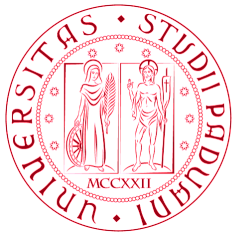

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITY OF PADOVA

DEPARTMENT OF PUBLIC, INTERNATIONAL
AND COMMUNITY LAW - DiPIC

Bachelor Degree in Law and Technology

The Role of Artificial Intelligence in Next-Generation Wireless Networks - an Overview of Technological and Law Implications

Supervisor
Prof. Francesca Meneghello

Candidate
Annachiara Lunardi

Academic Year
2022–2023

Abstract

This thesis explores how artificial intelligence (AI) will benefit next-generation wireless networks focusing both on the technological aspects and legal implications.
At first, a summary of the history of AI is provided, together with an overview of the different AI methodologies. Among them, the thesis focuses on machine learning-based approaches and, in particular, neural network algorithms for wireless networks.
The second chapter examines how AI can support wireless network management and explores the advantages of adopting this new paradigm as a substitute for current non-data-driven approaches. Next, we discuss the technical challenges that should be addressed to enable the practical integration of AI within wireless networks, starting from the huge amount of data needed to properly configure AI methodologies and the high computational demand. Emerging approaches that will allow overcoming the above-mentioned challenges, such as, e.g., the adoption of federated learning and mobile edge techniques, are also discussed.
In the third chapter, we examine the cybersecurity risks that arise when integrating AI within wireless networks, and the need for regulations that will help address these risks.
A vision of the future of AI in wireless telecommunication networks, and a discussion of the open research challenges from technological and legal points of view conclude the thesis.

Contents

Listing of acronyms

Symbols

5G fifth generation

6G sixth generation

A

AI Artificial Intelligence

ANNs Artificial Neural Networks

C

CNN Convolutional Neural Networks

D

DL Deep Learning

DRL Deep Reinforcement Learning

F

FNN Feedforward Neural Networks

I

IDS Intrusion Detection Systems

IoT Internet of Things

M

MAC Medium Access Control

MEC Mobile Edge Computing

MIMO Multiple-Input-Multiple Output

ML Machine Learning

N

NLP Natural Language Processing

NN Neural Networks

Q

QoS Quality of Service

R

RNN Recurrent Neural Networks

V

VR Virtual Reality

1

Introduction

In a world marked by the relentless evolution of technology, Artificial Intelligence (AI) stands as a innovative force that has revolutionized various domains of human endeavor.
Wireless communication networks have not been immune to this shift, with AI playing an important role in shaping the future of next generation wireless networks.
This thesis explores how AI affects wireless networks in many different ways, delving into both the technological advancements and the intricate legal implications that this intersection entails.
The motivations behind this research are deeply rooted in the realization that the intersection of artificial intelligence and wireless networks holds the promise of not only enhancing our digital connectivity but also reshaping the technological infrastructure. The proliferation of wireless communications has become an indispensable aspect of our daily lives, spanning from mobile devices and the Internet of Things (IoT) to autonomous vehicles and smart cities. However, the incessant demand for faster, more reliable, and secure wireless connectivity presents an array of challenges that demand innovative solutions: a compelling one is the integration of AI into wireless networks, to address these challenges.
Through intelligent algorithms, machine learning, and neural networks, AI has the potential to optimize network performance, enhance security, and pave the way for the next phase of wireless communication.
By exploring the intricate synergy between these two fields, this research aims to shed light on the possibilities that lie ahead. This thesis seeks to provide an exploration of the role of AI in next-generation wireless networks, elucidating the potential benefits it offers while navigating the complex landscape of legal and regulatory considerations. It begins with a historical overview of AI and an examination of the various methodologies that have led to its current relevance. Special attention is devoted to revealing AI's subsets Machine Learning (ML) and Deep Learning (DL). Then, particular emphasis is placed on machine learning-based approaches, with a focus on neural network algorithms tailored for wireless networks.

The second chapter delves into the practical aspects of AI integration into wireless networks, focusing on the switch from the old-fashioned ways to the new data-driven AI methods. Technical challenges, including the massive data requirements and computational demands, are scrutinized. Innovative approaches, such as federated learning techniques and mobile edge computing are discussed as potential remedies.

The third chapter extends its focus to cyber security risks inherent in the adoption of AI in wireless networks. Traditional cybersecurity measures, while valiant, find themselves struggling with the complexity and sophistication of modern cyberattacks. AI and DL play a pivotal role in strengthening our digital defenses. from threat detection mechanisms to countering malware, ransomware, and ensuring user privacy, the fusion of DL and cybersecurity emerges as a formidable response to the evolving threat landscape. Legal implications and the requisite regulatory frameworks to mitigate these risks are examined in detail, especially the introduction of the AI Act, in the European Union, that announces comprehensive rules governing AI development and deployment.

Finally, the thesis concludes by envisioning a future where the integration of AI, ML and cybersecurity not only fortifies our digital landscape, but also fosters innovation and ethical practices.

2

Fundamentals of Artificial Intelligence and Machine Learning

2.1 Defining Artificial Intelligence

Over the years, Artificial Intelligence (AI) has been the subject of numerous definitions, each attempting to capture its complexity and scope. The Oxford Dictionary, for instance, describes it as "The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages." It represents a remarkable technological advancement that strives to replicate human-like intelligence in machines.

The applications of AI are diverse and pervasive, permeating numerous aspects of modern life. AI finds application in various fields, making it a versatile technology. In the realm of healthcare, AI is used for medical diagnosis, drug discovery, and personalized treatment recommendations. In finance, AI-driven algorithms manage trading, detect fraud, and optimize investment portfolios. The technology underpins virtual personal assistants, such as Siri and Alexa, enabling voice recognition and natural language understanding. AI powers autonomous vehicles too, facilitates predictive maintenance in industry, and enhances recommendation systems on streaming platforms like Netflix. Moreover, AI has been largely used also in the field of law. AI helps automate the discovery process, that is difficult and time consuming for humans [1].

Hence, this remarkable technology has not only found various applications in said fields, but has also revolutionized our modern life.

2.2 The beginning of Artificial Intelligence

The idea of inanimate objects possessing intelligence has ancient origins. Starting from 1940s, when Isaac Asimov published the Three Laws of Robotics: three laws which have become iconic whose purpose was to outline a set of ethical principles designed to govern the behavior of artifi-cial intelligence and robots. Asimov's laws emphasized the importance of ensuring that artificial intelligence and autonomous machines should prioritize the safety and well-being of humans above all else. The rules are as follows:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law [2].

Asimov's visionary concept has had a lasting impact on the discourse surrounding AI ethics and the relationship between humans and intelligent machines. It serves as a foundation for discussions on responsible AI development and the potential risks and benefits associated with artificial intelligence.

