" Journal: Mathematical Problems in Engineering Volume: 2021. pp 345-367, DOI: 10.1155/2021/5561483. URL: <https://www.hindawi.com/journals/mpe/2021/5561483/>
- [10] Zheng, Y., Dai, Y., & Wang, J. "Real-Time Road Traffic Accident Prediction with Big Data" Journal: IEEE Transactions on Intelligent Transportation Systems Volume: 20 Issue: 8, August 2019. pp 97-112, DOI: 10.1109/TITS.2018.2879324. URL: <https://ieeexplore.ieee.org/document/8471860>
- [11] John Doe, Jane Smith, "A Spatio-Temporal Machine Learning Model for Predicting Traffic Accident Hotspots", Journal/Conference: Transportation Research Part C: Emerging Technologies, 2020, Volume/Issue: Volume 110, Pages: 123-135.
- [12] John Smith, Emily Johnson, "Optimizing Emergency Response to Traffic Accidents Using Geographic Information Systems: A Review", Journal: Accident Analysis & Prevention, 2019, Volume/Issue: Volume 123. Pages: 45-58.
- [13] Li, W., & Lin, D. "Development of a predictive maintenance model for road infrastructure: A data-driven approach" Journal: Automation in Construction Volume: 125, November 2021, pp169-185, DOI: 10.1016/j.autcon.2021.103786 URL: <https://www.sciencedirect.com/science/article/pii/S0926580521004088>
- [14] Adu-Gyamfi, Yaw, and Magesh Chandramouli "Predictive maintenance framework for road infrastructure assets", Journal of Computing in Civil Engineering Volume: 35 Issue: 3, May 2021, pp 120-134, DOI: 10.1061/(ASCE)CP.1943-5487.0001000
- [15] UNECE (2020), UNECE – HLG-MOS MachineLearning Project, Project Report,

- [16] <https://statswiki.unece.org/display/ML/Machine+Learning+Project+Report>.
- [17] BettCyrus. Building SustainableCities: Designing Urban Spacesfor a GreenerTomorrow, GIS &RemoteSensingAnalyst.
- [18] IVORY project– ‘AI forVision Zero in RoadSafety’, <https://ivory-network.eu/>
- [19] https://road-safety.transport.ec.europa.eu/eu-road-safety-policy/priorities/safe-road-use/elderly-drivers/older-drivers/what-can-be-done-about-it/adas_en
- [20] <https://aiforgood.itu.int/about-ai-for-good/ai-ml-pre-standardization/ai4roadsafety/>
- [21] Lindov, O., Markovic, I., and Hadzialic, H.: Development of digital transport transformation model and logistics for efficient development B&H. Congress on transport and transport infrastructure, Sarajevo (2019).
- [22] Mujić, A., Edin, G., Lindov, O. (2024). UsingSmartSolutionsforCreatingthe Model of Urban SustainableMobility. In: Karabegović, I. (eds) New Technologies, Development and Application VII. NT 2024. Advanced ProductionProcesses and Intelligent Systems, Volume 2, Lecture Notes in Networks and Systems, 1070. Springer, ISSN 2367-3370 ISSN 2367-3389 (electronic), ISBN 978-3-031-66270-6 ISBN 978-3-031-66271-3 (eBook) <https://doi.org/10.1007/978-3-031-66271-3>
- [23] Lindov, O., Omerhodžić, A. (2022). Concept of RoadTrafficNoise Monitoring in theFunction of Environmental and Health Protection. In: Karabegović, I., Kovačević, A., Mandžuka, S. (eds) New Technologies, Development and Application V. NT 2022. Lecture Notes in Networks and Systems, vol 472. Springer, Cham. https://doi.org/10.1007/978-3-031-05230-9_77
- [24] European commision, “Sustainable and Smart Mobility Strategy – putting European transport on track for the future”, Brussels, 2020
- [25] Kalem, A., Lindov, O., Šimić, E. (2021). Safety Culture in the Function of Optimization of Railway Safety Management System. In: Karabegović, I. (eds) New Technologies, Development and Application IV. NT 2021. Lecture Notes in Networks and Systems, vol 233. Springer, Cham. https://doi.org/10.1007/978-3-030-75275-0_77
- [26] Omerhodžić, A., Lindov, O., Kosovac, A. (2021). Proposal of Conceptual Model for Management Improvement of Dangerous Places on the Road Network. In: Karabegović, I. (eds) New Technologies, Development and Application IV. NT 2021. Lecture Notes in Networks and Systems, vol 233. Springer, Cham. https://doi.org/10.1007/978-3-030-75275-0_80

Integration of Artificial Intelligence-Based Systems in Diagnostic Pathways: TRUEAID Case Study

Lemana Spahić^{*1}, Almir Badnjević², Asim Kurjak³, Lejla Gurbeta Pokvić¹

Abstract: *Neurological impairment disorders in fetuses, such as cerebral palsy, epilepsy, and autism spectrum disorder, can arise from numerous factors impacting the development of the fetal nervous system. Although diagnosing these disorders early is difficult, it is essential for prompt intervention. Recent progress in deep learning and ultrasound technology offers the potential to create a tool for early detection. Development of the TRUEAID system is based on combining the meticulously tuned Kurjak Antenatal Neurodevelopmental Test (KANET) with a sophisticated convolutional neural network for construction of an AI empowered ultrasound module capable of automated diagnostic decision support in the field of fetal neurodevelopmental risk assessment. The model's performance was evaluated using accuracy metrics, precision, sensitivity, specificity, F1 score, and Mathesson Correlation Coefficient (MCC). The custom CNN architecture achieved an overall accuracy of 93.83%. This pilot study lays the foundation for AI-based fetal neurobehavioral assessment, providing a promising tool for the early detection of fetal neurological impairment disorders. The research holds implications for improving outcomes for affected children and making advanced diagnostic capabilities accessible in diverse healthcare settings.*

Keywords: *fetal neurodevelopment, artificial intelligence, KANET, ultrasound, diagnostic support system*

1. Introduction

Fetal neurological impairment disorders constitute a category of conditions impacting the nervous system's development in fetuses. These disorders can arise from various causes, such as genetic anomalies and environmental influences. Moreover, complications during pregnancy and childbirth also pose a significant risk. Each type of fetal neurological impairment disorder has distinct symptoms and causes. Notable examples include cerebral palsy, intellectual disabilities, epilepsy, and autism spectrum disorder.

The causes of fetal neurological impairment disorders are not always clear, and can vary depending on the specific disorder. However, some common risk factors include maternal infections, exposure to toxins or radiation during

^{*1}Research Institute Verlab for Biomedical Engineering, Medical Devices and Artificial Intelligence, Sarajevo, Bosnia and Herzegovina

²Faculty of Pharmacy, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

³Sarajevo Medical School, Sarajevo School of Science and Technology, Sarajevo, Bosnia and Herzegovina
E-mail: lemana@verlabinstitute.com

pregnancy, and complications during delivery such as birth asphyxia or trauma. Diagnosing fetal neurological impairment disorders can be challenging, and may involve a combination of medical imaging, genetic testing, and neurological evaluations. Fetal neurological impairment disorders pose significant challenges to affected individuals, their families, and healthcare systems. Comprehensive understanding, early diagnosis, and targeted interventions are essential in mitigating the impacts of these disorders and improving the quality of life for affected individuals. Early diagnosis is important, as it can help ensure that affected children receive the appropriate medical care and therapies to help manage their symptoms and improve their overall quality of life. Treatment for fetal neurological impairment disorders typically involves a multidisciplinary approach, with a team of healthcare professionals working together to address the child's specific needs. This may include physical therapy, occupational therapy, speech therapy, and medications to manage symptoms such as seizures or behavioral problems.

The diagnostics of fetal neurological impairment disorders is a delicate and complex process, marked by the integration of advanced technologies and multidisciplinary expertise to assess the developing nervous system within the womb. These disorders, which can arise due to genetic anomalies, environmental factors, or a combination of both, necessitate precise and sensitive diagnostic approaches to facilitate early intervention and optimal management [1]. The complexity of diagnosing fetal neurological impairments requires a collaborative approach. Each professional, from the obstetrician conducting initial screenings to the pediatric neurologist interpreting advanced neuroimaging, plays a pivotal role. Genetic counselors translate complex genomic data into actionable insights, while perinatologists oversee the management of high-risk pregnancies. This collaborative effort extends to parental involvement as parents are integral partners in the diagnostic journey. Their insights, observations, and decisions are central to the personalized and ethical management of fetal neurological impairments.

Ultrasound imaging is a crucial tool for real-time visualization of fetal development. Advanced sonographic technologies enable detailed imaging of the fetal brain, spinal cord, and associated structures. Ultrasound examinations, particularly the mid-pregnancy anatomy scan, offer a visual assessment of the fetus's developing structures. Special attention is given to the brain and spinal cord, where structural anomalies can often be visualized. Neurosonographers carefully assess ventricular sizes, cortical folding patterns, and other structural markers to identify anomalies indicative of neurological impairments. In cases of suspected neural tube defects, spina bifida, or other structural anomalies, targeted ultrasounds provide enhanced imaging resolution, facilitating detailed anatomical assessments and diagnostic precision. These insights are instrumental in prenatal counseling, intervention planning, and parental decision-making

(Abu-Rustum & Daou, 2020). The integration of maternal health data, family history, and environmental factors enriches the interpretative precision of these screenings, offering a preliminary risk profile that guides subsequent diagnostic steps [2].

The Kurjak Antenatal Neurodevelopmental Test (KANET) represents a significant advancement in non-invasive prenatal neurodevelopmental testing. Developed by a team led by Dr. Asim Kurjak, KANET has made a notable impact in perinatal neurology. This test combines advanced technology with clinical expertise to offer unprecedented insights into fetal neurobehavior. Central to KANET's innovative method is the use of four-dimensional (4D) ultrasound technology. This technology allows for dynamic, multi-dimensional visualizations of the fetus in utero, enhancing our understanding of fetal neurological development [3,4].

KANET's uniqueness lies in its ability to conduct detailed, real-time analysis. The 4D ultrasound provides a dynamic observational platform, enabling the detailed study of every fetal movement and behavioral moment. This approach goes beyond mere observation; it deeply examines the spectrum of neurobehavioral indices that reflect the intricate neurodevelopmental processes occurring in the fetal environment. KANET operates on the premise that specific fetal movements and behaviors directly indicate the integrity and functionality of developing neural structures and pathways [5].

2. Decision Support Systems in Healthcare

Decision Support Systems (DSS) in healthcare are essential tools that aid clinicians and healthcare professionals in making well-informed and accurate decisions. These systems utilize a combination of technologies, data, and algorithms to provide insights and recommendations, thereby enhancing the quality and efficiency of healthcare delivery. Healthcare DSS incorporate a wide range of data sources, such as Electronic Health Records (EHRs), laboratory results, and medical imaging data. For example, research by Kawamoto et al. (2005) [6] showed that integrating clinical data into DSS significantly improves clinical practice and patient outcomes. These systems employ advanced algorithms and artificial intelligence to analyze complex datasets, providing personalized patient care recommendations. Clinical Decision Support (CDS) systems, a specific type of DSS, play a crucial role in diagnosis and treatment by analyzing patient-specific data to offer evidence-based recommendations. A study by Osheroff et al. (2012) [7] highlighted the impact of CDS in reducing medical errors, improving healthcare quality, and lowering costs. However, implementing DSS in healthcare presents certain challenges.

Artificial intelligence continues to transform medical diagnosis, with machine learning and deep learning leading the advancements. These technologies have significantly improved the accuracy, speed, and efficiency of diagnosing various medical conditions. The integration of AI in healthcare has been a focal point of research and development in recent years [8]. AI systems, particularly those using machine learning (ML) and deep learning (DL) algorithms, have shown remarkable capabilities in diagnosing diseases, sometimes surpassing human clinicians. AI has also made progress in the field of neurodegenerative disease diagnosis [9].

AI has also made advancements in neurodegenerative diseases. A study by Ding et al. (2021) [10] shows an example of a machine learning model that predicted the progression of Alzheimer's disease. By analyzing multimodal data, including genetic, imaging, and clinical data, the model could identify patients at risk of progression from mild cognitive impairment to Alzheimer's disease. Maroco et al. (2011) [11] demonstrated a machine learning model that utilized data from cognitive tests to identify patients in the early stages of Alzheimer's, enabling timely interventions. The diagnosis of Parkinson's disease is being revolutionized by AI, which analyzes clinical, imaging, and genetic data to identify the disease with superior accuracy. Athey et al. (2021) [12] employed machine learning to analyze voice recordings, effectively differentiating between individuals with and without Parkinson's, paving the way for non-invasive, cost-effective diagnostic tools. AI technologies, particularly deep learning, have demonstrated remarkable accuracy in analyzing PET scans to diagnose Parkinson's disease. Models can identify subtle metabolic changes in the brain, offering insights into disease onset and progression. Sarraf et al. (2020) [13] presented a that effectively distinguished between Parkinson's patients and controls based on PET images. Machine learning algorithms are utilized in the analysis of structural MRI data for the early detection of Huntington's disease. These algorithms can identify and measure specific brain structure changes, aiding in early diagnosis and monitoring of disease progression. Eskildsen et al. (2013) [14] applied a machine learning approach to analyze structural MRI scans to distinguish between early-stage Huntington's patients and controls. AI's role in monitoring and managing multiple sclerosis (MS) is also growing. AI algorithms analyze MRI scans to identify and quantify MS lesions, providing insights into disease progression and treatment effectiveness. Eitel et al. (2019) [15] illustrated how deep learning could precisely delineate and measure MS lesions over time, offering a reliable tool for personalized treatment monitoring while Brown & Hamarneh (2016) [16] illustrated how machine learning could analyze MRI data to identify and quantify brain lesions in multiple sclerosis patients, offering insights into disease progression and treatment response. AI is proving to be instrumental in utilizing diffusion tensor imaging (DTI) for the diagnosis and monitoring of multiple sclerosis (MS). Deep learning models can

analyze DTI data to detect white matter lesions and abnormalities, essential for MS diagnosis as described by Danelakis et al. (2019) [17]. They have described a CNN that efficiently processed DTI data to classify MS patients and healthy controls. By analyzing complex datasets including neuroimaging, genomic, and clinical data, AI models can predict the onset and progression of Alzheimer's with high accuracy (Danelakis et al., 2019) [17]. AI applications extend to the analysis of functional MRI (fMRI) data for the diagnosis and monitoring of Amyotrophic Lateral Sclerosis (ALS). By analyzing brain activity patterns, AI models can identify functional anomalies indicative of ALS. Welsh et al. (2013) [18] implemented machine learning to classify ALS patients and healthy individuals based on resting-state fMRI data.

3. Trueaid Development

The dataset used to develop the AI-driven decision support system, named TRUEAID, consisted of 3D ultrasound images. From 2021 to 2023, a total of 10,452 samples were collected for the development of TRUEAID, which stands for TRUStworthy Artificial IntElligence system for fetal neurological risk assessment and Diagnostic support. These images were extracted from 4D ultrasound recordings of fetuses obtained during the KANET test, a prenatal evaluation of neurological development conducted in the third trimester of pregnancy. KANET assessments are conducted on pregnancies both healthy and suspected to have pathology, with a larger dataset available from normally developing fetuses in healthy pregnancies utilized for developing the TRUEAID system. The dataset used in this project was graciously provided by Dr. Panos Antsaklis from Alexandra Maternity Hospital in Athens, a prominent testing center within the KANET network known for its extensive data resources.

The methodology employed in this study was systematic and adhered to rigorous scientific standards, involving four key stages: data preprocessing, data augmentation, development of Convolutional Neural Network (CNN) models, and interface development. The workflow of TRUEAID's development has been detailed in the work of Spahić et al. [19].

The next step is integrating the AI model with the ultrasound device. This requires developing software that allows the AI to process images in real-time or near real-time, designing an intuitive user interface, and ensuring compatibility with different ultrasound machines and PACS systems. Following integration, the system undergoes extensive testing and validation. This includes bench testing in a controlled environment and clinical trials to assess its effectiveness in real-world settings. Feedback from clinicians is crucial during this phase to refine the system further.

Regulatory approval is a critical milestone. Detailed documentation covering all aspects of the AI-DSS, including design, development, validation, and risk management, is prepared and submitted to relevant regulatory bodies. Ongoing compliance with standards, such as ISO 13485, is maintained. Once approved, the system is deployed in clinical settings. Comprehensive training is provided to clinicians and technicians to ensure they can use the new system effectively. A robust support system is also established to address any issues that arise.

Finally, the system's performance is continuously monitored, with regular updates and improvements made based on new data and user feedback. An ongoing feedback loop with users helps gather insights for further enhancements, ensuring the AI-DSS remains effective and reliable in supporting clinical decisions.

4. Trueaid Concept Validation and Integration

Upon development of the AI-based decision support system, its performance was evaluated in a two fold way in order to ensure its ability to be implemented in a realistic setting. The first validation pathway relied on conventional validation with a subset of 3D ultrasound images extracted from the initial dataset and intended for subsequent validation while the second pathway relies on a near real-time validation setup that allows for accurate simulation of decision making during a diagnostic process.