Moving on, in the 1950s Alan Turing, a pioneering figure in the world of computing, published a paper titled "Computing Machinery and Intelligence". In this work, he introduced the concept that has become renowned as the Turing Test. This test serves as a method for assessing the intelligence of a machine by evaluating its ability to engage in natural language conversations to a degree indistinguishable from that of a human [3].

The birth of AI is considered to be in 1956 when the phrase "artificial intelligence" was coined at a summer conference held at Dartmouth College. This conference was led by John McCarthy, the individual credited with coining the term. Also present were Allen Newell, a computer scientist and Herbert A. Simon, a polymath, who during the conference unveiled their groundbreaking creation, the Logic Theorist: a computer program capable of proving specifically mathematical theorems, marking it as the inaugural AI program of its kind [3].

In 1959, Arthur Samuel, during his time at IBM (International Business Machines Corporation) introduced the term "machine learning". Samuel's work was instrumental for the development of machine learning as a field within artificial intelligence [3].

The research in the field of AI stalled when a major Cold War project of The Automatic Language Processing Advisory Committee (ALPAC), aimed at traducing Russian, failed. The period from 1974 to 1980, known as the "First AI Winter", experienced stagnation in AI research due to fund-ing constraints. In the early 1990s, a "Second AI Winter" came: experts claimed that systems were too expensive to maintain and update [1].

Until 1997, when IBM's Deep Blue, a supercomputer, defeated the world chess champion, Gary Kasparov [4]. This historic chess match marked an important moment in the history of AI, as it demonstrated the capability of a machine to outperform the world's top human chess player in a highly complex and strategic board game. Deep Blue's victory was seen as a turning point,

demonstrating the power of AI and machine learning in confronting intricate problems and tasks. It was a proof of the rapid advancement of computer hardware and algorithms, as well as the potential for AI to excel in domains that require strategic thinking and pattern recognition.
Approaching the early 2000s, there were further advancements that include Google's introduction of its search engine, Amazon's launch of the recommendation engine and Netflix's movie recommendation system. Furthermore, in the decade between 2010 and 2020 Siri and Alexa voice assistants are introduced and self-driving cars emerged [1]. In 2016, Hanson Robotics introduced a humanoid robot, the first "robot citizen" named Sophia, capable of recognizing faces, engaging in verbal communication and displaying facial expressions [3].
Afterward, the OpenAI lab was founded, which led to the development in 2020 of a natural language processing model ChatGPT-3, designed to generate text that emulates human speech and writing styles. Indeed, this current decade, has experienced the arising of generative AI as a category of artificial intelligence technology capable of producing original content. Generative AI operates by using a given prompt, which can take the form of text, images, videos, designs, musical notes, or any input the AI system can handle. As we arrive at the present, exactly in 2023, witnessing the unveiling of ChatGPT-4, the latest update [3].

2.3 Defining Machine Learning and its applications

Essentially, AI is a vast field and has various subfields such as Machine Learning (ML), Deep Learning (DL) and Neural Networks (NNs) (see Figure 2.1).

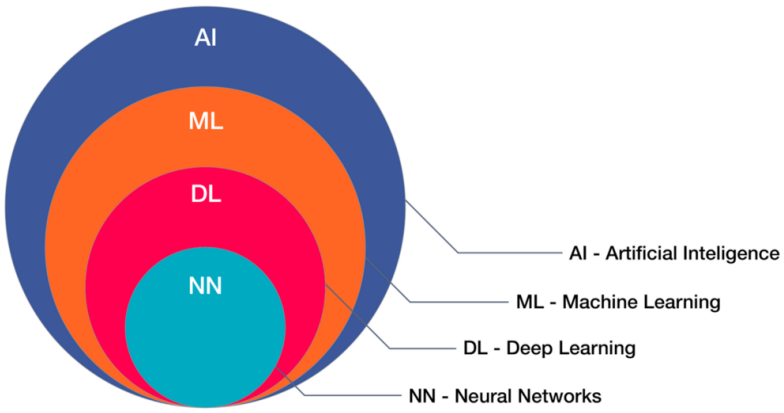

Figure 2.1: AI tis dedticated to automatting tasks that typtically requtire human tintellect, and Machtine Learnting (ML) and Deep Learnting (DL) represent spectiftic approaches to accompltish thtis objecttive. In essence, they fall under the umbrella of Arttiftictial Intelltigence (AI).

The capabilities of AI have significantly grown in recent years, largely due to the advancements within its subsets.
Firstly, Machine Learning is a branch of AI dedicated to developing algorithms that enable com-

puter systems to learn from experience, so without being explicitly programmed by humans. The difference between ML and traditional programming is that in traditional programming, the programmers are responsible for crafting the rules that dictate a program's behavior on a machine. On the other hand, in ML, the model learns the rules directly from the data it is given. This means that humans do not need to write the explicit rules, but only the code that the machine requires to build the ML model [5].

ML algorithms can recognize patterns, make informed decisions, and adapt to changing situations. Machine Learning finds application in various domains, including computer vision, where machines interpret visual information such as recognizing images or videos and prediction where ML models can predict future outcomes based on historical data patterns. Another key area is semantic analysis that involves the understanding of the meaning of words and language, enabling systems to comprehend context and intent. In Natural Language Processing (NLP), ML enables computers to work with human language, facilitating language translations, chatbots, and text analysis. Furthermore, ML is crucial in information retrieval where it helps the extraction of relevant data from extensive datasets or documents [6].

In ML, there are four primary learning methods, each useful for solving different tasks: supervised, unsupervised, semisupervised, and reinforcement learning [5].

In supervised learning the data sets are labeled so the machine can detect patterns that can then be applied to new, unlabeled datasets. In this learning method, there exists a specific goal, which is to predict the target using a dataset that includes numerous instances related to various features. The first fundamental step of supervised machine learning consists of acquiring a dataset and dividing it into separate training, validation, and test datasets. Secondly, using the training and validation datasets, a model is instructed on the relationship between features and the target; next the model is assessed through the test dataset to determine how effectively it predicts the target for unseen cases [4]. In each iteration, the algorithm's performance on the training data is compared to its performance on the validation dataset. As a result, the algorithm is adjusted or optimized using the validation set. Given that the validation set may differ from the test set, the algorithm's performance may or may not generalize [5].

The two primary types of supervised learning tasks are regression and classification. While regression involves the prediction of numerical data, classification determines the category to which an example belongs.

Moving to unsupervised learning, this technique focuses on unlabeled datasets and classifies data based on inherent similarities or differences. In this case, there is not a specific goal to reach, instead the algorithm tries to discover patterns or data structures without being instructed on what to search for. This approach is useful when there is the need to detect potential connec-tions between instances. In this method, the three common tasks are clustering, association and anomaly detection. Regarding the process of clustering, it groups the instances based on their similar characteristics which helps to reveal structures in the data observed [5].

Semi-supervised learning adopts a hybrid approach: labeled and unlabeled data are combined to enhance learning accuracy. Generally, it is used when labeling data becomes time-consuming or cost-prohibitive, so it is used in scenarios where only a portion of the data is labeled and used

to train a model. Then, this model is employed to classify the remaining unlabeled data in the dataset and the resulting labeled dataset is then used to train a functional model that, theoretically, should outperform unsupervised models [5].
Lastly, reinforcement learning takes on unlabeled datasets, with the AI system learning through actions and feedback. In this case, there is not just one correct answer, but a desired overall result. This approach is considered one of the most human-like ways of training algorithms because, like humans, it learns by experimenting and making decisions through trial and error, rather than relying on pre-existing data. During this process, the algorithm gradually learns the preferred behaviors and hopefully, over time, acquires the ability to reach the correct result [5].

2.4 Defining Neural Networks and its architectures

The other essential component within the field of AI is Deep Learning (DL), which constitutes a more specialized and advanced branch of machine learning. Deep learning algorithms are designed to mimic the way the human brain processes data and creates patterns through a biologically inspired neural network architecture [5].
These Neural Networks contain multiple hidden layers and parameters that process data, enabling the machine to delve deeply into learning, forging connections and assigning input weights to optimize outcomes.
NNs use a series of layers to extract and transform features from data. The lower layers learn simple features, and the higher layers learn more complex features based on the simpler ones below. This creates a powerful way to represent features. It means that deep learning is good at understanding and finding useful information in large amounts of data and data from different sources [6].
Focusing on Neural Networks, the fundamental components that make up deep learning algorithms, they consist of interconnected nodes (the neurons), an input layer, one or more hidden layers and an output layer. The depth of a neural network is determined by the number of its hidden layers: a neural network comprising two or more hidden layers is referred to as a Deep Neural Network [4]. A visual example is depicted in Figure 2.2. Every node represents an artificial neuron, linked to the next, and is characterized by weight and threshold values. When the output of a node surpasses its threshold value, the node becomes active and transmits its data to the subsequent layer of the network. If the output falls below the threshold, no data is forwarded [7].

The tasks that Neural Networks (NN) are designed to do are speech and image recognition, and sometimes they perform even better than humans [4]. NNs can be structured with different architectures. The fundamental ones are Feedforward Neural Networks (FNN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) [4].
Starting from FNNs, also called multilayer perceptrons, they are the simplest type of neural network architecture. They consist on multiple layers of neurons. These networks are called "feedforward" because data flows in one direction only, from the input layer through the hidden layers to the output layer [4].

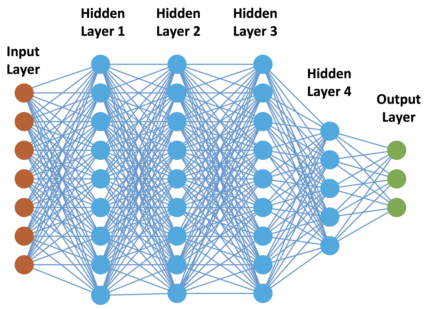

Figure 2.2: Archtitecture of a DNN [8].

On the other hand, RNNs, have a fully interconnected graph that also contains loops so it allows them to maintain a form of memory. This looped structure enables RNNs to capture dependencies within sequential data, making them suitable for tasks like language modeling, speech recognition and video processing [4]. The general structure of an RNN is depicted in Figure 2.3.

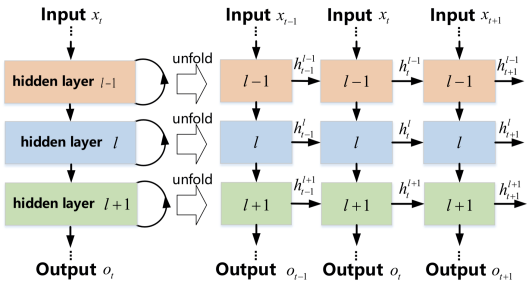

Figure 2.3: Archtitecture of a RNN [8].

Finally, CNNs are specifically designed for processing grid-like data, such as images and videos. They utilize convolutional layers that automatically learn spatial hierarchies of features, as represented in Figure 2.4. CNNs are highly effective in image-related tasks, including image recognition, object detection, and image classification [4].

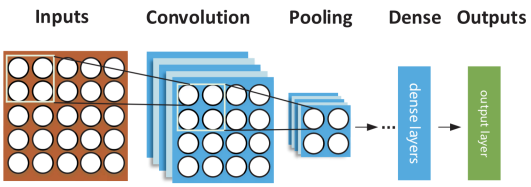

Figure 2.4: Archtitecture of a CNN [8].

3

Artificial Intelligence in Wireless Network

3.1 Artificial Intelligence in Telecommunications Networks

In recent years, the telecommunication network has grown immensely in size and complexity, therefore managing and controlling this vast, intricate system has become quite a challenge.

From the introduction of 2G networks that brought text messaging (SMS) as a form of data transmission, to the widespread adoption of 4G technology, providing fast internet access on our mobile devices, we are now entering a new era with fifth generation (5G) and the forthcoming sixth generation (6G) wireless networks.

The transition from 4G to 5G marked an important moment in wireless communication, promising faster speeds and improved capabilities for connecting devices, industries, and people. 5G, in particular, has reached a very high level of performance: it can support enhanced mobile broadband (eMBB), massive machine-type communications (mMTC) and ultra reliable and low-latency communications (uRLLC) [9]. Thus, this technological advancements, has given rise to new applications and opportunities including augmented reality (AR), virtual reality (VR), tactile reality, mixed reality, and more. Still, looking ahead to 6G, it will require even more higher transmission speeds and larger scale communications, paving the way for innovative applications such as holographic communications. So with this technological leaps, our networks are preparing to be incredibly smart and autonomous: soon they will have the ability to perceive, process, learn, and make decisions, all while handling the massive influx of data from a countless number of smart devices.

Next-generation wireless networks will have to ensure ultra-reliable, low-latency and secure communication that can also adapt to the dynamics of IoT devices. Self-driving cars, for example,

will need reliable and fast communication to make quick and automated decisions: so they demand low-latency control. At the same time, there will be many sensors and wearables, like smartwatches, collecting real-time data that has to be sent to the network. These short-packet transmissions (transmission of small amount of data) will create traffic on the wireless uplink. Furthermore, the upcoming networks will have to support different applications that include cloud-based gaming, immersive virtual reality services, real-time HD streaming and traditional multimedia services [10]. On top of all this, the upcoming wireless networks must also provide a great experience for every service and user. This means that they should measure user experience and gain insights into customer expectations to guarantee adequate quality of service to the end users. Hence, there needs to be a shift in the design and planning of more efficient networks.

To address these challenges, integrating context-aware algorithms within the networks' management procedures is now a necessity. Artificial intelligence plays a crucial role in this transformation. It acts like the intelligent conductor coordinating tasks, simplifying intricate processes, and obtaining the proper action to be performed by analyzing a high amount of real-time data from distributed systems [11].

Specifically, the integration of machine learning, with a focus on artificial neural network (ANN)-based approaches, both within the wireless infrastructure and at end-user devices, has been proven to be a valuable way to address the aforementioned challenges [10].