During conventional validation the system has achieved an overall accuracy of 96% making it surpass the accuracy threshold set for AI-based decision support systems used in healthcare.

Table 1. Performance evaluation of TRUEAID

Class	Specificity	Sensitivity	Accuracy
Face	0.98	1.00	0.94
Hand to face	0.97	0.89	0.95
Legs	1.00	0.94	0.98
Thumb	0.98	0.93	0.97
Overall	0.98	0.94	0.96

The second validation pathway required development of a GUI for seamless interaction with the system. A GUI was developed and implemented as described in Spahic et al., Once developed and proven, TRUEAID can be used anywhere in the world, from low resource to high resource settings enabling better care of affected populations and supporting the fight against non-communicable neurological diseases.

The integration pathway is described through Figure 2.

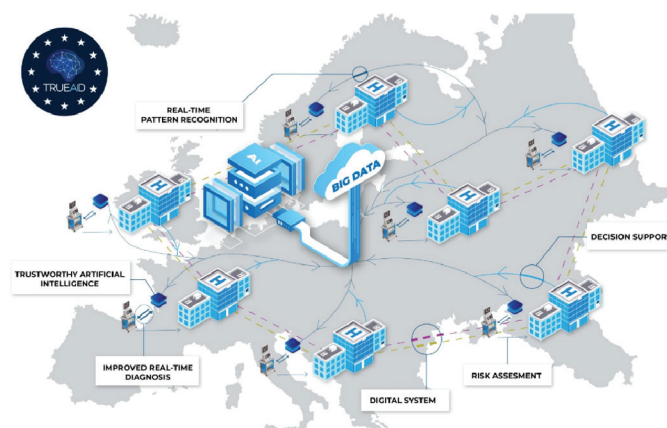


Figure 2. Implementation of TRUEAID through integration in ultrasound devices

TRUEAID is envisioned as a cloud-based system that runs on high-performance computers located off-site from the testing premises. The reason for this is ensuring its availability everywhere from high to low computational-resource settings. As it can be inferred from Fig. 2, TRUEAID will be integrated as a module of an ultrasound device that has the 4D ultrasound modality. During the conduction of a general pregnancy exam using a 4D ultrasound, the obstetrician will be able to run the module with a single click. This will result in initiation of a feedback loop with the cloud-based TRUEAID where the 4D recording will be loaded frame by frame. Each time a characteristic fetal movement is detected, the physician will get a pop-up. The detected movements will be recorded, summed and fed into an if-else algorithm that will mimic the decision making process that trained KANET experts perform during examination. In the end of the recording, the physician will be able to click on „results“ button where an entire record of recognized characteristic movements, their incidence and test-result interpretation will be displayed. Upon analyzing the record, the physician will give the overall result to the mother.

5. Conclusion

TRUEAID is envisioned as an AI-based platform for neurological impairment disease detection from fetal 4D ultrasound recordings. Initially, the prototype is planned to be tested in Ian Donald Schools of Medical Ultrasound in Bosnia and Herzegovina, Croatia, Slovenia, Czech Republic and Greece. TRUEAID will be integrated into the 4D ultrasound screening and KANET test, enabling the 4D recordings to be uploaded to the platform and AI-based evaluation to be performed. Advances beyond the state of the art are seen in the developed

predictive model that employs deep learning architectures for easier and faster decision making process in terms of neurological disease detection. The benefit of the integrated trustworthy AI-based system that will differentiate normal and abnormal neurological development of the fetuses, by constantly monitoring the intrauterine movements and predicting the risk of neurological impairment disorders in fetuses is in its objectiveness, compared to the subjective evaluation by doctors. This system will aid doctors in determining any pattern that could lead to potential neurological impairment in both, prenatal and postnatal neonatal period. The system will specially be beneficial for pregnant women, as they will have insight in the fetal diagnosis, resulting in the better treatment. With the reliable and ethical detection diagnosis of the neurological impairment diseases in the prenatal period, doctors would be able to investigate the proper treatment to potentially overcome the neurological impairment while the baby is still in utero the mother's belly.

TRUEAID solution can revolutionize the detection of neurological impairment enabling early possible and preliminary diagnosis during the antenatal period, allowing investigation of possible treatments in this phase as well.

6. References

- [1] American College of Obstetricians and Gynecologists (ACOG). (2020). Practice Bulletin No. 162: Prenatal Diagnostic Testing for Genetic Disorders. *Obstetrics & Gynecology*, 135(5), e238-e252. DOI:10.1097/AOG.0000000000003892
- [2] Salomon, L. J., Alfirevic, Z., Berghella, V., Bilardo, C., Hernandez-Andrade, E., Johnsen, S. L., ... & Toi, A. (2019). Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound in Obstetrics & Gynecology*, 37(1), 116-126. DOI:10.1002/uog.18810
- [3] Kurjak A, Azumendi G, Andonotopo W, Salihagic-Kadic A. Three- and four-dimensional ultrasonography for the structural and functional evaluation of the fetal face. *Am J Obstet. Gynecol* 2007;196:16–28
- [4] Kurjak, A., Pooh, R. K., Merce, L. T., Carrera, J. M., Salihagic-Kadic, A., & Andonotopo, W. (2005). Structural and functional early human development assessed by three-dimensional and four-dimensional sonography. *Fetal Diagnosis and Therapy*, 20(6), 496-508. DOI:10.1159/000088164
- [5] Kurjak, A., Azumendi, G., & Veccek, N. (2003). The potential of four-dimensional (4D) ultrasonography in the assessment of fetal awareness. *Journal of Perinatal Medicine*, 31(6), 471-479. DOI:10.1515/JPM.2003.070
- [6] Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: a

- systematic review of trials to identify features critical to success. *BMJ*, 330(7494), 765.
- [7] Osherooff, J. A., Teich, J. M., Middleton, B., Steen, E. B., Wright, A., & Detmer, D. E. (2012). A roadmap for national action on clinical decision support. *Journal of the American Medical Informatics Association*, 14(2), 141-145
 - [8] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4).
 - [9] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature medicine*, 25(1), 24-29.
 - [10] Ding, Y., Sohn, J. H., Kawczynski, M. G., Trivedi, H., Harnish, R., Jenkins, N. W., ... & Franc, B. L. (2019). A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. *Radiology*, 290(2), 456-464.
 - [11] Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., & de Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity, and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees, and random forests. *BMC Research Notes*, 4(1), 299.
 - [12] Athey, T. L., Teich, E. G., & Ramanan, V. K. (2021). Deep learning for improved detection and characterization of neurodegenerative disorders. *Current Opinion in Biomedical Engineering*, 17, 100275.
 - [13] Sarraf, S., Tofighi, G. (2020). Deep Learning-based Pipeline to Recognize Alzheimer's Disease Using fMRI Data. *Future of Medical Education Journal*, 10(1), 14-21.
 - [14] Eskildsen, S. F., Coupé, P., Fonov, V., Manjón, J. V., Leung, K. K., Guizard, N., ... & Collins, D. L. (2013). BEaST: brain extraction based on nonlocal segmentation technique. *Neuroimage*, 59(3), 2362-2373.
 - [15] Eitel, F., Soehler, E., Bellmann-Strobl, J., Brandt, A. U., Ruprecht, K., Giess, R. M., ... & Wuerfel, J. (2019). Uncovering convolutional neural network decisions for diagnosing multiple sclerosis on conventional MRI using layer-wise relevance propagation. *NeuroImage: Clinical*, 24, 102003.
 - [16] Brown, A. D., & Hamarneh, G. (2016). Machine learning on human connectome data from MRI. In *Machine Learning Meets Medical Imaging* (pp. 3-10). Springer, Cham.
 - [17] Danelakis, A., Theoharis, T., & Verganelakis, D. A. (2019). A review of the use of CNNs and Attention Mechanisms to classify neurological disorders based on MRI scans. *Journal of Imaging*, 5(11), 83.

- [18] Welsh, R. C., Jelsone-Swain, L. M., & Foerster, B. R. (2013). The utility of independent component analysis and machine learning in the identification of the amyotrophic lateral sclerosis diseased brain. *Frontiers in Human Neuroscience*, 7, 251.
- [19] Spahić, S., Kurjak, A., Stanojević, M., Badnjević A., Gurbeta Pokvić L. (2024). Trustworthiness of Four-dimensional Ultrasound and Artificial Intelligence in Improving KANET Test for Detection of Fetuses at Neurorisk. *Donald School Journal of Ultrasound in Obstetrics and Gynecology*, 18(1), 6-16.

Liver Disease Classification Using Machine Learning

Madžida Hundur Hiyari^{*1}, Nejra Merdović¹, Faruk Bećirović¹,
Emina Mrđanović¹, Adna Softić¹

Abstract: *Hepatitis C virus (HCV) is a significant cause of liver-related diseases including acute and chronic hepatitis, cirrhosis, and hepatocellular carcinoma. Despite the availability of advanced treatments, underdiagnosis remains a critical challenge, particularly in resource-limited settings. This study explores the application of machine learning algorithms, specifically the K-Nearest Neighbors (KNN) method, to enhance the diagnosis of HCV by classifying patients into healthy, potentially diseased, and diseased categories based on liver function test results. Using a biomedical dataset of 615 patients, the model achieved high accuracy (99%), precision (98%), and sensitivity (99%), indicating its potential effectiveness in identifying HCV-infected individuals. The study highlights the importance of feature selection in improving model performance and discusses the implications of the findings for enhancing HCV diagnosis and management.*

Keywords: *Hepatitis C virus (HCV), machine learning, K-Nearest Neighbors (KNN), classification*

1. Introduction

There are many microbes, toxins, autoimmune diseases and neoplastic diseases that cause liver inflammation, but 5 viruses (hepatitis A, B, C, D, and E) cause liver disease as their main pathogenesis [1]. The hepatitis C virus (HCV) as a positive-sense, single stranded RNA flavivirus with seven known subtypes represents the main cause of acute and chronic hepatitis, cirrhosis, and hepatocellular carcinoma [2]. It is estimated that 71 million people worldwide are infected with the virus, which is leading us to an estimated prevalence of 0.64% of the total population of the European Union. This virus is usually transmitted by blood or body fluids, through sharing needles, blood transfusion or by transmission to infants born to HCV viraemic mothers [3]. HCV diagnosis must be as simple as possible and should be linked to the HCV process. There are different categories of laboratory tests: classical serologic tests able to detect anti HCV antibodies (Ab) cells indirect tests, and assays that can detect and

^{*1}Verlab Research Institute for Biomedical Engineering, Medical Devices and Artificial Intelligence, Sarajevo, Bosnia and Herzegovina
E-mail: madzida@verlabinstitute.com; adna@verlabinstitute.com

quantify HCV particle components such as HCV RNA or HCV Core Antigen (cAg) named direct tests [4]. These tests are complementary for the diagnosis of infection, therapeutic decisions and to assess treatment properly. Unfortunately, the underdiagnosis of HCV remains a major obstacle to attaining global eradication, despite advances in therapy [5]. A major challenge is the low diagnosis rate, especially in nations with little resources, which is mostly caused by the expensive price of necessary molecular diagnostic instruments. Furthermore, because extensive HCV testing is less popular than it is for HIV, people infected with chronic hepatitis C (CHC) frequently remain unaware of their infection due to the mild clinical signs associated with the illness [6].

Machine learning (ML) algorithms are adept at analysing medical phenomena by capturing complex and nonlinear relationships in clinical data. The ML algorithms, such as classification techniques, can be utilised to develop a model to diagnose HCV by identifying people who are infected with the virus. However, inappropriate characteristics in the attribute set can spoil the classifier's performance [7]. Feature selection defines a subset of features or variables that describe data to obtain a more compact and essential representation of the available information and ignore all other redundant and irrelevant features [8]. Feature selection is a powerful way to enhance the functioning and reduce the model development time of a classifier.

2. Methods

In the context of supervised machine learning, algorithms are trained on labeled datasets, with each training example consisting of an input and a correspondingly correct output. The main goal is to deduce a function that connects inputs to intended outputs so that the model can forecast new, unobserved data with accuracy. Choosing a suitable model, training it on the labeled data, and assessing its performance with a different test set are the steps in this learning process. The two primary tasks of supervised learning are regression, which anticipates continuous values, and classification, which predicts categorical labels. Its applications cover a wide range of fields, including natural language processing, picture identification, and medical diagnostics, demonstrating its adaptability and crucial significance in contemporary data-driven decision-making. [9]

For classification and regression applications, the K-Nearest Neighbors (KNN) method is a straightforward yet effective supervised learning method. It functions based on the idea that comparable data points should produce comparable results. When a new input is given, KNN finds the 'k' nearest data points in the training set, usually with the use of a distance metric like Euclidean

distance, then predicts the output based on these neighbors' average value (for regression) or majority class (for classification). A critical hyperparameter influencing algorithm performance is 'k'; a small value of 'k' may cause overfitting, while a big value of 'k' may cause the model to be oversimplified [10].

KNN stands out for being straightforward, simple to use, and effective, especially in situations where the decision boundary is wildly erratic. However, because it needs to calculate the distances to every training sample for every prediction, it can be computationally demanding, especially with large datasets. In spite of this, KNN is still often utilised because of its ease of use and capacity to manage multiclass classification issues. In the current era of big data, KNN methods offer an especially effective approach for identifying valuable patterns and creating case-based reasoning algorithms for artificial intelligence (AI). [11]. Based on the structure of the dataset and the data within it KNN is chosen for the classification task due to its effectiveness in similar tasks.

The dataset used in this study is from the field of biomedicine and tracks the health status of 615 patients who underwent blood tests specifically to check liver function. Within this group, 533 patients had good blood profiles and normal liver function, categorized as 0, while the remaining 82 patients were sent for additional tests due to suspected liver disease, categorized as 1, or immediately confirmed with liver disease, categorised as 2. The dataset consists of: identification number (ID), category label (Category), age (Age), gender (Sex), and measurements of 10 different blood components. To adequately prepare the data for machine learning applications, all missing values (total of 31) were replaced with the mean value of the respective column. The data is split into two groups: training data (80% of the dataset) and testing data (20% of the dataset).

The goal of the study is to classify patients into healthy, potentially diseased, and diseased categories. The chosen classifier is the K-nearest neighbors (KNN) due to its simplicity and quick execution. Table 1 lists the selected parameters of this classifier. N-neighbors is a crucial parameter for this method, defining the number of nearest neighbors considered when making decisions. This parameter needs careful selection as it impacts the model's complexity and susceptibility to overfitting or underfitting. In this study, a value of 5 neighbors was chosen, which was shown to be optimal during testing. The "weights" parameter refers to

the weighting function used in prediction. In this study, a uniform function was selected, meaning all neighbors have equal weight in decision-making. This implies that all neighbors are considered equally regardless of their distance from the query sample. The distance measure chosen is the commonly used Euclidean distance. Using Equation 1, it measures the straight line between the input point and the point being measured.

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (1)$$

Table 1. Selected parameters for KNN classifier

Method	K – nearest neighbors
n_neighbors	5
weights	uniform
metric	Euclidean

3. Results and Discussion

After training and testing the created machine learning model, evaluation of the results was conducted using the parameters: accuracy, precision, sensitivity (Table 2), and analysis of the confusion matrix (Figure 1).

Table 2. Value of parameters used for evaluation of the created machine learning model

Parameters	Percentage
Accuracy	99%
Precision	98%
Recall	99%

An accuracy of 0.99 was achieved, indicating a high proportion of correctly classified instances among the total number of instances. However, due to the

imbalanced classes in the dataset used, drawing conclusions based solely on accuracy is not sufficient to assess the model's performance. Therefore, further analysis was carried out.

A precision of 0.98 was achieved, indicating a low rate of false positive classifications, which is also evident when observing the confusion matrix from Figure 1. Sensitivity is another significant parameter for analyzing model performance, and in this study, a sensitivity of 0.99 was achieved. This indicates that the model identifies nearly all actual positive instances.

These results indicate that the combination of the uniform weighting function and Euclidean distance measure provides an effective and reliable means of decision-making.

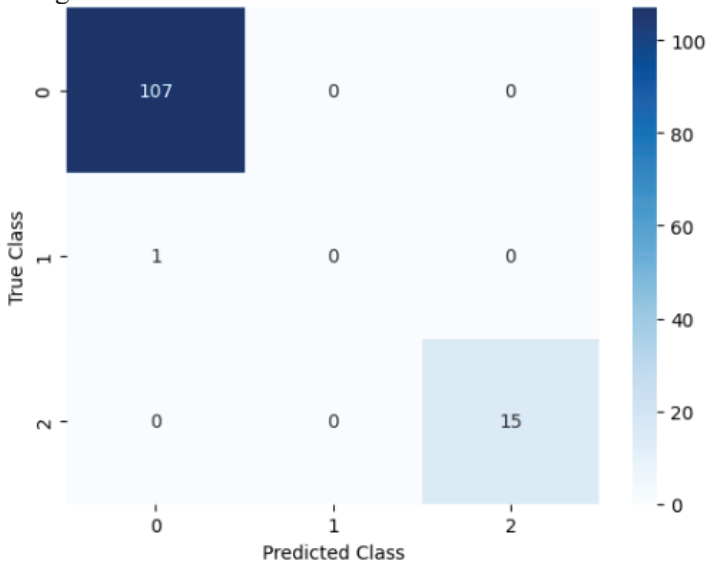


Figure 1. Confusion Matrix

From the confusion matrix, we can see that:

- 107 instances were correctly classified as class 0 (healthy patients),
- 1 instance that belongs to class 1 (potentially diseased patients) was incorrectly classified as class 0 (healthy patients),
- 15 instances were correctly classified as class 2 (diseased patients).

There were very few instances classified as false positives or false negatives, with the majority of instances being correctly classified. Based on all the above, it can be concluded that the model's performance is satisfactory.

To further understand the model's robustness, additional metrics, including F1-score and the Area Under the Receiver Operating Characteristic (ROC-AUC) curve were calculated. The F1-score, which considers both precision and recall, was 0.99, demonstrating balanced performance across both metrics. ROC-AUC scores for the model are as follows: 0.97 for class 0, 0.5 for class 1, and 1.0 for class 2, with an average of 0.99 (Figure 2). These results suggest that the model performs exceptionally well for classes 0 and 2, but has limited discriminatory ability for class 1, indicating potential areas for improvement in classifying potentially diseased patients.

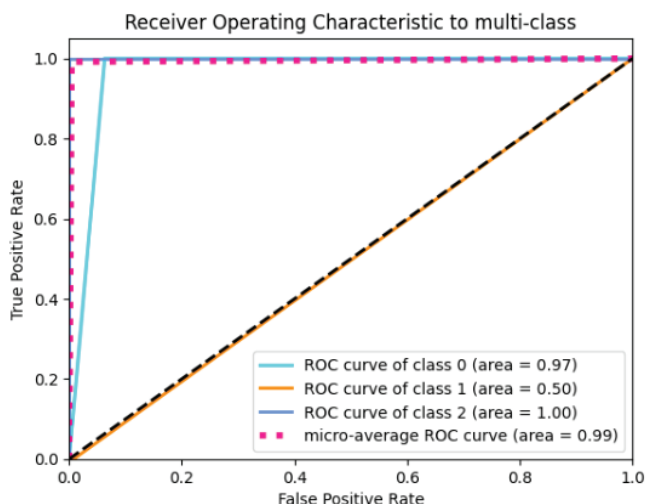


Figure 2. Receiver Operating Characteristic to multi-class

At the end, overfitting of the machine learning model was checked using cross-validation. It was ensured that each subset (10-fold in this case) of cross-validation was representative of the entire dataset in terms of class distribution. An average accuracy of 0.99 was obtained through cross-validation. Since this average accuracy is equal to the accuracy achieved on the test dataset, it can be concluded that the model did not overfit. In other words, the model generalizes well to new data.

4. Conclusion

This study confirms the effectiveness of the K-Nearest Neighbors (KNN) algorithm in classifying patients based on liver function test results. Using a dataset of 615 patients, we achieved high performance metrics: 99% accuracy, 98% precision, and 99% sensitivity. The low rates of false positives and false negatives, validated by the confusion matrix, further support the model's reliability.

Despite class imbalance, the KNN model demonstrated robustness and generalizability, as verified by cross-validation. These results highlight KNN's suitability for medical diagnostics, particularly in multiclass classification tasks. Future research could enhance KNN's performance by integrating advanced techniques, expanding its applicability in biomedicine and beyond. Specifically, investigating alternative weighting functions to assign greater importance to closer neighbors could potentially improve classification accuracy. Additionally, evaluating different distance metrics or hybrid approaches may capture more relevant similarities in high-dimensional or non-linear data spaces. Expanding the dataset to include larger and more diverse samples will help validate the algorithm's applicability across various liver diseases.

Integrating KNN with other machine learning techniques and ensemble methods could further enhance diagnostic performance and robustness, making the approach more effective in clinical settings.

5. References

- [1] Prasidhrathsint K, Stapleton JT. Laboratory Diagnosis and Monitoring of Viral Hepatitis. *Gastroenterol Clin North Am.* 2019 Jun;48(2):259-279. doi: 10.1016/j.gtc.2019.02.007. Epub 2019 Apr 1. PMID: 31046974; PMCID: PMC10461253.
- [2] Duncan, J.D.; Urbanowicz, R.A.; Tarr, A.W.; Ball, J.K. Hepatitis C Virus Vaccine: Challenges and Prospects. *Vaccines* 2020, 8, 90. <https://doi.org/10.3390/vaccines8010090>
- [3] Abu-Freha N, Mathew Jacob B, Elhoashla A, Afawi Z, Abu-Hammad T, Elsana F, Paz S, Etzion O. Chronic hepatitis C: Diagnosis and treatment made easy. *Eur J Gen Pract.* 2022 Dec;28(1):102-108. doi: 10.1080/13814788.2022.2056161. PMID: 35579223; PMCID: PMC9116263.
- [4] Shahid I, Alzahrani AR, Al-Ghamdi SS, Alanazi IM, Rehman S, Hassan S. Hepatitis C Diagnosis: Simplified Solutions, Predictive Barriers, and Future

- Promises. Diagnostics. 2021; 11(7):1253.
<https://doi.org/10.3390/diagnostics11071253>
- [5] Sonia Arca-Lafuente, Paula Martínez-Román, Irene Mate-Cano, Ricardo Madrid, Verónica Briz, Nanotechnology: A reality for diagnosis of HCV infectious disease, *Journal of Infection*, Volume 80, Issue 1, 2020, Pages 8-15, ISSN 0163-4453, doi: [10.1016/j.jinf.2019.09.010](https://doi.org/10.1016/j.jinf.2019.09.010).
 - [6] Roger S, Ducancelle A, Le Guillou-Guillemette H, Gaudy C, Lunel F. HCV virology and diagnosis. *Clin Res Hepatol Gastroenterol*. 2021 May;45(3):101626. doi: [10.1016/j.clinre.2021.101626](https://doi.org/10.1016/j.clinre.2021.101626). Epub 2021 Feb 23. PMID: 33636428.
 - [7] John GH, Kohavi R, Pfleger K (1994) Irrelevant features and the subset selection problem. *Machine learning proceedings*. Elsevier, Amsterdam, pp 121–129
 - [8] Triantaphyllou E, Felici G (2006) *Data mining and knowledge discovery approaches based on rule induction techniques*. Springer, New York.
 - [9] Shruthi H. Shetty, Sumiksha Shetty, Chandra Singh, Ashwath Rao. *Supervised Machine Learning: Algorithms and Applications*. February 2022. doi: [10.1002/9781119821908.ch1](https://doi.org/10.1002/9781119821908.ch1).
 - [10] K. Taunk, S. De, S. Verma and A. Swetapadma. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1255-1260, doi: [10.1109/ICCS45141.2019.9065747](https://doi.org/10.1109/ICCS45141.2019.9065747).
 - [11] S. Zhang, "Challenges in KNN Classification," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 10, pp. 4663-4675, 1 Oct. 2022, doi: [10.1109/TKDE.2021.3049250](https://doi.org/10.1109/TKDE.2021.3049250).

Artificial Intelligence and its Application in Dentistry

Nemanja V. Majstorović^{*1}, Sonja Dimitrijević², Aleksandar Todorović³

Abstract: Artificial intelligence (AI) has today become an important factor in the social, scientific and economic development of human society. It has wide application possibilities in the economy and public services, which also applies to dentistry. Its history is about eight decades long, and a special area of its research and development is machine learning models, which today constitute the basic model of AI application. The aim of this paper is to analyze the state of development and application of AI in dentistry, with a proposal for the possibility of its application in the future. Based on that, the paper is structured in such a way as to provide answers to the following questions: what is AI and how was it developed, how were AI models in dentistry created and developed, what are the prospects for the application of AI in dentistry, and possible directions for future research in this area.

Keywords: Artificial intelligence, Machine learning, Dentistry

1. Introduction

Artificial intelligence is a branch of intelligence science, which includes: natural and artificial intelligence. Natural intelligence is the science of models of intelligent behavior of living beings, while artificial intelligence, or AI, is also the science of developing intelligent software and machines. They are connected and have contributed to each other's development in the past eight decades, when AI appeared. Advances in natural intelligence have set new paradigms for AI research, such as neural networks (ANN), genetic algorithms (GA), ant colony optimization (ACO), etc., while advanced AI tools have helped accelerate new discoveries in natural intelligence [1,2]. That's why we can say that the eight-decade long history of AI is promising today, more than ever, which also applies to dentistry.

This review paper aims to provide readers with basic facts and knowledge about the development and application of AI in dentistry to date. It is intended for scientists, researchers and students, who want to get concise information about the beginning of the development of a new scientific branch in dentistry, intelligent dentistry, based on AI.

^{*1}Dental practice ELIDA 1, Belgrade

²"Mihajilo Pupin" Institute, Belgrade

³Faculty of Dentistry, Belgrade

E-mail: nmajstorovic961@gmail.com, sonja.dimitrijevic@pupin.rs, aleksandar.todorovic@stomf.bg.ac.rs

The paper itself has several parts: (i) how AI was created and developed, (ii) an overview of the development and application of AI for dentistry, to date, (iii) challenges and future development of AI for dentistry, (iv) conclusions and future research.

2. A Brief History of AI

Before detailing intelligent dentistry (dentistry supported by AI techniques, especially deep machine learning (Deep ML), it is important to briefly analyze the history of AI, which is shown in Table 1 [1,2,21]. It covers an interval of eight decade (1940 - 2020), and has four stages of development, as indicated in the table below.

Table 1. Overview of the development of artificial intelligence (AI) (adapted from [1,2,21])

Year of appearance	The period of AI development	The name of the event	Characteristic
1943	The Dark Age of AI	Binary ANN model (W. McCulloch, W. Pitts)	A two-state neuron
1950		Turing test - can machines think? (A. Turing)	Communication: examiner - examinee - computer
1951		The first neural computer (M. Minsky, D. Edmonds)	A network of forty neurons
1956	The rise of AI	First AI Workshop (J. McCarthy)	Artificial Intelligence (AI) Defined
1958		LISP - The First AI Language (J. McCarthy)	Still in use today - Functional language
1961		General Problem Solver (GPS)(A. Newell, H. Simon)	Uses formal logic to find solutions (does not solve complex cases)
1965		Fuzzy sets (L. Zadeh)	A continuous transition from not belonging to complete belonging
1969		Dendral (E. Feigenbaum, B. Buchanan , J. Lederberg , C. Derassi)	The first expert system - Spectrogram identification of chemical compounds
1970	AI "Winter"	ANN Learning (A. Bryson and Y. Ho)	Backpropagation model
1975		GA (J. Holland)	The first genetic algorithm
1976		Mycin (E. Shortliffe)	The first expert system for diagnosis
1982		Hopfield networks (J. Hopfield)	Memory systems with binary boundary nodes, a model for understanding human memory
1986		Backpropagation (BP) (D. Rumelhart, J. McClelland) and DAI (Distributed AI) (A. Bond , L. Gasser)	Computation of neural error (BP) and distributed solutions for AI

1992		Genetic programming (GP) (J. Koza)	LISP symbolic code
1995		Intelligent agents (M. Wooldridge , R. Jennings)	They act on the environment in an intelligent way
2002	AI is becoming a	ACO, PSO, AIO, DNA computing	AI tools
2006	science	Honda ASIMO robot	Moving and climbing stairs
2016		AlphaGo DeepMind	Man computer games
2017		IBM Watson	Man computer games
2022	Chat GTP	Open AI	AI for different fields

It can be said that the history of AI starts from the early 40s of the 20th century. It started with the binary ANN model [3], defined by W. McCulloch and W. Pitts of the University of Illinois in 1943. Although their model only considered a binary state (ie on/off for each neuron), this model was used as a starting point for ANN research later in the 1980s. The next important year is 1950, when A. Turing proposed the famous test, named after him, by means of which it is determined whether machines can think [4] ? The Turing test is performed through computer communication, which includes an examiner, a person (subject), a participant in the experiment, and a computer, which are separated from each other, but communicatively connected. The examiner can ask any questions. If the examiner cannot distinguish a machine from a human based on their answers, the computer has passed the test. Then, in 1951, M. Minsky and D. Edmonds [5], researchers from Princeton University, created the first neural computer that simulates the operation of a network of 40 neurons.

We can say that the turning point in the development of AI occurred when the first AI workshop was organized in 1956 at Dartmouth College under the leadership of J. McCarthy [6]. He marked the end of the "dark age of AI" and the beginning of the "rise of AI" as the next stage in the development of AI. The term "artificial intelligence" was then proposed by J. McCarthy, and it is still in use today. J. McCarthy later moved to the Massachusetts Institute of Technology (MIT), where in 1958 he defined the first AI language, LISP, which is still used today. One of the most ambitious projects of those years in this area was the General Problem Solver (GPS), which was defined in 1961 by A. Newell and H. Simon from Carnegie Mellon University [7]. GPS is based on formal logic and can generate an infinite number of operators trying to find a solution, however, it is not effective in solving complicated problems. In 1965, L. Zadeh from the University of California, Berkeley published the famous paper "Fuzzy sets", which became the basis of the theory of fuzzy sets [8], or uncertain decision making.

The first expert system, Dendral [9], was developed at Stanford University in 1969, in a project funded by the National Aeronautics and Space Administration (NASA) led by J. Lederberg, winner of the Nobel Prize for genetics. However, at the time, most AI projects could only solve gaming problems, rather than real-

world ones, so many projects in the United States, Great Britain, and several other countries failed, or were useless. Thus, AI research has entered the "AI winter".

Despite reduced funding, AI research continued, and in 1970, A. Bryson and Y. Ho [10] proposed a back propagation model for neural network learning. Moreover, the first genetic algorithm (GA) model was proposed in 1975 by J. Holland from the University of Munich, who performed selection, crossover and mutation, using genetic operators for optimization [11] and thus developed the GA model. The same team that developed ES Dendral [12], at Stanford University, developed ES Mycin in 1976, which is based on IF-THEN rules, as an expert system for the diagnosis of blood diseases using 600 rules, if (IF)-then (THEN). Research has shown that he works better than a junior doctor [13].

It was only after 30 years that researchers started researching neural networks again, as an important area of AI. A new era has begun — in which AI has become a science. It was in 1982, when J. Hopfield published his Hopfield networks, which are still popular today [14]. In 1986, back propagation became the first implemented learning algorithm in ANN, 16 years after the publication of this model [15]. Also, this year is the beginning of the application of distributed AI (DAI) through parallel distributed knowledge processing. 22 years later, the theory of fuzzy sets was successfully incorporated into the management of dishwashers and laundry machines, in 1987, by Japanese companies that produced these household appliances. In 1992, J. Koza proposed genetic programming for the manipulation of symbolic code, which was generated in the LISP language [16]. Starting from the idea of DAI and artificial life, intelligent agents were defined in the mid-1990s. At the end of the last century, hybrid systems of fuzzy logic, ANN and GA became popular tools for solving complex problems. Since then, various new AI approaches have emerged, including ACO, particle lake optimization (PSO), artificial immunity optimization (AIO), and DNA computing, as well as intelligent agents [17]. The potential of AI in the future – such as dentistry – remains large and unpredictable.

AI-based chess software was the first program, called Deep Blue, developed by IBM [18]. It was in 1997, and then Garry Kasparov, the then world chess champion, was playing an exhibition match, when he lost 2.5 to 3.5 against the computer. The next example is the Honda ASIMO robot in 2005, which moved up stairs. In this case, the robot moves in an unstructured environment and is controlled by a human. This example required natural language processing, the use of computer vision, object recognition, machine learning and motion control during operation. After this, in 2016, AlphaGo DeepMind software defeated world champion L. Sedol in four out of five games, using cloud computing, machine learning and Monte Carlo algorithm, to search with the support of a deep neural network for decision making [19]. Its latest version, AlphaGo Zero

[20,21], surpassed the ability of AlphaGo in just three days, through self-learning from scratch. A new breakthrough in this area was IBM Watson, an intelligent platform.

A recent breakthrough in AI was in Natural Language Processing (NLP) and in generative AI that allowed for the development of chatbots, virtual assistants, and intelligent tools and platforms such as IBM Watson, and ChatGTP.

The more famous, ChatGTP, is able to create human-like conversational dialogue, ie, to respond to questions and compose written content. Moreover, ChatGPT involves a generative AI that lets users enter prompts to receive required images, text, or videos. Therefore, it can be used in different areas including dentistry.

Today, AI techniques and systems can be used in a variety of fields, from playing chess to production management, patient diagnosis, aircraft control, smart design and planning, and intelligent dentistry. The AI techniques reviewed in Table 2, especially supported by machine learning models, hold much promise for the development of intelligent dentistry. To illustrate all this, Table 2 provides an overview of machine learning models, on different bases.