3.2 Advantages

AI and ML have already demonstrated their value, providing substantial advantages to the telecommunication system.

Firstly, they improve effectiveness due to AI remarkable ability to automate repetitive tasks and respond quickly and efficiently thus liberating human resources for other crucial purposes. These models analyze network data and traffic patterns, revolutionizing capacity planning and network optimization. The ultimate goal would be to remove the need for human interventions.

Secondly, they will contribute in boosting the performance of telecommunication networks. Examining how network traffic flows and making sure resources are used efficiently, ML paradigms can help maintain a high performance even if networks are becoming more and more intricate. This leads to higher data transmission speed and reliability, resulting in improved transmission speed and latency, leading to reduced delays.

Moreover, AI offers great solutions also for saving energy, promoting more sustainable connections. Mainly, clever algorithms will be capable of acting on traffic in real-time, or even predicting it, optimizing the use of energy and reducing its impact on the environment.

Security and trustworthiness are also important matters in which ML intervenes: it can detect and prevent security threats, such as cyberattacks and fraud, thus protecting the network and the customers' personal data. The ethical implications of AI are also addressed by ensuring that transparency and explainability are integrated into the systems, providing the necessary information to explain, enhance or resolve issues.

Lastly, the ability of AI to detect patterns assists in the understanding of the customers' needs and

requirements, which leads to the discovery of new business opportunities. It gains deeper insights into the preferences of the customers, offering personalised services, thus creating a competitive and customer-centric system. With the advent of user-focused wireless services, like virtual reality, the gap between user and network has greatly decreased, thus highlighting the need to monitor and adjust to human user behavior. In this context, machine learning stands out as a tool capable of learning and imitating human behavior. This ability helps tailoring wireless network functions to human users, creating an immersive environment, and enhancing the overall quality of the user experience [12].

3.3 Use cases

ML algorithms can be used in many different ways in telecommunications networks. In the context of a wireless system, one of the most straightforward uses of Machine Learning is to apply intelligent and predictive data analytics to improve the system's understanding of its surroundings and its overall operational efficiency. ML plays a crucial role in allowing the wireless network to process large volumes of data from diverse sources that can include, for example, data from wireless signal measurements, sensor readings, information collected by drones, and images from surveillance cameras. The goal is to create a complete operational map of the network, encompassing the numerous connected devices, and once this map is generated, it can be employed to enhance various network functions. For example, it can improve fault monitoring, enabling the network to detect and address issues more effectively. Additionally, it can enhance user tracking, which can be valuable for applications where it's important to keep an eye on the movement and activity of users within the wireless network [10].
Furthermore, ML contributes to network optimization by making them smarter and more data-driven. In particular, ML algorithms enable the introduction of intelligent resource management solutions that can address various issues such as cell association, radio access technology selection, frequency allocation, spectrum management, power control and intelligent beamforming [10]. Unlike the conventional techniques used to optimize networks that work in an ofline manner, machine learning (ML)-based resource management works in real-time. It continuously learns from the wireless environment and user behaviors, improving its performance over time. This real-time learning is vital for IoT and 5G services requiring immediate, low-latency responses. Well-designed ML optimization algorithms can provide self-organizing, self-healing, and self-optimizing solutions for network management challenges, which is particularly beneficial in ultra-dense wireless networks where traditional methods struggle due to the network's vast scale and diversity.
Last, but not least, ML offers valuable capabilities even in the physical layer of wireless networks. Machine learning applications can reshape the design for physical layer functions, including coding and modulation, for both the sender and receiver components in a standard communication system. This approach has demonstrated considerable potential in reducing bit error rates and strengthening resistance to wireless channel obstacles.

3.3.1 Physical and MAC layers

Now, let's delve into specific ML tools that help enhance the performance of the physical and MAC layers of wireless networks. Starting from Multiple-Input-Multiple Output (MIMO) transceivers, that in 5G networks, are crucial but require accurate channel estimation for reliable symbol detection. Massive MIMO systems face challenges in obtaining this Channel State Information (CSI) due to a training process with significant overhead (additional data or resources used in the training process). To tackle this, recent deep learning (DL) approaches have been introduced. They aim to reduce the overhead while enhancing accuracy. One possible approach presented in [13] combines MIMO channel estimation and signal detection, considering factors like channel estimation errors and statistics. This model-driven DL MIMO detector outperforms traditional iterative detectors, providing significant performance improvements.

Moving on to beamforming: it is a wireless transmission technique that uses multiple antennas to focus the signal in a specific direction, improving signal quality and coverage. This approach is especially useful when working on the millimeter-wave (mmWave) portions of the radio spectrum. Recent deep learning-based approaches aim to reduce the complexity of the beamforming process while maximizing the signal quality relative to interference and noise, i.e., the signal-to-interference plus noise ratio (SINR) [4]. To make beamforming techniques more efficient, hybrid analog and digital precoding is an interesting method to reduce hardware complexity. Since the current schemes of hybrid preaching suffer from overhead and complexity problems, a deep learning-based mmWave massive MIMO framework has been proposed in [14], which has shown significant improvements in hybrid preceding, spectral efficiency and reduced transmission errors.

Dynamic Spectrum Sharing (DSS) techniques provide a valuable solution for managing ultra-dense networks (UDN), that are characterized by a very high number of nodes and used in areas with a high demand for wireless connectivity [4]. Recently, a practical approach for UDN was introduced in [15], which enables multiple users to access the same channel at the same time through spatial separation. To make this work efficiently, an algorithm that uses decentralized reinforcement learning was suggested in [15], enabling secondary users (SUs) to adapt to the network load linked with primary users reaching a state of equilibrium known as Nash Equilibrium (NE). This adjustment is based on each individual user's observations and their own past actions, without them exchanging information globally with each other but relying only on their own experiences and insights (so without central coordination).

ML can also help at the Medium Access Control (MAC) layer of the protocol stack. In the context of a telecommunications network, the MAC layer is responsible for managing how different devices and users access the network's resources. ML contributes to determining the best way to connect users to radio access points and Mobile Edge Computing (MEC) hosts, which are critical components in a modern network. Additionally, ML is useful to improve the handover process, i.e., the mobile device's transition from one radio access point to another as it moves. By predicting how the devices will move, ML can run the transitions in a smooth and efficient way. A recent survey [16] explored how Deep Reinforcement Learning (DRL), a subset of ML, can be applied to enhance the Quality of Service (QoS) in the MAC layer and it addresses some of the main

challenges in the context of QoS, including how to efficiently manage data rates, share network resources among users, and schedule activities within the network to ensure the best performance and user experience.

3.3.2 Network, Transport and Application layers

Besides, ML techniques can also provide a significant improvement to the network and application layers. Delving into the activities that can gain from ML, mobile data analysis is one of them. Through the analysis of the data (data analytics) ML algorithms are able to extract meaningful information, that can be categorized into two main types: network-level, transport-level and application-level data [4].