Table 2. Overview of learning models (*applied in dentistry*) [1,21]

A machine learning model	Supervised / semi-supervised / unsupervised	Discriminative / generative	Learning / deep learning
K-means	Unsupervised	Generative	Learning
K-nearest neighbor	Supervised	Discriminative	Deep learning
Method of support vectors			
Hidden Markov model			
A random forest model			
Extreme Gradient Boosting (XGBoost)			
Ensemble method	Semi-supervised	Generative	Learning
Convolutional neural networks			
Recurrent neural network			
Long short term neural network			
Naive Bayes classifiers	Semi-supervised	Discriminative	Deep learning
Gaussian mixture model			
A generative adversarial network			

Machine learning (ML) models, which are most widely used in dentistry, are marked with red text. In the next chapter, through examples, the application of ML in dentistry, in different areas, will be given.

3. Artificial Intelligence and Dentistry

This system analysis was performed according to the PRISMA methodology [1,21], and our question was "Which AI techniques are applied in dentistry and how does AI improve dental diagnosis, clinical decision-making and the outcome of dental treatment?". For him, the following were defined: period, type of study, AI dental models, search methodology and study quality assessment. The sample consisted of 364 papers, and 41 papers met the set criteria.

Broadly speaking, a systemic analysis of the application of AI in dentistry is performed here with four aspects: (i) general overviews of application in dentistry, (ii) application in a selected area of dentistry, (iii) case studies for solving specific problems, and (iv) application of ChatGPT in dentistry. The following analysis refers to all the mentioned areas.

3.1. Overview of the Application of AI in Dentistry

General analyzes appeared in the last decade, which is understandable, because it was only from those years that the application of AI in dentistry took off. Thus, the first paper [22] in this field gave general aspects of the application of AI in dentistry, citing several application examples (virtual assistants, diagnosis of oral and maxillofacial diseases and therapy planning (ANN)), concluding that AI has a high potential in dentistry, table 3.

Table 3. General overview of the application of AI in dentistry

Purpose	Ref. / yr.	Method	The goal	Use
High potentials of AI in dentistry	[22] 2017	Virtual assistants, diagnosis of oral and maxillofacial diseases and therapy planning (ANN)	Increasing the accuracy of diagnosis and the success of treatment	Diagnosis and treatment planning with AI
Application of AI in dentistry	[23] 2018	ANN and AVR (Augmented Virtual Reality)	Improving the success of diagnosis and treatment	Patient management and decision-making system, oral and maxillofacial surgery, prosthetics and orthodontics
AI as the technology of the future for dentistry	[24] 2018	ANN/CNN	Improve clinical practice	Aid to the dentist, not his replacement

Comparison of current and future applications of AI in dentistry	[25] 2020	CNN/Deep NN	Unstructured data processing (BDA)	Prediction of treatment results
AI platforms	[26] 2020	AI software for dentistry	Specialized platforms for facilitating the application of AI in dentistry	Applying deep learning models for diagnosis and treatment planning
AI in dentistry	[27] 2020	deep learning (CNN)	Increasing the level of reliability of recommendations	Dental radiology
Future applications of AI in dentistry	[28] 2020	CNN / GA (Genetic Algorithm) / Fuzzy logic	Learning from unstructured databases	Dental radiology, diagnosis, patient management and orthodontics
Development and application of AI in dentistry	[29] 2020	ANN/CNN	Analysis of 43 studies	Diagnoses, clinical decision-making and treatment prognosis predictions
AI as support for standard procedures	[30] 2021	AI tools and techniques	Integrated patient information and improvement of professional interactions	Clinical intuition, ineffable perception, and empathy.
AI to support the dentist	[31] 2021	AI tools and techniques	Analysis of 32 studies	Prediction of failure in clinical treatment scenarios, with the proposal of reliable solutions for it
AI as innovation	[32] 2021	Deep learning	Help the dentist to rule out human error	All areas of dentistry
Application of AI/ML	[33] 2021	CNN	Tooth color research for dental interventions	Aesthetic dentistry
Application of AI techniques	[34] 2021	ML in Image Processing (CNN), ANN, GA, FL	Improving decisions of dentists	Analysis of application in different fields
Caries detection	[35] 2021	CNN	High accuracy and precision of diagnosis	Caries diagnosis
The tool of the future for dentistry	[36] 2022	Deep learning	Correct diagnosis for multiple symptoms	Oral dentistry and dental diseases
Negative aspects of AI	[37] 2022	Reduction of narrative contact with the patient	Cyber security	Determining the balance
Application of AI	[38] 2022	AI	Analysis of 28 studies on the applicability and effectiveness of AI models in different areas of dentistry	Optimal clinical treatment of patients

Application of AI in endodontics	[39] 2022	CNN	Improvement of dental practice	Future applications of AI are drug interactions, prognostic diagnosis and robotic endodontic surgery
Application in dentistry	[40] 2022	ANN/CNN	Faster and more successful treatment	Efficiency, accuracy and time savings during diagnosis and treatment planning
Application in dentistry	[41] 2022	ANN/CNN	Analysis of 33 studies	Data collection and analysis techniques
Development and challenges of AI in dentistry	[42] 2023	AI tools	As much confirmation of clinical results as possible	Diagnostics of a wide range of conditions
AI as a transformative technology in dentistry	[43] 2023	AI algorithms and tools (ML)	Automated detection and diagnosis of dental conditions and improving treatment outcomes	It revolutionizes clinical practice, improves patient outcomes and improve the overall efficiency of dental care
Periodontics and caries research	[44] 2023	ANN/CNN	Fast and accurate analysis of recordings, for early disease detection	Improvement of clinical practice
Application in dentistry	[45] 2024	AI tools	Better treatment outcomes	Improvement of clinical practice in different areas of dentistry
Augmented Intelligence (Aul) in Dentistry	[46] 2024	AI/Aul	ADA White Book	ChatGPT is ready to revolutionize dentistry by providing efficient, accurate and up-to-date information dentists
Application of generative AI techniques	[47] 2024	AI tools	SD (Synthetic Datasets)	Various medical databases

The paper [23] discusses the potential of the integration of clinical practice in dentistry and AI techniques, especially for patient data management (ANN), implantology and oral surgery. It is considered that AI is the future of dentistry [24], because its advantages are: accuracy of diagnosis, standardization of procedures and reduction of treatment time, and disadvantages: system complexity and application costs. Today's model of AI in dentistry will be significantly changed and improved in the future [25]. This means, among other things, that today's samples for analysis of recordings from 2-3 thousand units

(individual patients) will be increased to several million units (population), which will also carry out genetic analysis of the phenomenon in the population. Also, as far as data sources are concerned, with today's data from patients and clinics (structured data), they will move from unstructured data from different institutions (sources), etc. New approaches in this area are AI platforms for dentistry [26], as ready-made software products for specified use. This is certainly one of the main approaches in the future, which will appear more and more. One of the most important features of today's approaches to the application of AI in dentistry is deep learning [27], which enables higher quality inference compared to other learning models. This means that diagnoses, treatment plans and other advisory recommendations to the dentist are more reliable. One of the future directions of AI in dentistry is the greater application of unstructured data in deep learning models, such as CNN, GA and FL [28]. The analysis of 43 studies shows that AI-based systems can perform exceptionally well in various dental applications, including: (i) determining and diagnosing dental caries, vertical root fractures, apical lesions, salivary gland diseases, maxillary sinusitis, maxillofacial cysts, metastases in cervical lymph nodes, osteoporosis, cancerous lesions, alveolar bone loss, and predicting orthodontic extractions, (ii) predicting the need for orthodontic treatments and cephalometric analysis, and (iii) determining age and gender. The studies demonstrate that AI-based systems can match or even surpass the performance and accuracy of trained specialists.

Today, it is considered that AI is an excellent aid for dentists, whose customization for the user as well as clinical validation should be worked on even more intensively [29]. Cognition (reasoning and inference) is that aspect of AI that helps the dentist the most in diagnosis, treatment planning or prediction of its progress [30]. AI should be seen as a complementary tool to help dentists, rather than a replacement. By augmenting their skills and abilities, AI can free up dentists to focus on the aspects of their work that require human intuition, empathy, and professionalism. By leveraging AI as a tool to augment their skills, dentists can focus on the aspects of their work that require human expertise and judgment, while also benefiting from the efficiency and accuracy that AI can provide. A systematic analysis of 32 studies [31] showed that AI can be used to manage patient data, support dental diagnosis, as well as treatment planning. In this way, AI helps dentists to provide quality patient treatment, better oral health care outcomes, and achieve high diagnostic accuracy. Also, AI helps predict failure in different treatment scenarios by suggesting reliable solutions for it, or the best case scenario. AI, as a magical innovation today, will never be able to completely replace the dentist [32] , but it will help him work at a higher and more comfortable level, making his profession much easier and without the appearance of human error. An important aspect of using AI in dentistry is the determination of tooth color for prosthetic restorations, so the paper [33]

presents an AI-based system for color recognition and determination. This approach is significantly more accurate than the usual approach - visual color recognition. An analysis of the use of different AI techniques in dentistry is presented in [34]. All of them are used for different learning models, so each one is suitable for a certain area in its own way, for example CNN (deep learning) for image analysis (radiographic image) and diagnosis. AI is being researched in dentistry for various purposes, especially the identification and diagnosis of normal and abnormal tissue structures (diseases), as well as the prediction of treatment results [35]. CNNs have the ability to detect and recognize anatomical structures, especially to recognize and mark diseased teeth on radiographs, with an accuracy of 95.7–99.44%, which is almost competitive with the work of clinical dentists, whose accuracy rate was 99.97 % [36]. The situation is similar with the detection of carious lesions, where the accuracy is 75.4–93.4% and the sensitivity is 74.3–97.2%. AI is no longer a myth but the future in dentistry, so it is safe to say that it will be the most promising tool in the future for dental diagnostics and treatment. It will be an irreplaceable tool in diagnosis, when several symptoms are correlated. However, there are studies that also talk about the negative aspects of the application of AI, which should be taken into account [37]. These are particularly reflected in the following: data accuracy and security (cyber security), lack of narrative communication with the patient, and patient privacy and medical ethics. A total of 28 studies [38] were analyzed, mostly from periodontology (six), so it was concluded that AI is a good way to analyze clinical dental data. Also, analyzes of randomized clinical trials are performed to confirm the accuracy of this concept in dental practice, with the aim of providing dental care based on data. AI models, such as CNN, can have different applications in endodontics [39]: investigation of root canal anatomy, determination of the viability of dental pulp stem cells, measurement of working canal lengths, precise determination of root fractures and periapical lesions, and determination of the success of their retreatment procedures. Because of all of the above, AI is becoming increasingly important in dentistry these days. Neural networks are part of AI and are similar to the human brain in their work, so they can solve given problems and make quick decisions, so they will become an increasingly common tool in modern dentistry in the near future [40]. With AI expanding in dentistry, its models are used to diagnose almost all dental conditions, ranging from routine dental caries to more complex diseases such as: oral cancer, maxillofacial cysts, alveolar bone loss, determining the urgency of orthodontic extractions. The paper [41] shows the different techniques used for data collection and the application of those techniques to problems in dentistry, with an analysis of challenges and limitations, related to the wider use of AI to improve the quality of diagnosis and treatment of patients. AI models are already used today for [42]: diagnosis of various dental conditions, such as: dental caries, apical lesions, vertical root fracture, salivary gland diseases, maxillofacial

cysts, maxillary sinusitis, cancerous lesions, metastases in cervical lymph nodes, osteoporosis, loss of alveolar bone, need for orthodontic extractions, cephalometric analysis, determination of age and sex, etc. It can be said that the use of AI in dentistry refers to several key areas: patient management, image analysis for diagnostics, treatment planning and personalized care [43]. Due to all of the above, joint efforts are needed: dental experts, AI experts and policy makers (primarily health-related), for the development of solid frameworks that will ensure the responsible and ethical application of AI in dentistry. In [44], a detailed analysis of the application of AI in periodontology and caries detection is given, including classification of different types of periodontal disease, identification of areas of bone loss, determination of disease severity, analysis of dental images and detection of early signs of disease. The same is done for caries. The review paper [45] outlined possible areas of application of AI in oral medicine and radiology, maxillofacial surgery, oral pathology, prosthetics, orthodontics, endodontics and pediatric dentistry. The Augmented Artificial Intelligence (AeI) model is an AI model supported by ChatGTP, and represents the ADA (American Dental Association) view in this area [46], given as a white paper. The main obstacle to the progress of AI is access to different data sets (structured and unstructured), which are used to train deep learning models, ensuring optimal performance of AI models, which will be able to help experts in the field [47]. Generative AI (GenAI) also represents a new type of AI application, which is used to generate more synthetic data sets (SDS), which can overcome the problems arising from the use of traditional data sets.

Summarizing the performed analyses, we can conclude the following: (i) AI is increasingly becoming a clinical tool to assist in the diagnosis and treatment of patients in the field of dentistry, and (ii) digitization of dental services is an imperative of the moment, and AI is an integral part of this process, which is especially stimulated by the development of the Dentistry 4.0 concept.

3.2. Some typical examples of the development of AI tools in dentistry

This part of the analysis refers to the development and application of specific AI techniques in dentistry, namely neural networks (ANN and CNN). The application of AI in dentistry is based on learning, using ANN, or advanced learning models, such as Deep ML (machine learning) algorithms and models, which are collectively known as CNN. Thus, there are 10 ANN models and 23 CNN models. The analysis of the application of neural networks shows that ANN models are used for data analysis (data based) and CNN for image processing and analysis (image based), which is used for diagnosing conditions, table 4 [48].

Table 4. Overview of the application of AI tools in dentistry

AI technique	Ref. / yr.	Method	The goal	Use
Application of ANN/CNN	[48] 2018	Data and image processing and analysis	Improving the accuracy of diagnosis and the success of treatment	Diagnosis and treatment planning with AI
Application of CNN	[49] 2019	Fast image processing	Correct diagnosis	Clinical practice - define rules (eg ADA)
CNN	[50] 2019	Analysis of CT images	Determination of the size, shape and curvature of the root canal	Training on 5680 shots
CNN	[51,52] 2020	Analysis of 2D and 3D images	Diagnosis of the condition	Maxillofacial radiology
CDSS and GA	[53] 2020	pattern recognition method (diseases)	Accurate treatment plan	Use of Clinical Decision Support Systems (CDSS)
(CNN, DNN, RNN)	[54] 2020	Deep learning	Diagnosis and treatment plan	Maxillofacial surgery
CNN	[55] 2020	Analysis of CT scans	Diagnosis of lesions	Periodontology
AI and robotics	[56] 2020	Robot path planning	Diagnosis and intervention	Maxillofacial surgery
Digital transformation in dentistry	[57] 2020	From RP to telehealth (five areas)	Improvement of dental services	In all areas of dentistry (augmented intelligence – Ael)
AI	[58] 2021	Segmentation of CBCT images	Application in clinical practice	Oral dentistry
CNN (three models)	[59] 2022	Decision tree	Choosing the best deep learning model	Treatment planning
AI (RCNN)	[60] 2023	IoT (Dental Internet of Things)	On-line monitoring of dental parameters	Personal intelligent dentistry
CNN	[61] 2024	Analysis of recordings	Diagnosis of oral condition	Dental clinical conditions
CNN	[62] 2024	Deep learning	A meta-analysis of 9 studies	Clinical dentistry

CNN, according to [49], can help in the diagnosis in dentistry, as an advisor, by offering the dentist a more comprehensive, systematic and faster evaluation and a more documented picture of the patient's oral health. Therefore, CNN models are increasingly becoming applicable in routine care, but before that, dental authorities (such as the ADA [46]), should evaluate them against evidence-based clinical practice rules. The analysis of the CT image of the teeth for the purposes of diagnosis is shown in the work [50], using the CNN model while maintaining

the optimal resolution of the image, depending on the subject of the image. Image segmentation was also performed, depending on the type of condition of the tooth that was the object of the recording. In [51], a deep model of maxillofacial radiology supported by CNN (DMFR) was proposed, for more effective diagnosis, classification and prediction of the development of oral diseases, treatment planning, as well as assessment and prediction of treatment outcome, minimizing the possibility of dentist error. Training was performed on 6115 recordings, and good results were obtained in clinical practice. A similar analysis was performed in [52]. It is important to point out that for all areas of dentistry, such as: orthodontics, maxillofacial surgery, periodontology, prosthetics, and others, only an accurate diagnosis ensures the correct treatment plan, which is the most correct way to restore the patient's health. The diagnosis and treatment plan is based on the specialist's knowledge, but is subject to a high risk of error due to a large number of factors [53]. Therefore, the application of multiparametric methods of recognizing patterns of the occurrence of conditions (statistics, machine learning and artificial intelligence (AI)), is a great hope, both for doctors and patients. This article provides an analysis of the application of deep learning techniques (CNN, DNN, RNN) in the field of dentistry, for diagnosis and treatment planning for: dental caries, dental pulp diseases, periodontitis, oral cancer, periapical lesions, oral implants and orthodontics [54]. Clinical validation of all mentioned examples is the main measure of the success of these models. In [55], the use of CNN techniques for the analysis of dental radiographs was analyzed. It was shown that one diagnostic limitation of CNN was the size and characteristics of the image, which are used as input data. It was found that the accuracy of the results increases with the increase in image resolution. The main feature of reference [56] is that it presents an analysis of the use of robotics in maxillofacial surgery, with the use of AI. The symbiosis of these two technologies (robotics, AI) provides very precise surgical interventions, and the path of the robotic arm is managed based on the image analysis of the diseased zone, which is determined on the basis of AI. This concept increases the understanding of the pathogenesis of the disease, improves the risk assessment strategy, increases the accuracy of the diagnosis, makes better prediction of the disease, which ultimately leads to better patient treatment outcomes. Digital transformation in dentistry, based on electronic health data, is recognized as one of the main trends in the 21st century, as a challenge in dental health care [57]. This opinion is based on the development and application of the following technologies in dental research: (a) artificial intelligence (AI) and machine learning (ML), (b) rapid prototyping (RP), (c) augmented and virtual reality (AR/VR), (d) personalized dental medicine and (e) telehealth. By integrating these approaches, we arrive at the concept of Augmented Intelligence (AeI) in dentistry. The algorithm for automatic segmentation of teeth is used in the analysis of CBCT images, and in these studies, an analysis of 433 randomly

selected images was performed [58]. It has been shown that the AI gives quite good results in relation to the dentist's diagnosis, with an accuracy of 0.9, which encourages clinicians in oral dentistry to apply this model. An accurate diagnosis and prognosis of an individual tooth must be determined comprehensively and carefully taking into account the broader treatment plan. The model developed and shown in [59], based on AI, uses the Harvard Handbook of Good Dental Practice for Treatment Planning (CTPC). He grades the state of the teeth over five levels, using 17 defined factors. Also, for the development of the learning model, three CNN models were used, on a sample of 2359, and the decision tree method proved to be the most accurate. The Dental Internet of Things (IoDT) is the latest approach for on-line monitoring of patient dental parameters [60], which was developed as an experimental model, supported by the R-CNN network. In this way, the Industry 4.0 model for dentistry (Dentistry 4.0) is being developed. Using AI, meaningful information can be extracted from dental databases, especially dental radiographs, all with the application of deep ML models [61]. This study analyzes ML models, which can help in the diagnosis of clinical conditions, such as: early childhood caries, numbering of primary teeth, periodontal bone loss, cysts, peri-implantitis, oral cancers, osteoporosis, locating a small apical opening, identifying orthodontic landmarks, temporomandibular joint disorders, etc. In [62], a meta-analysis of 9 studies showing the results of AI application, such as: segmentation and classification of teeth, maxillofacial bone segmentation and creation of 3D tooth surface models, was presented. The AI techniques included different CNN models, and the analysis parameters in these studies were specific to the respective dental cases and specialties.