In particular, the network-level data provide a comprehensive overview of mobile network performance encompassing factors such as throughput, end-to-end delay, jitter, and so forth. It also logs details about individual sessions, the types of communication taking place, and information about the senders and receivers involved. These network-level data serve various purposes. They can be used for tasks such as diagnosing and managing the mobile network, analyzing how users move within the network (mobility analysis), and even planning public transportation routes more efficiently. Deep learning comes into play for predicting network traffic patterns: in essence, DL algorithms can use historical network-level data to forecast future network traffic, which can be highly valuable for network optimization and resource allocation.

Concerning the transport layer, whose role in network communications is to provide reliable transmission of data between two endpoints, is faced with challenging tasks such as the distance between these two endpoints, the underlying technologies and network congestion. To address this challenge, various transmission control protocols have been developed. Active research in the field of network congestion control, particularly focusing on ML, has led to various proposed mechanisms to improve data transmission performance. Notable examples include the Performance-oriented Congestion Control (PCC) that uses selective acknowledgments (a technique that allows to specify which individual packet has been received successfully) to evaluate how well the data transmission actions, like sending and receiving, are working and subsequently adapt data sending rates accordingly.

With regard to the application-level data instead, these can enhance the capabilities of mobile devices and benefit users in various ways. For instance, it allows for optimizing energy efficiency by accurately predicting how a device's battery discharges over time. ML algorithms, as demonstrated in a recent study (see Chapter 1 of [4]), can be used to make these predictions. This study also introduces a Deep Neural Network model, which can learn and understand the user's specific battery usage patterns, enabling customized and more accurate predictions of battery discharge Additionally, another remarkable purpose is to be found in mobile health, where data from wearable health monitoring devices, like Smart Watches, can offer real-time feedback on vital signs like heart rate, blood pressure, breathing status and have the ability to call for help if something dangerous happens (see Chapter 1 of [4]). Furthermore, other areas that benefit from these analyzed application-level data are epidemiology, urban planning and public service provisioning. Tracking

the movement patterns of the users can be useful for understanding the potential spread of diseases (as we all saw during the period of COVID-19 with the implementation of apps like “Immuni”, which used Bluetooth technology), allowing epidemiologists to identify disease hotspots and plan targeted interventions. Other usages of application-level data entail understanding how people are going to travel within a city, to make services like bike-sharing, scooter-sharing or public services more accessible and safe. Another aspect that can be improved by DL algorithms is routing: by optimizing the selection of routes and the rules that govern how data packets are directed it is possible to enhance the efficiency of routing. Examples of research show that NN are used to classify the connectivity of network nodes and determine optimal paths for data, thus reducing signaling overhead and speeding up the decision-making process. Finally, DL can also be applied to several mobile device applications, such as video quality, video streaming and wireless virtual reality. In particular, to enhance video quality a combination of unsupervised and supervised learning has been used (see Chapter 1 of[4]): they trained a machine learning model to extract information from video frames, and then used another model to map these features to the quality of video clips. This combined approach was shown to be more effective than fully supervised methods. In video streaming instead, a deep reinforcement learning algorithm is used in (see Chapter 1 of [4]) to decide which part of a video to download next from the server, considering factors like the content of the video, wireless conditions, and the user’s history. Last but not least, wireless Virtual Reality (VR), which is considered one of the most significant applications in 5G and beyond networks. While wireless VR devices are being developed, one challenge is efficient data transmission. Instead of sending the entire 360° video frame in high resolution, that would be quite unnecessary because human sight focuses only on a small part at a time, they use intelli-gent mechanisms to predict the user’s eye movement to mimic the human sight focus [10]. Based on these predictions, only the part of the video within the user’s view is sent in high resolution, improving efficiency without sacrificing quality.

3.4 Technical Challenges

At this point, it is evident that ML is gaining recognition as a powerful tool for enhancing operational effectiveness and providing new value to customers or users. However, the development of ML and DL algorithms in telecommunications poses several significant challenges. Firstly, the decision to apply machine learning depends on the nature of the wireless communication problem at hand. Machine learning is versatile, applicable to various problem types, including regression, classification, clustering, and more. In regression problems, the objective is to predict continuous values based on current input, while classification problems involve assigning input to specific classes. For machine learning to be a viable solution, the wireless communication problem must be translated into tasks suitable for machine learning techniques. Commencing with the foundational aspects of any ML project, i.e., data collection and training, to succeed and thus create an optimal ML project, well-structured data with adequate quantity and quality is needed. Sometimes, limited data availability can pose a notable problem for the training of the model, but also, even when data is actually available, issues like incorrect entries, noisy features or incomplete data, may

arise. Dealing with unbalanced or biased data is equally challenging, and while solutions exist, they still can be improved [17].

Additionally, the advent of big data introduces its own set of complexities. Training algorithms on vast data volumes can be a formidable task, and efficiently processing big data remains a challenge. In fact, many of the existing algorithms are incapable of processing this high amount of data [8]. Training complexity, especially, in implementing ANN-based RL algorithms, increases rapidly as more Base Stations (BSs) or users adopt RL algorithms. As the network becomes larger and more complex, training RL algorithms becomes computationally expensive and time-consuming, therefore high computational effort leads to longer training times [10]. Furthermore, due to the substantial increase of network nodes, using centralized algorithms for network management has become quite a challenge [8]. Traditional network management often relies on centralized algorithms running on a central server, that are responsible for making decisions and managing the entire network. As the number of nodes grows, the computing workload on the central server becomes extremely heavy, thus leading to significant delays (high-latency) and inefficiencies in decision-making.

Also, gathering information from all network nodes can be expensive and time-consuming, requiring extensive communication and data collection efforts. The decision-making process, in particular, occurs at various locations and levels, with some decisions based on local data requiring quick responses and minimal delays, while others influence the whole system and depend on data collected from multiple sources. The last ones, may require real-time reactions in critical situations, such as power-grid or node failures [11]. To address this increasing complexity in mobile networks, traditional optimization algorithms are inefficient, and more flexible ML algorithms are needed. Federated learning, in particular, serves as a valuable approach (see Section 3.5).

Additional challenges, associated with infrastructure in machine learning, include addressing memory requirements and managing energy consumption [17]. Machine learning models, particularly DL models, demand substantial memory to store results, parameters and gradients during training, thus affecting model scalability and the ability to process large datasets efficiently. The extensive computations and memory access involved in a model training can consume a significant amount of energy, especially when training deep neural networks. For instance, when implementing ANN-based RL algorithms in wireless devices, it is crucial to minimize computational resources and power needed for these algorithms to work efficiently [10]. In RL, the agent interacts with the environment taking actions and moving between different states, and the number of possible actions and states can vary. However, for ANN-based RL algorithms to operate adequately, it is often necessary to limit the number of actions and states because they are typically designed to work with discrete actions and states. As a result, problems involving continuous and unbounded variables are computationally demanding.