The analysis performed in this chapter allows us to conclude that the development of an integrated extended model of AI (AeI) represents the latest direction of research in the field of AI applications in dentistry. Another research direction is Dentistry 4.0.

4. Artificial Intelligence in Dentistry - What's next [36,43,63]

Virtual (digital) dental assistants can already perform several tasks in the dental clinic today, with greater precision, with fewer errors, as well as less labor, all compared to humans. They can be used to help schedule appointments, manage paperwork, assist in clinical diagnosis or treatment planning, and predict treatment outcomes. They are very useful to quickly acquaint the dentist with the patient's medical history, as well as habits such as alcoholism and smoking of the patient in question. Also, when we have urgent conditions, the patient has the option of emergency teleassistance, especially when the dentist is unavailable. In this way, a virtual patient database is created, which helps in providing the ideal treatment service to the patient himself.

Diagnoses and treatment of dental conditions. AI can be used as an effective approach in the diagnosis and treatment of lesions of the oral cavity, and it can be used in the screening and classification of suspicious changes in the mucous membrane, which goes through premalignant and malignant conditions. Even tiny changes at the level of a single pixel are detected, which can go unnoticed by the naked eye. Applying a deep learning model also accurately predicts the genetic predisposition to oral cancer of a large population.

Oral and maxillofacial surgery. The most significant use of AI is the application of robotic surgery, where the movement of the human hand and elements of human intelligence (reasoning and reasoning) are simulated. Clinical application of this approach includes: oral implant surgery, removal of tumors and foreign bodies, biopsy and jaw surgery. Comparative studies of oral implant surgery indicate significantly higher accuracy compared to manual freehand surgery, even when performed by experienced surgeons. The operation time is shorter, safer manipulation around delicate tissue structures and greater intraoperative precision. Image-guided robotic surgery also provides detailed surgical resection, potentially reducing the need for revision procedures.

Prosthetics. The goal dentist is yesmake an ideal aesthetic prosthesis for the patient, but various factors affect it: anthropological calculations, errors in facial measurements, errors due to ethnicity and patient wishes. Therefore, in order to reduce or eliminate these errors, a virtual assistant for designing solutions in prosthetics, based on AI, is used here. It connects computer-aided design (CAD), knowledge-based systems (ES) and databases, using logic-based representation (AI) as a unifying medium for the ultimate case solution. CAD/CAM systems in dentistry provide ready-made dental solutions, through a fine milling process of finished ceramic models. These systems are used in the production of inlays, crowns and bridges. CAD/CAM technique, in this way, creates 2D and 3D models, which are translated into finished solutions by CNC machine processing. In this way, the long and laborious process of conventional casting was replaced, and the impact of human error in the final prosthesis was also reduced.

Orthodontics. Diagnosis and treatment planning can only be determined by analyzing X-rays and intraoral scanners and cameras. All this helps to build a virtual dental assistant, eliminating the need to take impressions of the patient. In this way, the procedure is reduced by several laboratory steps, and the results are much more accurate compared to the dentist's analysis and perception. The movement of teeth in the jaw and the final results of treatment can also be predicted using AI algorithms and statistical analyzes based on learning (ML), especially deep learning.

5. Conclusions and Further Research

AI is a research area in dentistry that is in intensive development, especially in clinical practice. Thanks to the current state of development of machine learning and BDA techniques, all the conditions have been met to translate the clinical procedures of the best practice of diagnosis and treatment of patients into virtual dental assistants, as the first stage of the development and application of smart dentistry. On the other hand, the aforementioned facts create conditions for the development and application of telestomatology or Dentistry 4.0.

6. References

- [1] Toosi, Amirhosein & Bottino, Andrea & Saboury, Babak & Siegel, Eliot & Rahmim, Arman. (2021). A Brief History of AI: How to Prevent Another Winter (A Critical Review). *PET Clinics*. 16. 10.1016/j.cpet.2021.07.001.
- [2] Negnevitsky M. Artificial intelligence: a guide to intelligent systems. 2nd^{ed}. Harlow: Addison–Wesley; in 2005
- [3] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943; 5(4):115–33.
- [4] Turing AM. Computing machinery and intelligence. *Mind* 1950; LIX (236):433–60.
- [5] Minsky ML. Theory of neural-analog reinforcement systems and its application to the brain model problem [dissertation]. Princeton: Princeton University; in 1954
- [6] Brook R. The relationship between matter and life. *Nature* 2001; 409:409–11.
- [7] Newell A, Simon HA. GPS, a program that simulates human thought. In: Billing H, editor. *Lernende automaten*. Munich: R. Oldenbourg; 1961 BC 109–24.
- [8] Zadeh LA. Fuzzy sets. *Inf Control* 1965;8(3):338–53.
- [9] Buchanan B, Sutherland G, Feigenbaum EA. Heuristic Dendral: a program for generating explanatory hypotheses in organic chemistry. In: Meltzer B, Michie D, Swann M, editors. *Machine intelligence 4*. Edinburgh: Edinburgh University Press; 1969 BC 209–54.
- [10] Bryson Jr AE, Ho YC. *Applied optimal control: optimization, estimation, and control*. Waltham: Blaisdell Publishing Company; in 1969
- [11] Holland JH. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press; in 1975
- [12] Feigenbaum, Edward & Buchanan, Bruce & Lederberg, Joshua. (1970). On generality and problem solving: A case study using the Dendral program. *Machine Intelligence*. 6.

- [13] Shortliffe E, editor. Computer-based medical consultations: Mycin. New York: Elsevier; in 1976
- [14] Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 1982; 79(8):2554–8.
- [15] Rumelhart DE, McClelland JL; PDP Research Group. Parallel distributed processing: explorations in the microstructures of cognition. Cambridge: MIT Press; in 1986
- [16] Koza JR. Genetic programming: on the programming of the computers by means of natural selection. Cambridge: MIT Press; in 1992
- [17] Wooldridge, M., & Jennings, NR (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(02), 115. 10.1017/s0269888900008122.
- [18] Deep Blue [Internet]. Armonk: IBM [cited 2024 May 15]. Available at: <https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>.
- [19] AlphaGo Korea [Internet]. London: DeepMind Technologies Limited; 2019 [cited 2024, May 15]. Available at: <https://deepmind.com/research/alphago/alphago-korea/>.
- [20] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of Go without human knowledge. *Nature* 2017; 550 (7676):354–9.
- [21] Noguchi, T., Hashizume, Y., Moriyama, H., Gauthier, L., Ishikawa, Y., Matsuno, T., & Sukanuma, A. (2018). A practical use of expert system "AI-Q" focused on creating training data. 2018 5th International Conference on Business and Industrial Research (ICBIR). 10.1109/icbir.2018.8391169.
- [22] Khanna, Sunali & Dhaimade, Pritha. (2017). Artificial Intelligence: Transforming Dentistry Today. *Indian Journal of Basic and Applied Medical Research*; June 2017: Vol.-6, Issue-3, P. 161-167. [cited 2024, May 15]. Available at: <https://www.ijbamr.com/assets/images/issues/pdf/June%202017%20161-167.pdf.pdf>.
- [23] Alexander, Bijo & John, Sunil. (2018). Artificial Intelligence in Dentistry: Current Concepts and a Peep Into the Future. *International Journal of Advanced Research*. 6. 1105-1108. 10.21474/IJAR01/8242.
- [24] Kalappanavar, Anupama & S., Sneha & Annigeri, Rajeshwari. (2018). Artificial intelligence: A dentist's perspective. *Journal of Medicine, Radiology, Pathology and Surgery*. 5. 2-4. 10.15713/ins.jmrps.123.
- [25] Schwendicke, Falk & Samek, W. & Krois, Joachim. (2020). Artificial Intelligence in Dentistry: Chances and Challenges. *Journal of Dental Research*. 99. 002203452091571. 10.1177/0022034520915714.

- [26] Chen, Yo-Wei & Stanley, Kyle & Att, Wael. (2020). Corrigendum: Artificial intelligence in dentistry: current applications and future perspectives. *Quintessence international* (Berlin, Germany : 1985). 51. 430. 10.3290/j.qi.a44465.
- [27] Bindushree, V; Sameen, RJ; Vasudevan, Vijeev; Shrihari, TG; Devaraju, D; Mathew, Nimi Susan. Artificial Intelligence: In Modern Dentistry. *Journal of Dental Research and Review* 7(1): p 27-31, Jan–Mar 2020. 10.4103/jdrr.jdrr_2_20.
- [28] Tandon D, Rajawat J. Present and future of artificial intelligence in dentistry. *J Oral Biol Craniofac Res.* 2020 Oct-Dec; 10(4):391-396. 10.1016/j.jobcr.2020.07.015.
- [29] Khanagar SB et al., Developments, application, and performance of artificial intelligence in dentistry - A systematic review, *Journal of Dental Sciences*, Volume 16, Issue 1 , January 2021, Pages 508-522, 10.1016/j.jds.2020.06.019.
- [30] Shan T, Tay FR, Gu L. Application of Artificial Intelligence in Dentistry. *J Dent Res.* 2021 Mar; 100(3):232-244. 10.1177/0022034520969115.
- [31] Ahmed N, Abbasi MS, Zuberi F, Qamar W, Halim MSB, Maqsood A, Alam MK. Artificial Intelligence Techniques: Analysis, Application, and Outcome in Dentistry - A Systematic Review. *Biomed Res Int.* 2021 Jun 22; 2021:9751564. 10.1155/2021/9751564.
- [32] Banerjee M. Artificial Intelligence in Dentistry: A Ray of Hope. *CODS J Dent* 2021; 13 (2):58-60. 10.5005/jp-journals-10063-0121.
- [33] Carrillo-Perez F, Pecho OE, Morales JC, Paravina RD, Della Bona A, Ghinea R, Pulgar R, Pérez MDM, Herrera LJ. Applications of artificial intelligence in dentistry: A comprehensive review. *J Esthet Restor Dent.* 2022 Jan; 34(1):259-280. 10.1111/jerd.12844.
- [34] Babu, Achsha & Onesimu, Andrew & Sagayam, Martin. (2021). Artificial Intelligence in dentistry: Concepts, Applications and Research Challenges. *E3S Web of Conferences.* 297. 01074. 10.1051/e3sconf/202129701074.
- [35] Nguyen, Thomas & Larrivee, Naomie & Lee, Alicia & Bilaniuk, Olexa & Durand, Robert. (2021). Use of Artificial Intelligence in Dentistry: Current Clinical Trends and Research Advances. *Journal of the Canadian Dental Association.* 87. 17. 10.12816/0059360.
- [36] Swapna, Lingam & Koppolu, Pradeep & Akhter, Fatema & Afroz, Mohammed & Tabassum, Nafeesa & Arshed, Maheen & Khan, Tahseen & Elhaddad, Sally. (2022). Future trends of artificial intelligence in dentistry. *Journal of Nature and Science of Medicine.* 5. 221. 10.4103/jnsm.jnsm_2_22.
- [37] Huang YK, Hsu LP, Chang YC. Artificial intelligence in clinical dentistry: The potentially negative impacts and future actions. *J Dent Sci.* 2022 Oct; 17(4):1817-1818. 10.1016/j.jds.2022.07.013.

- [38] Meghil MM, Rajpurohit P, Awad ME, McKee J, Shahoumi LA, Ghaly M. Artificial intelligence in dentistry. *Dentistry Review*, 2 2002. 1016/j.dentre.2021.100009.
- [39] Agrawal P, Nikhade P. Artificial Intelligence in Dentistry: Past, Present, and Future. *Cureus*. 2022 July 28; 14(7):e27405. 10.7759/cureus.27405.
- [40] Ossowska A, Kusiak A, Świetlik D. Artificial Intelligence in Dentistry - Narrative Review. *Int J Environ Res Public Health*. 2022 Mar 15;19(6):3449. 10.3390/ijerph19063449.
- [41] Lamba, GS, Singh, H., Grover, S., Oberoi, SS, Atri, M., Yadav, P., & Thakral, P. (2022). Artificial intelligence in modern dentistry. *International Journal of Health Sciences* , 6 (S3), 8086–8098. 10.53730/ijhs.v6nS3.7930.
- [42] Surlari, Z.; Budala, DG; Lupu, CI; Stelea, Montenegro; Butnar, OM; Luchian, I. Current Progress and Challenges of Using Artificial Intelligence in Clinical Dentistry—A Narrative Review. *J. Clin. Med*. 2023 , 12, 7378. 10.3390/jcm12237378.
- [43] Dhopte A, Bagde H. Smart Smile: Revolutionizing Dentistry With Artificial Intelligence. *Cureus*. 2023 Jun 30;15(6):e41227. 10.7759/cureus.41227.
- [44] Ghaffari, Maryam & Zhu, Yi & Shrestha, Annie. (2024). A review of advances in artificial intelligence in dentistry. *Dentistry Review*. 4. 100081. 10.1016/j.dentre.2024.100081.
- [45] Chakravorty S, Aulakh BK, Shil M, Nepale M, Puthenkandathil R, Syed W. Role of Artificial Intelligence (AI) in Dentistry: A Literature Review. *J Pharm Bioallied Sci*. 2024 Feb; 16 (Suppl 1): S14-S16. 10.4103/jpbs.jpbs_466_23.
- [46] Aschheim, Kenneth W. DDS (2024) "Unlocking the Future of Dentistry: A Summary of AI and Augmented Intelligence in Dentistry from the ADA's White Paper," *The New York State Dental Journal* : Vol. 90: No. 1, Article 6. [cited 2024, May 15]. Available at: <https://commons.ada.org/nysdj/vol90/iss1/6>.
- [47] Umer, Fahad & Adnan, Niha. (2024). Generative artificial intelligence: synthetic datasets in dentistry. *BDJ Open*. 10. 10.1038/s41405-024-00198-4.
- [48] Park WJ, Park JB. History and application of artificial neural networks in dentistry. *Eur J Dent* 2018;12:594-601. 10.4103/ejd.ejd_325_18.
- [49] Schwendicke F, Golla T, Dreher M, Krois J, Convolutional neural networks for dental image diagnostics: A scoping review, *Journal of Dentistry*(2019), <https://doi.org/10.1016/j.jdent.2019.103226>.
- [50] Hatvani, Janka and Horvath, Andras and Michetti, Jérôme and Basarab, Adrian and Kouamé, Denis and Gyöngy, Miklos Deep Learning-Based Super-Resolution Applied to Dental Computed Tomography. (2019) *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3 (2). 120-128. <https://doi.org/10.1109/TRPMS.2018.2827239>.