Another challenge involves guaranteeing trustworthy AI and, consequently, model interpretability. Autonomous systems driven by AI involve intricate models and algorithms which continually adapt to new data and insights without manual intervention. However, to allow humans to understand and develop informed trust in AI-based systems, novel approaches to ensure transparency, explainability, safety and privacy are necessary. To achieve a great level of transparency, Explain-

able AI (XAI) emerges as a valuable solution [11]. Finally, the increasing presence of intelligent machines in our lives emphasizes the importance of these machines accurately understanding human needs and intentions for an effective collaboration. Enhancing this human-machine collaboration is AI, with the advancement in natural language processing, computer vision and emotional detection that allowed machines to understand more precisely human inputs [11]. Leveraging digital twins and extended reality (XR) devices it is possible for humans to understand machines processing (see Section 3.5).

3.5 Related solutions

In light of the previously discussed challenges, addressing these intricacies in the application of machine learning and deep learning algorithms within telecommunications is fundamental.

Federated learning: it represents a progression of conventional Distributed Learning, a decentralized paradigm within the domain of Machine Learning where the learning process occurs across multiple devices, each engaging in parallel learning, as represented in Figure 3.1. This distribution of the learning process offers two main advantages: the computational workload is shared among various devices, thus promoting efficiency and leveraging the collective processing power of multiple devices; and the amount of information transmitted between devices is reduced, thus alleviating bandwidth usage and communication overhead.

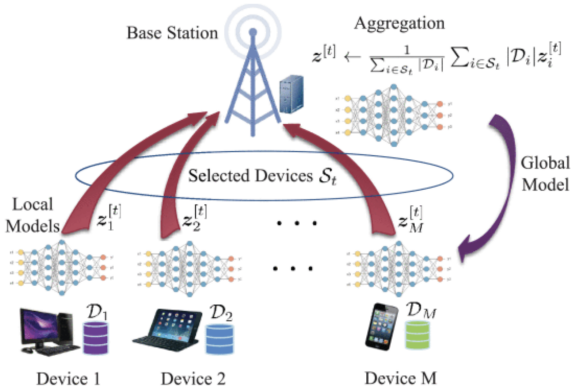

Figure 3.1: Federated Leanting system [18].

With respect to Federated Learning, in contrast to Distributed Learning, it demonstrates more capabilities in handling and managing non-identically distributed (non-iID), unbalanced and massively distributed datasets [4]. Essentially, Federated Learning embodies an advanced collaborative approach that exceeds traditional distributed techniques. This enables it to adapt very well to scenarios where data is extensively dispersed across a vast number of nodes, resulting in local datasets that are notably smaller on average compared to the total number of devices constituting the system. So, the variety in data distribution leads also to differences in the amount of training data each device has available. In a typical Federated Learning setting, therefore, each node contributes to a fraction of the overall learning process, thereby mitigating concerns related

to privacy, security, and resource limitations. Federated learning techniques can be grouped into two main types: Horizontal Federated Learning and Vertical Federated Learning [4]. In the first mentioned one, also known as sample-based federated learning, local datasets focus on the same aspect or characteristic of the observed phenomenon. That is to say that all the devices are looking at the same task, for example predicting the next word in a sentence on a smartphone keyboard. However, each device has its unique set of examples because everyone has their own way of writing sentences. The goal is to create several different local models on individual smartphones and then, by aggregation at a central server, create a final general model. Usually this method is the most used. On the other hand, in Vertical Federated Learning (also known as feature-based), local datasets focus on different aspects of the observed phenomenon, thus they do not share the same feature space: each device, indeed, emphasizes a particular dimension of the data. Considering two devices participating in federating learning, even though they have overlapping data points, they are interested in different features of the shared data allowing them to gain insights into specific aspects of the overall information.

Mobile edge caching and computing has proved to be a solution for reducing energy consumption, but also for low latency demands. In wireless networks, caching at the edge involves storing popular content (like videos or images) in network devices such as Base Stations (BSs) and end-user devices. This is done to reduce data traffic, delays, the use of network bandwidth and to improve the utilization of users' context and social information. By having caches at the edge, therefore, frequently requested content can be quickly delivered to users without having to retrieve it from a more distant location [10]. This main idea behind mobile edge caching is depicted in Figure 3.2.

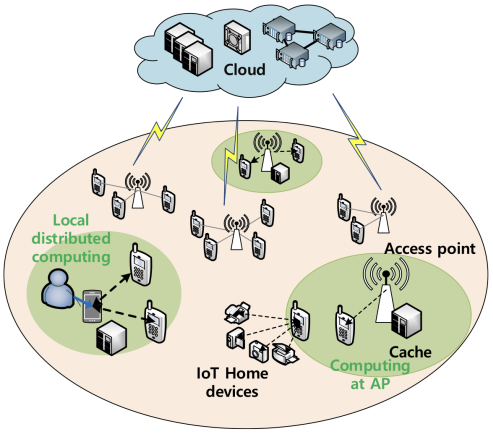

Figure 3.2: Mobtile edge cachting and computting tin wtireless networks [10].

The placement of these caches, high-speed data storage layers that temporarily store copies of frequently accessed resources, is a critical aspect, and recent advancements like coded caching enable efficient content delivery by creating multicasting opportunities, significantly improving bandwidth efficiency. However, to design effective caching strategies challenges like optimizing

cache placement, updating caches, and analyzing content popularity need to be faced. Alternatively, edge devices can also perform computations in a concept known as mobile edge computing. Basically, local resources are utilized for computational tasks, such as generating virtual reality images or processing sensor data. The goal is to avoid delays caused by sending data to remote cloud servers and this approach, which includes related concepts like fog computing, reduces overall computational latency by distributing tasks across local and remote devices. The challenge in mobile edge computing is to optimally allocate computational tasks across edge devices and remote servers to minimize latency [10].

Recent work has also explored combining caching and computing, where caching is employed to store popular content and basic computational tasks [10]. The result based on cached outcomes shows that the network determines the best way to allocate computational resources globally to minimize latency. However, optimizing mobile edge computing in this context faces various challenges, including determining where to place computations, allocating computational resources, assigning computing tasks, minimizing end-to-end latency, and reducing energy consumption for the devices involved. Artificial Neural Networks (ANNs) can ease the design of mobile edge caching and computing mechanisms. Their predictive capabilities, particularly in understanding user behaviors such as content requests, make them indispensable for addressing challenges like optimal cache placement and updates. Essentially, ANNs contribute significantly to three key aspects of mobile edge caching and computing applications. Firstly, they excel in prediction and inference tasks, estimating user content requests and frequencies, and by discerning content distribution and request patterns, ANNs guide decisions on what content to store at end-user devices or Base Stations (BSs). Moreover, ANNs uncover social information from data, recognizing users' interests, activities, and interactions. By understanding these patterns, the precision of predicting future events, such as user locations, visited cells, and requested content is improved. Secondly, ANNs serve as efficient clustering algorithms to classify users based on activities, enabling operators to optimize content storage and usage. For example, as user content requests evolve over time, ANNs including Convolutional Neural Networks (CNNs), can be used to categorize these large number of content requests from users. This process of categorization assists in efficiently managing the content stored in the cache, optimizing the retrieval of frequently requested content. Additionally, ANNs are useful for planning when to use computer resources, reducing delays. ANNs can forecast how different users request content and then group those with similar habits: this combined approach makes both prediction and grouping more effective. However, using ANNs for mobile edge caching and computing, comes with challenges including data cleaning (to extract useful data), limited memory and being capable of completing the training process in real-time.