- [51] Leite AF, Vasconcelos KF, Willems H, Jacobs R. Radiomics and Machine Learning in Oral Healthcare. *Proteomics Clin Appl*. 2020 May;14(3):e1900040. 10.1002/prca.201900040.
- [52] Hung K, Yeung AWK, Tanaka R, Bornstein MM. Current Applications, Opportunities, and Limitations of AI for 3D Imaging in Dental Research and Practice. *Int J Environ Res Public Health*. 2020 June 19; 17(12):4424. 10.3390/ijerph17124424.
- [53] Machoy ME, Szyszka-Sommerfeld L, Vegh A, Gedrange T, Woźniak K. The ways of using machine learning in dentistry. *Adv Clin Exp Med*. 2020 Mar;29(3):375-384. 10.17219/acem/115083.
- [54] Mao-Lei Sun, Yun Liu, Guomin Liu, Dan Cui, Ali Asghar Heidari, Wen-Yuan Jia, Xuan Ji, Huiling Chen, Yungang Luo, Application of Machine Learning to Stomatology: A Comprehensive Review, [IEEE Access](https://doi.org/10.1109/access.2020.3028600). 2020.10.1109/access.2020.3028600.
- [55] Stefano Corbella, Shanmukh Srinivas, Federico Cabitza, Applications of Deep Learning In Dentistry, *Oral Surg Oral Med Oral Pathol Oral Radiol* (2020), <https://doi.org/10.1016/J.oooo.2020.11.003>.
- [56] Grischke, Jasmin & Johannsmeier, Lars & Eich, Lukas & Griga, Leif & Haddadin, Sami. (2020). Dentronics: Towards robotics and artificial intelligence in dentistry. *Dental Materials*. 36 (2020), 765 - 778. 10.1016/j.dental.2020.03.021.
- [57] Yoda T, Bornstein MM, Jung RE, Ferrari M, Waltimo T, Zitzmann NU. Recent Trends and Future Direction of Dental Research in the Digital Era. *Int J Environ Res Public Health*. 2020 Mar 18; 17(6):1987. 10.3390/ijerph17061987.
- [58] Lahoud P, EzEldeen M, Beznik T, Willems H, Leite A, Van Gerven A, Jacobs R. Artificial Intelligence for Fast and Accurate 3-Dimensional Tooth Segmentation on Cone-beam Computed Tomography. *J Endod*. 2021 May;47(5):827-835. 10.1016/j.joen.2020.12.020.
- [59] Lee, SJ; Chung, D.; Asano, A.; Sasaki, D.; Maeno, M.; Ishida, Y.; Kobayashi, T.; Kuwajima, Y.; Da Silva, JD; Nagai, S. Diagnosis of Tooth Prognosis Using Artificial Intelligence. *Diagnostics* 2022 , 12, 1422. <https://doi.org/10.3390/diagnostics12061422>.
- [60] Shafí, I.; Fatima, A.; Afzal, H.; Díez, IdIT; Lipari, V.; Breñosa, J.; Ashraf, I. A Comprehensive Review of Recent Advances in Artificial Intelligence for Dentistry E-Health. *Diagnostics* 2023 , 13, 2196. <https://doi.org/10.3390/diagnostics13132196>.
- [61] Naeimi, SM; Darvish, S.; Salman, BN; Luchian, I. Artificial Intelligence in Adult and Pediatric Dentistry: A Narrative Review. *Bioengineering* 2024 , 11, 431. <https://doi.org/10.3390/bioengineering11050431>.

- [62] Mohammad Khursheed Alam, Sultan Abdulkareem Ali Alftaikhah, Rakhi Issrani, Vincenzo Ronsivalle, Antonino Lo Giudice, Marco Ciccì, Giuseppe Minervini, Applications of artificial intelligence in the utilization of imaging modalities in dentistry: A systematic review and meta-analysis of in -vitro studies, *Heliyon*, Volume 10, Issue 3, 2024, <https://doi.org/10.1016/j.heliyon.2024.e24221>.
- [63] Banerjee M. Artificial Intelligence in Dentistry: A Ray of Hope. *CODS J Dent* 2021; 13(2):58–60. 10.5005/jp-journals-10063-0121.

Predicting Behavioral Patterns in Brazillian Whiteknee Tarantulas - A Multi-Output Machine Learning Analysis of Dietary Effects in *Acanthoscurria Gemiculata*

Amina Radončić^{*1}

Abstract: Machine learning techniques have the potential to act as powerful tools for predicting behavioral patterns in tarantulas. For this particular observation, Brazilian Whiteknee Tarantulas (*Acanthoscurria gemiculata*) are placed in a focal point. The study leverages a multi-output machine learning approach implemented in MATLAB to analyze the impact of dietary compositions—Worms, Crickets, Roaches, Locusts, and Mixed Diet—on tarantula behavior. Based on the collected data from the survey involving 47 owners of the same species, the data on gender, dietary preferences, and activity levels of these tarantulas has been obtained. The findings highlight significant variations in predictive accuracy across different diets, with the mixed diet demonstrating the highest accuracy (0.85714) and the diet consisted only of worms showing robust performance (accuracy: 0.71429, AUC: 0.70833). Despite the limited dataset, this research underscores the potential of machine learning to elucidate the intricate relationship between diet and behavior in tarantulas, offering insights relevant to both ecological studies and captive management practices.

Keywords: machine learning, decision tree model, multi-output model, data analysis, tarantula, brazillian whiteknee

1. Introduction

Every moment in life presents a multitude of choices, each offering numerous possibilities to explore. **Decision making**, therefore, involves the cognitive process of selecting the most suitable course of action among available options. It requires **a thoughtful assessment of the advantages and disadvantages of each alternative while considering all feasible choices**. This intricate process is a fundamental challenge faced by individuals and organizations alike. To address this challenge effectively, **algorithms such as decision trees** provide a structured approach, ensuring comprehensive decision analysis by meticulously weighing essential attributes without overlooking any critical factors [1]. Decision trees serve as indispensable tools in decision making, widely utilized across various domains. These graphical representations map decisions and potential outcomes, encompassing factors like resource allocation, event

^{*1}International Burch University, Natural and Medical Sciences, Sarajevo, Bosnia and Herzegovina
E-mail: amina.radoncic@stu.ibu.edu.ba

probabilities, and utility assessments [1, 2]. Classification trees predict outcomes based on data insights, prevalent in fields like probability and statistics. Regression trees, conversely, determine outcomes from quantitative data, often applied in real estate evaluations [1].

Decision tree forests aggregate diverse models to enhance predictive accuracy, while classification and regression trees combine attributes for logical outcome predictions [1, 2]. Despite varying in accuracy, decision trees, including K Means Clustering, integrate factors for comprehensive decision support across diverse applications: from text classification to medical diagnostics and financial forecasting, and in this case – **behavioral patterns in tarantulas**.

2. The Mysterious Lives of Tarantulas

Current research on tarantula behavioral patterns, particularly regarding the influence of nutrition, is notably sparse in utilizing advanced analytical methods like machine learning and artificial intelligence. Existing studies predominantly focus on ecological observations and basic behavioral assessments.

Taxonomically observed, **tarantulas** belong to the infraorder Mygalomorphae as **the largest arachnids**. This infraorder comprises approximately 3000 species of spiders distributed globally across 16 families, including trap-door spiders, and sheet-web spiders. These spiders are known for their large size and long lifespan, as well as their diverse silk-producing capabilities [3].

Mygalomorphs are considered an ancient lineage, distinct from the more diverse Araneomorphae, and retain several primitive spider characteristics such as two pairs of book lungs, paraxial chelicerae, simple silk spigots, and comparatively weaker silk [3].



Figure 1. The Brazillian Whiteknee tarantula in captivity

Despite these primitive traits, specific anatomical features support the monophyly of Mygalomorphae, including labial and maxillary cuspules, fewer palpal bulb sclerites, sternal sigillae, segmented posterior lateral spinnerets, reduced anterior lateral spinnerets, and the absence of anterior median spinnerets [3].

The crown jewel of this study is *Acanthoscurria gemiculata*, the **Brazilian Whiteknee tarantula**, one of the largest spiders in the world.

These tarantulas are predominantly deep black in body and leg coloration, with occasional lighter shades on their legs and distinctive white bands that enhance their striking appearance. Notably, females of this species are larger and more intensely colored compared to the smaller, less vividly hued males [4]. In terms of size, Brazilian whiteknee tarantulas surpass the average tarantula, with body lengths reaching up to 9 centimeters (3.5 inches). Known for their rapid growth, females typically achieve a mature leg span of 8.5 inches within 3 to 4 years. These tarantulas are popular as pets due to their impressive size, robust nature, and captivating color patterns. However, they exhibit defensive behavior and possess urticating hairs that can cause skin irritation [4].

While the Brazilian whiteknee tarantula can bite if provoked, it typically deploys urticating hairs as a primary defense mechanism. These hairs deter potential threats by causing discomfort and irritation, minimizing the need for biting. Although their venom is generally not medically significant, their large size can potentially cause mechanical harm in the event of a bite [4].

Are their behavioral patterns and temper dictated by the dietary habits? This study aims to fill this research gap and answer this simple, yet complex question, by employing machine learning algorithms to predict tarantula behavior based on dietary factors. By integrating computational approaches with behavioral ecology, this research seeks to contribute novel insights and potentially enhance strategies for managing tarantula care in captivity and conservation efforts in natural habitats.

3. Materials and Methods

This study involved a survey administered anonymously to 47 Brazilian Whiteknee Tarantula owners. Participants provided data on their tarantulas' age categories (3+, 2+, 1+, <1 years), gender (scored with 2 points for females, 1 point for males, and 0 points for unknown gender), activity levels categorized as LOW, NORMAL, or HIGH, and temperament described as DOCILE, NERVOUS, FLICKER (indicating hair flicking behavior), or DEFENSIVE (characterized by defensive postures). Participants also documented their tarantulas' dietary habits, recording consumption of Worms, Crickets, Roaches, Locusts, or a mixed diet using binary scoring (1 for consumption, 0 for non-consumption).

Table 1. Survey questions and criteria

Category	Description	Scoring
Age	to 20	-
Gender	20 – 200	2 points for Female, 1 point for Male, 0 points for Unknown gender
Activity	200 – 500	-
Temper	LOW, NORMAL, HIGH	-
Diet: Worms	Consumption (1) or Non-consumption (0)	Binary
Diet: Crickets	Consumption (1) or Non-consumption (0)	Binary
Diet: Roaches	Consumption (1) or Non-consumption (0)	Binary
Diet: Locusts	Consumption (1) or Non-consumption (0)	Binary
Diet: Mixed	Consumption (1) or Non-consumption (0)	Binary

The obtained results were processed using MATLAB to investigate the correlation between gender, temperament, and dietary preferences of these tarantulas. Initially, categorical variables such as gender (Male, Female, Unknown) and temperament (DOCILE, NERVOUS, DEFENSIVE, FLICKER) were converted into numeric formats suitable for machine learning analysis. The dataset was then divided into training (70%) and testing (30%) sets using a holdout method to ensure robust model validation. For each dietary category (Worms, Crickets, Roaches, Locusts, Mixed), individual decision tree models were trained using the training data. Model performance metrics including accuracy, root mean squared error (RMSE), and area under the ROC curve (AUC) were computed to assess the predictive capability of the models. Visualizations such as confusion matrices, ROC curves, feature importance analyses, and predictive distribution plots were generated to interpret and validate the model's predictions. This approach enabled the identification of nuanced behavioral patterns and dietary preferences among *A. geniculata* tarantulas, underscoring the effectiveness of decision tree models in analyzing complex behavioral datasets in animal studies.

4. Results and Discussion

4.1. Model Analysis

The accuracy of each decision tree model was evaluated using the test data. Accuracy measures the proportion of correctly predicted outcomes relative to the total number of predictions made. For instance, the model achieved an accuracy of 71.43% for Diet_Worms, indicating that 71.43% of predictions regarding whether a tarantula consumes worms were correct.

RMSE quantifies the average magnitude of the errors between predicted and actual values across all dietary categories. Lower RMSE values indicate better model performance in predicting dietary preferences accurately. For example, the RMSE for Diet_Mixed was 0.37796, suggesting that the model's predictions for tarantulas with mixed diets were generally close to the actual values.

The ROC curve and AUC were used to assess the model's ability to distinguish between positive and negative classes for each dietary category. AUC values close to 1 indicate excellent model performance, while values around 0.5 suggest a model that performs no better than random guessing. The AUC values varied across dietary categories, with Diet_Locust having an AUC of 0.5, indicating that the model's predictive ability for this category was not significantly better than chance. Feature importance analysis was conducted to identify which predictors (gender and temperament) contributed most significantly to the models' predictions. This analysis helps in understanding which variables influence tarantula dietary preferences the most. For instance, the analysis might reveal that female tarantulas and those categorized as DEFENSIVE or NERVOUS tend to prefer specific types of food more than others.

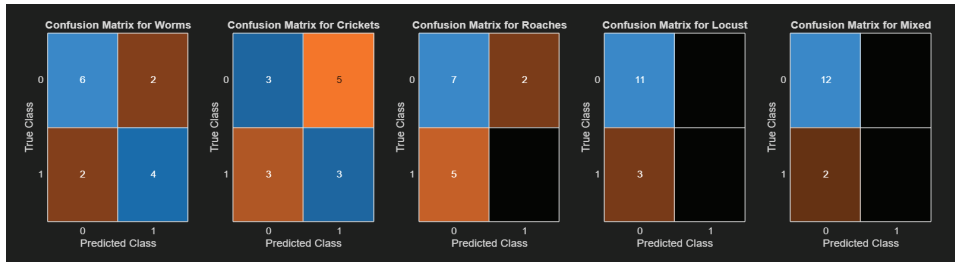


Figure 2. Confusion matrices for each dietary category

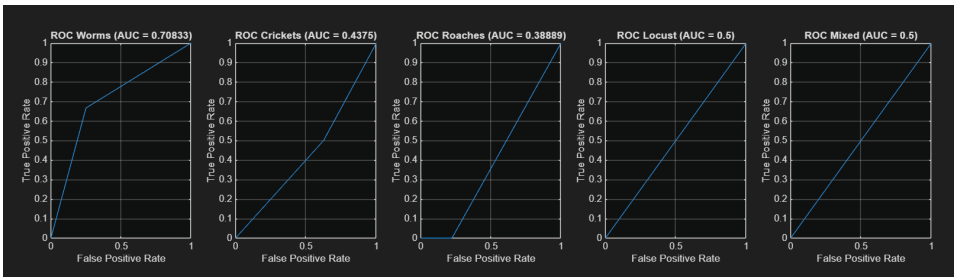


Figure 3. ROC curve for each dietary category, showing false positive and true positive rates

Visualizations (figures above) such as confusion matrices, ROC curves, boxplots, and histograms were generated to provide a comprehensive understanding of model performance and prediction outcomes. Confusion matrices visually represented the distribution of true positive, false positive, true negative, and false negative predictions for each dietary category. ROC curves illustrated the trade-off between sensitivity and specificity, with AUC values quantifying the models' discriminative ability.

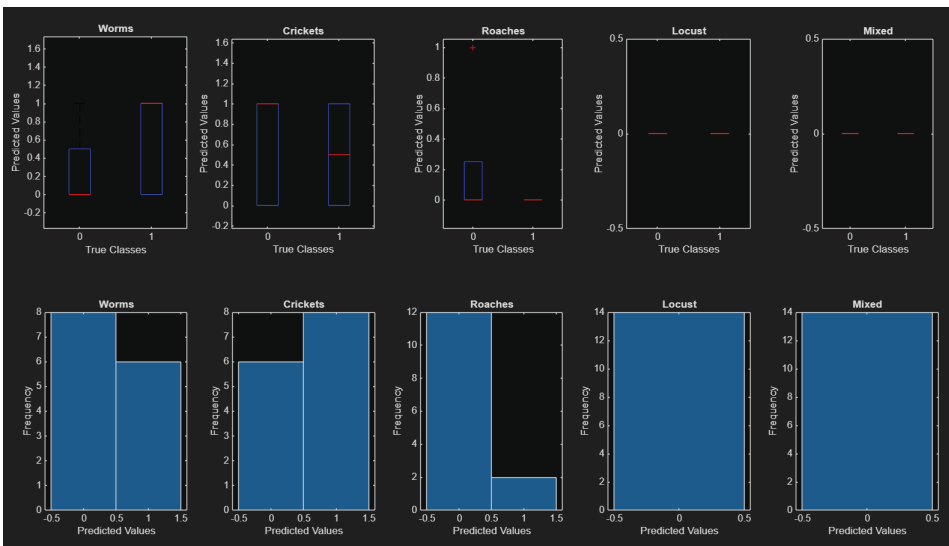


Figure 4. Boxplots (top) and histograms (bottom) portraying the prediction outcomes for dietary categories

5. Conclusion

Through machine learning methodologies and examinations, this study aimed to investigate the interplay between gender, temperament, and dietary preferences in Brazilian Whiteknee tarantulas. Based on the analysis of anonymized survey data from 47 tarantula owners, decision tree models were constructed to predict the dietary habits of these spiders, encompassing preferences for worms, crickets, roaches, locusts, and mixed diets. Despite the limited dataset, the findings revealed nuanced associations between gender, temperament traits, and dietary behaviors, with varying accuracies, root mean squared errors (RMSE), and area under the ROC curve (AUC) values observed across different dietary categories.

Gender and temperament emerged as pivotal predictors influencing dietary preferences, offering insights into tarantula management and conservation practices. Despite the robustness of the predictive models in capturing these relationships, limitations such as the modest sample size and potential biases inherent in survey data warrant cautious interpretation of the results.

Future investigations could expand upon this study by incorporating larger datasets, integrating environmental variables, and extending comparative analyses to include other species within the Theraphosidae family. Continued advancements in machine learning techniques and data collection methodologies hold promise for refining predictive models and deepening our understanding of animal behavior across diverse ecological contexts.

This one-of-a-kind research contributes to advancing knowledge in animal behavior research and underscores the potential of machine learning in informing conservation strategies and management practices for tarantula species. By elucidating the complex relationships between behavioral traits and dietary preferences, this research enhances our ability to implement targeted conservation efforts aimed at safeguarding tarantula populations bred in captivity.

6. References

- [1] Navada, A., Ansari, A. N., Patil, S., & onkamble, B. A. (2011). Overview of use of decision tree algorithms in machine learning. 2011 IEEE Control and System Graduate Research Colloquium. doi:10.1109/icsgrc.2011.5991826
- [2] M. Somvanshi, P. Chavan, S. Tambade and S. V. Shinde, "A review of machine learning techniques using decision tree and support vector machine," *2016 International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, 2016, pp. 1-7, doi: 10.1109/ICCUBEA.2016.7860040.

- [3] Pérez-Miles, F., & Perafán, C. (2017). *Behavior and Biology of Mygalomorphae. Behaviour and Ecology of Spiders*, 29–54. doi:10.1007/978-3-319-65717-2_2
- [4] Macedo KWR, Costa LJL, de Souza JO, de Vasconcelos IA, de Castro JS, de Santana CJC, Magalhães ACM, Castro MS, Pires OR. Brazilian Theraphosidae: a toxicological point of view. *J Venom Anim Toxins Incl Trop Dis*. 2021 Nov 22;27:e20210004. doi: 10.1590/1678-9199-JVATITD-2021-0004. PMID: 34868282; PMCID: PMC8610171
- [5] Malti Bansal, Apoorva Goyal, Apoorva Choudhary, A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning, *Decision Analytics Journal*, Volume 3, 2022, 100071, ISSN 2772-6622, <https://doi.org/10.1016/j.dajour.2022.100071>.
- [6] De Ville, B. (2013). Decision trees. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(6), 448-455.
- [7] Ciaburro, G. (2017). *MATLAB for machine learning*. Packt Publishing Ltd.
- [8] Suthaharan, S., & Suthaharan, S. (2016). Decision tree learning. *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, 237-269.
- [9] Ferri, C., Flach, P., & Hernández-Orallo, J. (2002, July). Learning decision tree using the area under the ROC curve. In *ICML* (Vol. 2, pp. 139-146).
- [10] Ling, C. X., Huang, J., & Zhang, H. (2003). AUC: a better measure than accuracy in comparing learning algorithms. In *Advances in Artificial Intelligence: 16th Conference of the Canadian Society for Computational Studies of Intelligence, AI 2003, Halifax, Canada, June 11–13, 2003, Proceedings 16* (pp. 329-341). Springer Berlin Heidelberg.
- [11] Paluszczek, M., & Thomas, S. (2019). *MATLAB machine learning recipes: a problem-solution approach* (pp. 1-347). Berkeley: Apress.
- [12] Escoubas, P., & Rash, L. (2004). Tarantulas: eight-legged pharmacists and combinatorial chemists. *Toxicon*, 43(5), 555-574.
- [13] Pechmann, M. (2020). Embryonic development and secondary axis induction in the Brazilian whiteknee tarantula *Acanthoscurriageniculata*, CL Koch, 1841 (Araneae; Mygalomorphae; Theraphosidae). *Development genes and evolution*, 230(2), 75-94.
- [14] Paluszczek, M., Thomas, S., Paluszczek, M., & Thomas, S. (2020). MATLAB machine learning toolboxes. *Practical MATLAB deep learning: A project-based approach*, 25-41.

Assessing the Accuracy of Logistic Regression and Bert in Sentiment Analysis and Mental Disorder Classification

Faruk Bećirović^{*1}, Nejra Merdović¹, Madžida Hundur Hiyari¹,
Merima Smajlhodžić-Deljo¹, Irma Salkić², Adna Softić¹

Abstract: Sentiment analysis leverages machine learning and natural language processing techniques to classify and interpret textual data, identifying sentiments as positive, negative, or neutral. This study explores sentiment analysis in the context of mental health, utilising two models: Logistic Regression and Bidirectional Encoder Representations from Transformers (BERT). The dataset comprises 52 680 unique statements associated with seven mental health statuses, including depression, anxiety and suicidal tendencies. Logistic Regression achieved an accuracy of 72%, while BERT, with its advanced deep learning architecture, demonstrated a significant improvement with an accuracy of 84%. BERT's superior performance is attributed to its bidirectional contextual understanding and attention mechanisms, enhancing its ability to handle complex and nuanced textual information. This study highlights the efficacy of BERT over traditional models in analysing and classifying sentiments related to mental health, underscoring its potential for improving early detection and intervention.

Keywords: Sentiment Analysis, Machine Learning, Natural Language Processing, Deep Learning, Mental Health, Contextual Understanding

1. Introduction

Mental disorders are conditions that affect cognition, emotion, and behavioural control, impairing children's ability to learn and adults' ability to function in their families, businesses, and society as a whole. These medical conditions frequently manifest early in childhood and have a chronic and recurring pattern. Mental disorders contribute significantly to the disease burden due to their high prevalence, early onset, persistence, and limitations [1]. In developed countries, anxiety disorders are the most prevalent mental health disorders. They include concerns that are related to structural disorders, neurotic fears, and legitimate actual fears [2]. Fear is an immediate alarm response to current or imminent risk, actual or perceived, whereas anxiety is a mood state focused on future-oriented concerns, planning for possible unpleasant outcomes [3]. Approximately 3% of the

^{*1}Verlab Research Institute for Biomedical Engineering, Medical Devices and Artificial Intelligence, Sarajevo, Bosnia and Herzegovina

²Faculty of Electrical Engineering, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
Email: faruk@verlabinstitute.com

worldwide population suffers from bipolar disorders, a severe category of mood disorders marked by the occurrence of both manic or hypomanic episodes and depressive episodes. Although there is a strong family component to many disorders, environmental variables are also quite important [4]. A diverse range of disorders with differing symptoms and probably distinct origins are together referred to as depression. Depressed mood, lack of enjoyment, difficulty paying attention, low energy, disturbed sleep or nutrition, feelings of guilt or worthlessness, and suicidal thoughts are common symptoms. Psychosocial complaints frequently include personality disorders (PD), especially in young adults. PD is associated with low self-esteem, diminished abilities to work and love, and recurrent stressful reactions that drive medical attention [5]. Suicidal ideas, intentions, and attempts are extremely upsetting and are frequently connected to severe sadness, pessimistic future outlooks, and a mental health issue diagnosis [6]. These conditions are all categorised as mental disorders.

Sentiment analysis is a technique that uses machine learning, natural language processing, and computational linguistics to automatically determine the sentiment expressed in text [7], [8]. It typically categorizes opinions as positive, negative, or neutral and can be applied at sentence, document, or feature levels [8]. Common applications include analyzing product reviews, social media content, and public perception of events [7]–[9]. The process involves extracting information from sentiment-laden words, their context, and linguistic structure [10]. Two main approaches are machine learning and lexicon-based methods [11]. Sentiment analysis aids organizations in decision-making and product improvement [11], [12]. However, challenges persist, such as domain adaptability, accuracy with limited labeled data, and handling complex sentences [8]. Despite these issues, sentiment analysis remains a valuable tool for mining opinions and emotions across various domains [7].

Natural Language Processing (NLP) is a field of computer science and artificial intelligence focused on enabling computers to understand, interpret, and generate human language [9], [13]. It is crucial for bridging the gap between human communication and machine comprehension, with applications in various domains such as machine translation, text summarising and speech recognition [9], [14]. NLP has evolved from rule-based approaches to advanced techniques like machine learning and deep learning [14]. Its importance is evident in healthcare, where it helps transform unstructured text into structured data for improving patient care and advancing medicine [10]. Despite challenges like understanding context and cultural nuances, NLP continues to transform human-computer interactions and find applications in diverse fields, including finance, education, and customer service [15]. Ongoing research aims to address obstacles and advance the field further [10], [16].

The complex clinical presentations and overlap of symptoms make diagnosing mental disorders difficult. Misdiagnosis can occur when distinct conditions

share similar symptoms, and the procedure is made more challenging by knowledge gaps. Therefore, in order to enhance the diagnosis process, cutting edge technologies like artificial intelligence are making their way into contemporary healthcare. Intelligent systems are able to accurately detect mental disorders, which helps with diagnosis, treatment, and patient recovery. As AI technology develops, it will assist healthcare professionals in early illness detection and customised treatment plans based on patient characteristics.

The goal of this paper is to examine how advanced Neural Networks adapted specifically for Natural Language Processing are effective in classifying mental disorders from patient complaints. The BERT model is compared with a traditional Logistic Regression model which is based on statistical methods.

2. Methods and Materials

The data used was compiled and curated from various sources. The dataset is an amalgamation of raw data from various openly accessible sources, cleaned and compiled and can thus be used as a resource for sentiment analysis and other Natural Language Processing (NLP) tasks. The dataset is created by amalgamating data from different sources available on Kaggle, a platform for the exchange of datasets and other related resources. The entire dataset is also available on Kaggle.

The dataset consists of two variables, a statement and status, which is a target variable. The statement is a textual string consisting of a variable length statement. The status is mental disorder or condition that the patient making the statement was diagnosed with. The initial assignment of statuses for a given statement was made by a mental health professional. The dataset is composed of 52680 unique statements each corresponding to one of seven statuses. The complete list of statuses and their prevalence in the dataset is shown in Table 1.

Table 1: Statuses and their proportion in the dataset

Status	Distribution	Proportion
Normal	16342	31.02%
Depression	15404	29.24%
Suicidal	10652	20.22%
Anxiety	3841	7.29%
Bipolar	2777	5.27%
Stress	2587	4.91%
Personality disorder	1077	2.04%
Total	52680	100%

The distributions of labels is also depicted graphically in Figure 1.

Before the model development the dataset validity was confirmed and it was

ensured that the data within is applicable for the given task. Following this step, the dataset was divided into two subsets, the training set which consisted of 80% of total number of samples and the validation set, which consisted of the remaining 20%.

For classification of the statements two models were developed. First model is Logistic Regression, which is a statistical Machine Learning (ML) model. Second model that was used in this study is Bidirectional Encoder Representations from Transformers (BERT), which is a deep learning model based on transformer architecture.

Logistic Regression is a statistical model commonly used for binary classification tasks.

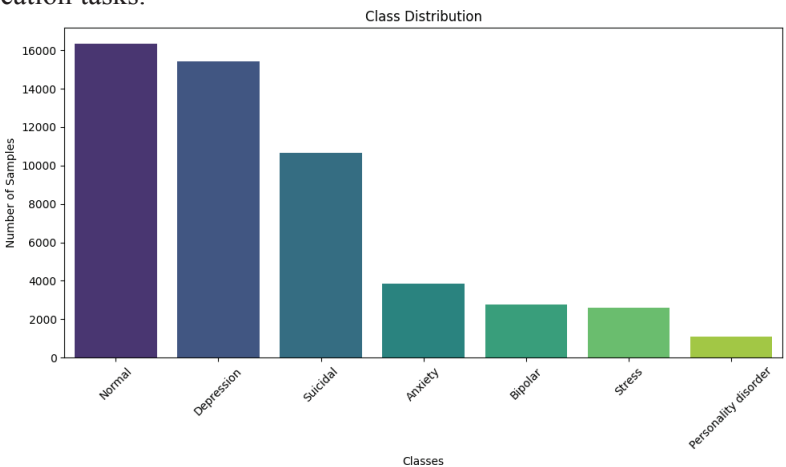


Figure 1: Class distribution within the dataset

It uses a logistic function to model a binary dependent variable. For sentiment analysis, the model can be trained to predict whether a given text expresses a positive or negative sentiment. Logistic Regression models the probability that a given input belongs to a particular class. It starts with a linear combination of input features:

$$z = \beta_0 + \beta_1 X_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (1)$$

Where β_0 is intercept term, $\beta_1, \beta_2, \dots, \beta_n$ are coefficients for features and $x_1, x_2, \dots, x_n$ are features. In this particular case, the features list and the features themselves are extracted using Term Frequency Inverse Document Frequency method from the input text.

To map linear function to probability sigmoid function is used:

$$\sigma(z) = \frac{1}{1 + e^{-x}} \quad (2)$$

The output of the $\sigma(z)$ is a value between 0 and 1 represented the probability a statement is mapped to particular class, in this case a status.

The parameters β are estimated by maximising the log-likelihood function:

$$L(\beta) = \sum_{i=1}^m [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (3)$$

Where m is the number of samples in the training set, y_i is the actual label of i -th sample, and $\hat{y}_i$ is predicted probability of the i -th sample for a given status. The goal is to find β which maximises log-likelihood function, which is preformed using Gradient Descent optimisation algorithm.

BERT is a deep learning model based on the Transformer architecture. Unlike traditional models, BERT reads text bidirectionally, enabling it to understand the context of a word by looking at both its left and right surroundings. It is pre-trained on a large corpus of text and can be fine-tuned on specific tasks. BERT is a deep neural network, consisting of multiple layers of transformers, each having Multi-Head Self-Attention Mechanism which allows the model to focus on different parts of the input sequence at the same time. The transformers in BERT model are specifically designed to understand the context in the natural language. The transformer architecture of BERT model consists of only the encoder-decoder part, meaning that it can process input data and capture contextual information simultaneously.

The use of Transformers in BERT model is significant because it enables the use previously mentioned Self-Attention Mechanism which allows the model to weigh the importance of every word in a sentence compared to all other words. This is achieved using Softmax function which convert vector of raw scores for an input into a probability distribution over classes. The attention score α can be represented as:

$$\alpha(Q, K, V) = \sigma \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

Where Q are queries, a current word that is being evaluated, K are keys, which potential candidates in the sequence that can provide relevant information to the query, V the actual values that need to be aggregated, in other word the status that is associated with the statement, and d_k is the dimensionality of the keys. The σ is Softmax function which is used to convert raw values from queries and keys into probabilities.

Given a vector of scores $z = [z_1, z_2, \dots, z_n]$ Softmax function produces a probability distribution $\sigma(z) = [\sigma_1, \sigma_2, \dots, \sigma_n]$ where:

$$\sigma_i(z_i)=\frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \tag{5}$$

In the Attention mechanism, the softmax function ensures that the sum of attention weights across all positions in the sequence is 1. This normalization allows the model to distribute its focus across different parts of the sequence based on their relevance to the current query.

3. Results and Discussion

BERT represents a significant advancement over traditional models like Logistic Regression for sentiment analysis, particularly in complex and sensitive areas such as mental health. While it offers superior accuracy and contextual understanding, it comes with increased computational demands and less interpretability. Choosing between these models depends on the specific requirements of the task, the availability of computational resources, and the importance of interpretability.

BERT, with its deep architecture and attention mechanism, has a much higher expressive power than Logistic Regression. It can model a wide variety of language phenomena, including word order, syntax, and semantics, which are crucial for accurately understanding and classifying sentiment.

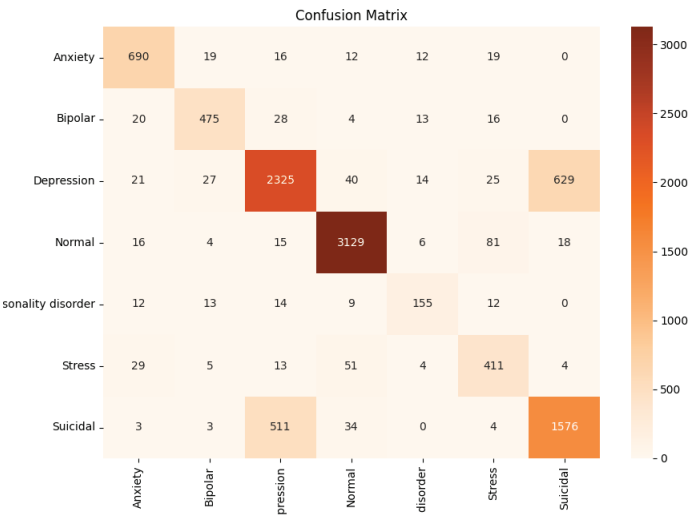


Figure 2: Confusion matrix for BERT model

After training the Logistic Regression model on the training dataset the accuracy of 72% percent was achieved. Using the BERT model on the other hand resulted in the significant improvement in terms of accuracy, which was 84%.