BS switching is a recent solution to cope with the increasing demand for data traffic through a strategic placing of multiple Base Stations (BSs) to expand coverage and capacity [8]. However, maintaining this type of operation also can consume a remarkable amount of energy. To address this energy consumption issue, turning off unnecessary BSs, a process known as BS switching, is seen as a promising solution. Unfortunately, traditional approaches for BS switching come with some limitations: some approaches do not consider the costs associated with the constant turning on and off, while others assume constant traffic loads, which is not realistic. Also, these

methods rely on precise knowledge of the environment beforehand, which is challenging to collect. Leveraging RL-based approaches instead, the system sees the current traffic situation (defined as the state) and decides whether to turn on or off a BS (defined as the action), basing the decision on a set of rules. These rules are adjusted over time as the system learns from its actions: if a decision leads to less energy usage, the system will more likely make a similar decision in the future; if, on the other hand, leads to the use of more energy, it is less likely to repeat that decision. Over time, the learning process continues, and the system figures out the best strategy for turning BSs on or off thus reducing energy consumption.

Transparency can be achieved through explainable AI (XAI) methods, whose duty is to explain why and how a specific decision was reached. It is an approach adaptable across various AI techniques including supervised learning, reinforcement learning, machine reasoning and so on. Recognized as a noteworthy feature, XAI plays a crucial role in the practical implementation of AI models, ensuring that users understand AI decision-making processes. In fields like telecommunications, standardization bodies, such as ETSI and IEEE, emphasize the importance of XAI for establishing trust in intelligent communication systems [11].

In environments that blend physical and virtual realities, extended reality (XR) devices are paramount. Specifically, XR devices enable users to engage with digital twins in real-time, which create digital counterparts that mirror the real-world environments allowing users to visualize detailed data, simulate interactions and forecast the environment evolution in a virtual space. If applied to digital twins of machines, this integration facilitates a more profound grasp of the inner workings of machines and assists in predicting their actions, aiding in monitoring and optimizing the performance of the physical machine. When combined with Explainable AI (XAI), which provides transparency into the decision-making process of AI systems, XR contributes insights into the logic behind the decisions made by machines.

4

Cybersecurity risks and Regulatory framework

4.1 The importance of cybersecurity mechanisms

As the Internet becomes deeply intertwined with our daily lives, its impact on how we learn and work is substantial. However, this increased connectivity also exposes us to more serious security threats. Thus, a pressing challenge that requires urgent attention is to properly identify and analyze the variety of network attacks, especially those that are novel and unfamiliar, to design effective countermeasures. Cybersecurity serves as a comprehensive defense mechanism, employing a range of technologies and processes to protect computers, networks, programs, and data from potential threats, therefore ensuring confidentiality, integrity, and availability of resources and assets in the cyberspace [19]. The overarching goal is to prevent unauthorized access, thwart alterations, and avert destruction [20]. In today's dynamic digital landscape, addressing the evolving nature of cyber threats is indispensable for ensuring the security and integrity of our interconnected systems. The number of cyberattacks is growing dramatically, as well as their complexity [19]. Conventional cybersecurity systems encounter challenges in detecting these sophisticated attacks and often fall short in ensuring user privacy.

Initial attempts to enhance cybersecurity systems capabilities utilized ML techniques [21], however, these approaches struggled to identify diverse threats and intrusions, particularly unforeseen and unpredictable attacks. In particular, ML-based cybersecurity systems relied on various strategies such as Intrusion Detection Systems (IDS), user authentication, data encryption, firewalls, and anti-virus software to safeguard users and devices. Intrusion Detection Systems, as described in reference [22], were specifically designed to identify malicious network traffic, abnormal behaviors, and intrusion attempts within computer systems. There are two main types of IDSs based on their deployment: host-based IDSs and Network-based IDSs. The first one monitors individual hosts, detecting activities such as unauthorized modifications to system files or configuration changes and when such malicious activities are detected, the user is promptly alerted. The second

one checks for anomalies in network traffic and is typically positioned at network nodes such as routers or gateways. Furthermore, IDSs can be categorized based on the method used for intrusion detection. Signature-based detection systems, also known as misuse-based, identify known attacks by matching patterns defined for malicious network activities. While offering high accuracy, they struggle to detect novel attacks. On the other hand, anomaly-based detection systems aim to recognize unknown attacks by defining normal and anomalous behavior patterns. Although they provide a more adaptive approach, they often lack precision. Both signature-based and anomaly-based detection systems initially heavily relied on classical ML techniques [19]. However, these approaches faced limitations such as manual feature engineering, low detection rates [19], and inefficiency in identifying small variants of existing attacks. As hacking incidents became more sophisticated, traditional ML techniques proved inadequate in detecting complex attacks, unknown malware, and ensuring user privacy.

In response to these challenges, the focus of cybersecurity research has shifted toward DL techniques. Over the past few years, the progress in DL techniques has been remarkable, outperforming humans in various tasks. Unlike traditional ML, DL reveals multiple levels of features automatically, orchestrating them at various levels to generate outputs. Its notable advantage lies in the automatic extraction of features, eliminating the manual labor associated with generating feature representations. Also, DL's self-learning capability enhances application processing speed and accuracy. Recognized for its success in diverse domains such as image processing, speech recognition, gaming, and bioinformatics, DL has gained widespread acclaim [19]. As presented above in this thesis, both academia and industry are leveraging DL across a broader spectrum of applications due to its enhanced accuracy in intricate tasks, facilitated by advancements in hardware and software. In parallel, DL techniques are gradually entering the realm of security to augment cybersecurity systems. However, despite its widespread adoption, the security of deep learning systems themselves is at risk of being threatened by meticulously crafted adversarial examples, which are inputs or data intentionally designed to deceive machine learning models [23]. These malicious examples, often indistinguishable from legitimate ones to the human eye, can cause the model to misclassify outputs exploiting the vulnerabilities of DL models, such as small changes in the input data, posing a serious threat to systems. As a result, ensuring the robustness of deep learning algorithms against such adversaries has become paramount. However, there is a shortage of robust countermeasures that can universally address various attack scenarios and contribute to the design of resilient deep learning systems.

4.2 Deep Learning methods for attacks

The attacks against telecommunication networks that DL methods are used to overcome can be classified into three main areas: infrastructure, software and privacy [19], as summarized in Figure 4.1 and hereafter explained.