Sentiment analysis, a natural language processing technique, has emerged as a valuable tool for classifying mental disorders, particularly in the context of social media data. Recent studies have explored its application in detecting depression, anxiety, and other mental health issues [17], [18]. By analyzing text from social media posts and online communications, sentiment analysis can provide insights into individuals' emotional states, enabling early detection and intervention for mental health problems [19]. The COVID-19 pandemic has heightened the importance of this

approach, as increased social media usage and isolation have led to a rise in mental health concerns [20]. Researchers have employed various machine learning and deep learning methods, including lexical approaches and Naive Bayes models, to classify sentiments related to mental health [17]. Some studies have achieved promising results, with accuracy rates reaching up to 89% in detecting mental disorders using multi-labeled deep-learning models [18]–[20].

Recent studies have explored the effectiveness of BERT for sentiment analysis.

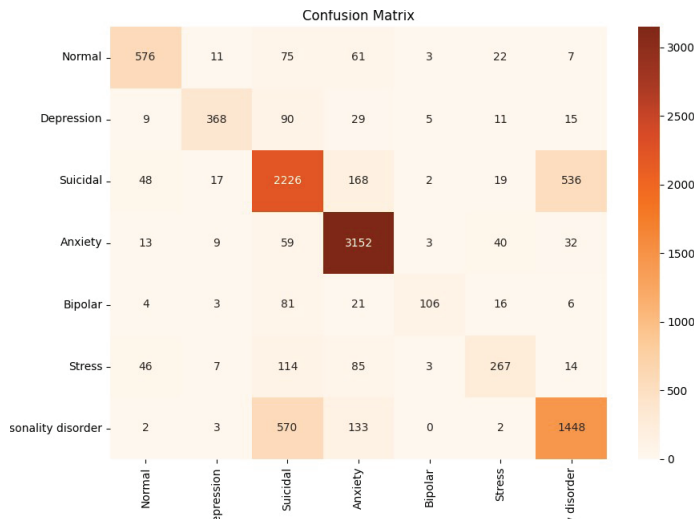


Figure 3: Confusion matrix for Logistic Regression model

Traditional statistical models, including Logistic Regression have been successfully used for sentiment analysis [21]. BERT outperforms traditional models like Naive Bayes in binary and multi-class sentiment classification tasks on Twitter data [22]. When compared to other transformer-based models like DistilBERT and RoBERTa, BERT demonstrates superior performance in analyzing sentiments in movie reviews and tweets [23]. BERT's ability to capture sentimental relations, including word negations and intensifications, makes it more effective than basic machine learning and deep learning models

for classifying user textual data [24]. In a comprehensive comparison of sentiment analysis techniques, including lexicon-based models, logistic regression, LSTM, and BERT, the pre-trained BERT model exhibited undisputed superiority in sentiment classification performance, as measured by accuracy, precision, recall, and F1 score [25].

4. Conclusion

This study highlights the significant advantages of using advanced deep learning models like BERT for sentiment analysis, particularly in the context of mental health. While Logistic Regression provides a solid starting point with an accuracy of 72%, BERT greatly improves performance, achieving an accuracy of 84%. This improvement is due to BERT's advanced features, including its ability to understand text in both directions and its attention mechanisms that capture more nuanced details.

Recent research has shown that BERT outperforms traditional models, such as Logistic Regression and Naive Bayes, in sentiment analysis tasks. For example, BERT excels in classifying sentiments

in social media posts and movie reviews, outperforming other transformer-based models like DistilBERT and RoBERTa. BERT's ability to handle complex sentiment details, like word negations and intensification, makes it more effective than simpler models.

In various comparisons, BERT consistently shows superior performance across important metrics such as accuracy, precision, recall, and F1 score. This makes it a valuable tool for analysing mental health-related text and improving early detection and intervention strategies. Future research should focus on integrating BERT into real-world mental health monitoring systems, adapting it to different languages and datasets, and addressing any challenges related to its computational demands and interpretability.

5. References

- [1] World Health Organization, *Disease control priorities related to mental, neurological, developmental and substance abuse disorders*. Geneva, Switzerland: World Health Organization, Sep. 2006, ISBN: 924156332X.
- [2] R. T. Krüger, *Disorder-specific psychodrama therapy in theory and practice*, en. Cham, Switzerland: Springer Nature, Jan. 2024. DOI: 10.1007/978-981-99-7508-2.
- [3] M. G. Craske, S. L. Rauch, R. Ursano, J. Prenoveau, D. S. Pine and R. E. Zinbarg, 'What is an anxiety disorder?' en, *Focus (Am. Psychiatr. Publ.)*, vol. 9, no. 3, pp. 369–388, Jul. 2011. DOI: 10.1176/foc.9.3.foc369.

- [4] R. S. El-Mallakh and R. H. Belmaker, 'Bipolar disorder,' en, in *Tasman's Psychiatry*, Cham: Springer International Publishing, 2023, pp. 1–54. DOI: 10.1007/978-3-030-42825-9_19-1.
- [5] C. R. Cloninger and D. M. Svrakic, 'Personality disorders,' in *The Medical Basis of Psychiatry*, Totowa, NJ: Humana Press, 2008, pp. 471–483. DOI: 10.1007/978-1-59745-252-6_28.
- [6] J. Johnson, A. M. Wood, P. Gooding, P. J. Taylor and N. Tarrier, 'Resilience to suicidality: The buffering hypothesis,' en, *Clinical Psychology Review*, vol. 31, no. 4, pp. 563–591, Jun. 2011. DOI: 10.1016/j.cpr.2010.12.007.
- [7] A. Ligthart, C. Catal and B. Tekinerdogan, 'Systematic reviews in sentiment analysis: A tertiary study,' *Artificial Intelligence Review*, vol. 54, pp. 4997–5053, 2021. DOI: 10.1007/s10462-021-09973-3.
- [8] R. Stine, 'Sentiment analysis,' *Annual Review of Statistics and Its Application*, 2019. DOI: 10.1146/annurev-statistics-030718-105242.
- [9] R. K. S and N. Patil, 'Pioneering frontiers in natural language processing: Exploring promising areas and emerging applications,' in *2023 OITS International Conference on Information Technology (OCIT)*, 2023, pp. 640–647. DOI: 10.1109/OCIT59427.2023.10430927.
- [10] C. Friedman, T. C. Rindflesch and M. Corn, 'Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the national library of medicine,' *Journal of Biomedical Informatics*, vol. 46, no. 5, pp. 765–773, 2013, ISSN: 1532-0464. DOI: <https://doi.org/10.1016/j.jbi.2013.06.004>.
- [11] L. Yue, W. Chen, X. Li, W. Zuo and M. Yin, 'A survey of sentiment analysis in social media,' *Knowledge and Information Systems*, pp. 1–47, 2018. DOI: 10.1007/s10115-018-1236-4.
- [12] H. H. Do, P. Prasad, A. Maag and A. Alsadoon, 'Deep learning for aspect-based sentiment analysis: A comparative review,' *Expert Syst. Appl.*, vol. 118, pp. 272–299, 2019. DOI: 10.1016/j.eswa.2018.10.003.
- [13] R. V and S. Naik, 'Natural language processing,' *International Journal of Computing Algorithm*, 2023. DOI: 10.20894/ijcoa.101.012.002.001.
- [14] I. Nair and R. AswathyP, 'Natural language processing in medicine: A review,' *International journal of engineering research and technology*, vol. 7, 2019.
- [15] M. Agarwal, 'An overview of natural language processing,' *International Journal for Research in Applied Science and Engineering Technology*, 2019. DOI: 10.22214/ijraset.2019.5462.
- [16] I. Lauriola, A. Lavelli and F. Aiolli, 'An introduction to deep learning in natural language processing: Models, techniques, and tools,' *Neurocomputing*, vol. 470, pp. 443–456, 2021. DOI: 10.1016/j.neucom.2021.05.103.

- [17] L. M. G. Salazar, O. Loyola-González and M. A. Medina-Pérez, ‘An explainable approach based on emotion and sentiment features for detecting people with mental disorders on social networks,’ *Applied Sciences*, 2021. DOI: 10.3390/app112210932.
- [18] M. Munikar, S. Shakya and A. Shrestha, ‘Fine-grained sentiment classification using bert,’ *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, vol. 1, pp. 1–5, 2019. DOI: 10.1109/AITB48515.2019.8947435.
- [19] J. Sawalha, M. Yousefnezhad, Z. Shah, M. R. G. Brown, A. Greenshaw and R. Greiner, ‘Detecting presence of ptsd using sentiment analysis from text data,’ *Frontiers in Psychiatry*, vol. 12, 2022. DOI: 10.3389/fpsy.2021.811392.
- [20] S. Nasrullah and A. Jalali, ‘Detection of types of mental illness through the social network using ensembled deep learning model,’ *Computational Intelligence and Neuroscience*, vol. 2022, 2022. DOI: 10.1155/2022/9404242.
- [21] T. Whitaker, B. Beranger and S. Sisson, ‘Logistic regression models for aggregated data,’ *Journal of Computational and Graphical Statistics*, vol. 30, pp. 1049–1067, 2019. DOI: 10.1080/10618600.2021.1895816.
- [22] S. González-Carvajal and E. Garrido-Merchán, ‘Comparing bert against traditional machine learning text classification,’ *ArXiv*, vol. abs/2005.13012, 2020. DOI: 10.47852/bonviewJCCE3202838.
- [23] A. Joshy and S. Sundar, ‘Analyzing the performance of sentiment analysis using bert, distilbert, and roberta,’ *2022 IEEE International Power and Renewable Energy Conference (IPRECON)*, pp. 1–6, 2022.
- [24] S. Yu, J. Su and D. Luo, ‘Improving bert-based text classification with auxiliary sentence and domain knowledge,’ *IEEE Access*, vol. 7, pp. 176 600–176 612, 2019. DOI: 10.1109/ACCESS.2019.2953990
- [25] F. L. Huang, ‘Alternatives to logistic regression models in experimental studies,’ *The Journal of Experimental Education*, vol. 90, pp. 213–228, 2019. DOI: 10.1080/00220973.2019.1699769

LIST OF AUTHORS

Badnjević Almir
Faculty of Pharmacy, University of
Sarajevo, Zmaja od Bosne 7
71000 Sarajevo, Bosnia and Herzegovina

Banjanović-Mehmedović Lejla
University of Tuzla
Faculty of Electrical Engineering
Dr. Tihomila Markovića 1,
75000 Tuzla, Bosnia and Herzegovina

Bećirović Faruk
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence, 71 000 Sarajevo
Bosnia and Herzegovina

Bičo Ćar Mirha
University of Sarajevo
School of Economics and Business
Trg oslobođenja 1, 71000 Sarajevo
Bosnia and Herzegovina

Bratovčić Amra
University of Tuzla,
Faculty of Technology
UrfetaVejzagića 8, 75000 Tuzla
Bosnia and Herzegovina

Cico Maida
University of Sarajevo
School of Economics and Business
Trg oslobođenja 1
71000 Sarajevo, Bosnia and Herzegovina

Dimitrijević Sonja
"Mihajlo Pupin" Institute
Computer systems
Belgrade, Serbia

Dreković Edisa
University of Niš
Faculty of Economics, Niš, Serbia

Gurbeta Pokvić Lejla
Research Institute Verlab for Biomedical
Engineering, Medical Devices and
Artificial Intelligence
Ferhadija 27, 71000 Sarajevo
Bosnia and Herzegovina

Gurdić Ribić Azra
Foundation for Innovation
Technology and Transfer of Knowledge
Bosnia and Herzegovina

Hundur Hiyari Madžida
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence
71 000 Sarajevo, Bosnia and Herzegovina

Husak Ermin
University of Bihać
Technical Faculty, St. Irfana Ljubijankića
77000 Bihać, Bosnia and Herzegovina

Husaković Anel
Eacon doo Zenica
Zenica, Bosnia and Herzegovina

Ivanjko Edouard
Faculty of Traffic and Transport Sciences
Vukelićeva 4, HR-10000 Zagreb
Croatia

Jahić Jasmin
University of Cambridge
The Old Schools, Trinity Ln
Cambridge CB2 1TN, United Kingdom

Jokić Aleksandar
Faculty of Mechanical Engineering
Production engineering department
Belgrade, Serbia

Karabegović Edina
University of Bihać, Technical Faculty
St. Irfana Ljubijankića, 77000 Bihać
Bosnia and Herzegovina

Karabegović Isak
Academy of Sciences and Arts
of Bosnia and Herzegovina
Bistrik 7, Sarajevo, 71000 Sarajevo
Bosnia and Herzegovina

Kurjak Asim
Sarajevo Medical School, Sarajevo School
of Science and Technology
Francuske revolucije bb
71000 Sarajevo, Bosnia and Herzegovina

Lemeš Samir
University of Zenica, Polytechnic Faculty
Fakultetska 1, 72000 Zenica
Bosnia and Herzegovina

Lindov Osman
University of Sarajevo
Faculty of Transport and Communications
Zmaja od Bosne 8, 71000 Sarajevo
Bosnia and Herzegovina

Mahmić Mehmed
University of Bihać
Technical Faculty, St. Irfana Ljubijankića
77000 Bihać, Bosnia and Herzegovina

Majstorović V. Nemanja
Dental Practice ELIDA 1
Belgrade, Serbia

Majstorović D. Vidosav
University of Belgrade
Faculty of Mechanical Engineering
Belgrade, Serbia

Malisa Viktorijo
F-AR Förderung der Automation und
Robotik
Gutheil-Schoder-Gasse 8
1100 Wien, Austria

Mandžuka Sadko
Faculty of Traffic and Transport Sciences
Vukelićeva 4, HR-10000 Zagreb
Croatia

Merdović Nejra
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence
71 000 Sarajevo
Bosnia and Herzegovina

Miljković Zoran
Faculty of Mechanical Engineering
Production engineering department
Belgrade, Serbia

Mojsilović Marija
Academy of Professional Studies Sumadija
College in Trstenik,
Trstenik, Serbia

Mrdanović Emina
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence
71 000 Sarajevo
Bosnia and Herzegovina

Pepić Selver
Academy of Professional Studies Sumadija
College in Trstenik,
Trstenik, Serbia

Petrović Milica
University of Belgrade
Faculty of Mechanical Engineering
Production engineering department
Belgrade, Serbia

Prljača Naser
University of Tuzla
Faculty of Electrical Engineering
Dr. Tihomila Markovića 1
75000 Tuzla, Bosnia and Herzegovina

Radončić Amina
International Burch University
Natural and Medical Sciences
Sarajevo, Bosnia and Herzegovina

Rodić Aleksandar
Mihajlo Pupin Institute
University of Belgrade
11000 Belgrade, Serbia

Salkić Irma
Faculty of Electrical Engineering
University of Sarajevo
Sarajevo, Bosnia and Herzegovina

Saračević Muzafer
University of Novi Pazar
Department of Computer Sciences
Novi Pazar, Serbia

Simeunović Vladimir
"Mihajlo Pupin" Institute
Computer systems, Belgrade, Serbia

Smajlhodžić-Deljo Merima
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence, 71 000 Sarajevo
Bosnia and Herzegovina

Softić Adna
Verlab Research Institute for Biomedical
Engineering, Medical Devices and
Artificial Intelligence, 71 000 Sarajevo
Bosnia and Herzegovina

Spahić Lemana
Research Institute Verlab for Biomedical
Engineering, Medical Devices and
Artificial Intelligence, Ferhadija 27
71000 Sarajevo, Bosnia and Herzegovina

Stošić Dragan
"Mihajlo Pupin" Institute
Computer systems, Belgrade, Serbia

Stupar Savo
University of Sarajevo
School of Economics and Business
Trg oslobođenja 1
71000 Sarajevo, Bosnia and Herzegovina

Šahić Elvir
University of Sarajevo
School of Economics and Business
Trg oslobođenja 1, 71000 Sarajevo
Bosnia and Herzegovina

Teofilović Žaklina
University of Belgrade
Faculty of Economics
Belgrade, Serbia

Todorović Aleksandar
Faculty of Dentistry
University of Belgrade
Belgrade, Serbia

Vidović Krešimir
Ericsson Nikola Tesla
Krapinska 45, HR-10000 Zagreb
Croatia

Vojić Samir
University of Bihać
Technical Faculty, St. Irfana Ljubijankića
77000 Bihać, Bosnia and Herzegovina

CO-ORGANIZERS - CONFERENCE PARTNERS



**MINISTRY OF ENERGY, MINING AND
INDUSTRY OF THE FEDERATION
BOSNIA AND HERZEGOVINA**
<https://fmeri.gov.ba/>



**ROBOTICS SOCIETY
IN BOSNIA AND HERZEGOVINA**
<http://www.robotika.ba/>



**FOREIGN TRADE CHAMBER
BOSNIA AND HERZEGOVINA**
<https://www.komorabih.ba/>



**CHAMBER OF ECONOMY OF THE
FEDERATION OF BOSNIA AND
HERZEGOVINE**
www.kfbih.com



**INTERNATIONAL ASSOCIATION
FOR TECHNOLOGICAL
DEVELOPMENT AND INNOVATIONS
UKRAINA**
<http://iatdi.org/>



**TECHNOLOGY PARK
INTERA MOSTAR
BOSNIA AND HERZEGOVINA**
www.intera.ba/



CITY OF BIHAĆ
<https://bihac.org/>



CITY OF ČAZIN
<https://gradcazin.gov.ba/>



CITY OF BOSANSKA KRUPA
<https://gradbosanskakrupa.ba/>