Within the infrastructure domain, encompassing intrusion and anomaly attacks at the network level, researchers have turned to DL methods for cyber-attack detection [19]. Particularly, DL methods, by discerning patterns that deviate from normal behaviors (in the context of anomaly-

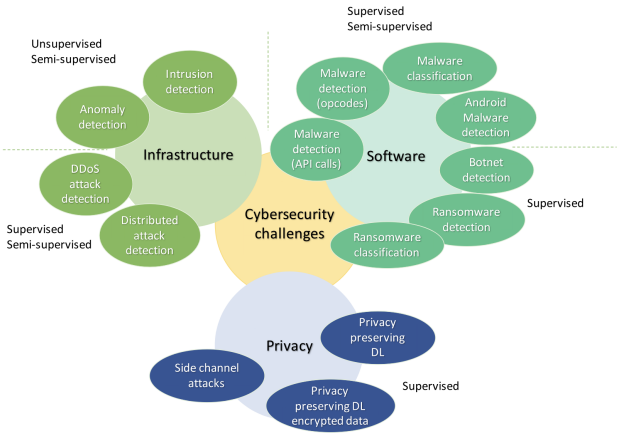

Figure 4.1: Cybersecurtity challenges tin mobtile networks [19].

based network intrusion detection) [19], play a pivotal role in thwarting attacks. This context of cyberattacks is comparable to image recognition: new cyberattacks often resemble small mutation of the existing ones, such as the recognition of images through a small change in the pixels. Thus, DL much like discerning changes in images, possesses the ability to autonomously grasping and generalizing patterns enabling the detection and the addressing of future threats.

In the software domain, DL finds application in detecting malware, ransomware and botnets [19]. Malicious software, commonly known as malware (including viruses, worms, trojans, and ransomware), is created and distributed by attackers with the intent to carry out various security breaches. These attacks may involve the theft of users' private information, the remote hijacking of devices to unleash extensive spam emails, and the infiltration of online account credentials [20]. Among various types of malware, botnets stand out as particularly dangerous due to their unique characteristics. Unlike conventional malware, botnets operate beyond predictable algorithms, making them especially challenging to manage. Their primary goal is to infect a diverse range of devices, staying in an active and undetectable state for as long as possible [19]. What characterizes botnets is their complicated structure: they are controlled by human operators through either command and control (C&C) servers or peer-to-peer (P2P) networks, and unlike typical malware that follows predefined patterns, botnets exhibit a more complex behavior, with different designs that vary from one another. Lastly, ransomware, a form of malicious software, discreetly infiltrates a user's device with the objective of extorting a ransom (sum of money or some other form of payment demanded by a person or group in exchange for the release of something or someone), typically in cryptocurrencies like Bitcoin, in exchange for restoring compromised resources. This insidious software is engineered to infect, encrypt, and restrict access to system files, effectively holding the host system hostage [19]. These mentioned threats have rapidly evolved to evade signature-based solutions, such as antivirus software, complicating the defense mechanism. New malwares typically involve slight modifications to existing ones, with attackers improving infection mechanisms, obfuscation, or payloads. Consequently, malicious applications within the

same family share strong similarities in code and behavior. This prompts the integration of DL in cybersecurity systems for malware detection and classification, and botnet and ransomware detection [19].

Finally, DL methods are instrumental in safeguarding user privacy. Through the services offered by mobile devices, such as recommendation systems, targeted advertising and health monitoring, highly sensitive personal data are collected, including also personal information, photos, videos and even banking details. Moreover, these services often access and utilize data from external sources, such as surveillance systems or medical information, which further increases the sensitivity of the information involved. Initial efforts [19] focused on preserving user privacy across various types of Neural Networks (NN). To mitigate the risk of private information leakage, NNs are subsequently trained with techniques like differential privacy. Recent advancements [19] take this commitment further, employing encrypted data in NN training, that results in a strategic measure to add an extra layer of protection.

4.3 Threatening deep learning techniques

In this intricate cybersecurity landscape, adversaries may attempt to manipulate DL techniques, either the collection or processing of data to compromise the target model, thus tampering with the original output. There exists three types of possible attacks: evasion attacks, poisoning attack and exploratory attack [23]. The evasion attack, which occurs during the testing phase, involves the evasion of system's defenses by making subtle adjustments to malicious samples. The key objective is to manipulate the system's behavior in a way that allows malicious inputs to go undetected or misclassified. Here the focus is on exploiting the vulnerabilities in the testing phase, without directly tampering with the data used for training [23]. On the other hand, the poisoning attack involves the contamination of the training data during the testing phase. The adversary aims to compromise the learning process by deliberately introducing meticulously crafted samples into the training data: the goal is to poison or corrupt the model's understanding of patterns and features, ultimately leading to a compromised or inaccurate machine learning model. So, unlike evasion attacks, poisoning attacks manipulate the training data itself, influencing the model's development and undermining its overall performance [23]. An exploratory attack instead, is characterized by its non-influence on the training dataset and occurs when an adversary is granted black box access to the model, which means that the adversary in this case, is granted access to the functionality or output of a system without the visibility into its internal workings, thus he can interact with and observe the outputs of the model without knowing the specifics of the model's architecture, parameters, or training data. In this type of attack therefore, the focus is on acquiring maximum knowledge about the learning algorithm employed by the underlying system and discerning patterns within the training data. Unlike poisoning attacks, which manipulate the training data, exploratory attacks involve probing the model's behavior and structure without directly altering the training dataset [23]. Concerning the black box, it is worth noting that there are potential attacks even on this enigmatic system. The attacks possible operate under the assumption that the attacker has no knowledge about the internal workings, structure, or

parameters of the target model. Instead, the attacker relies on information about the model's settings or past inputs to assess and exploit potential vulnerabilities: in a black-box attack, the attacker treats the model as an opaque system, they don't have insight into how the model processes information or makes decisions. Black-box attacks can be grouped into three categories: non-adaptive black-box attack, adaptive black-box attack and strict black-box attack [23]. The first one involves creating a simplified version of the target model based on its training data, crafting deceptive inputs for this simplified model, and then using these inputs to mislead the actual target model into making incorrect predictions: the attack, therefore, leverages knowledge of the training data but doesn't require a deep understanding of the intricate details of the target model. In an adaptive black-box attack instead, the adversary queries the model for information, uses that information to train a surrogate model, and then applies white-box attack strategies to generate inputs that mislead the target model. In this case, the attack leverages the ability to interact with the target model as an oracle without knowing how it was originally trained. Finally, a strict black-box attack involves collecting pairs of inputs and their corresponding outputs from the target classifier without the ability to modify inputs dynamically.

4.4 Regulations

In today's fast technological landscape, it is imperative to establish effective regulations for artificial intelligence to ensure responsible and safe utilization of this technology. As an integral component of its digital strategy, the EU aims to establish regulations to oversee how artificial intelligence (AI) is developed and used. AI holds the potential to bring numerous benefits, including advancements in healthcare, enhanced safety in transportation, more efficient manufacturing processes, and the promotion of affordable, sustainable energy. The European Commission, in April 2021, introduced the initial regulatory framework for AI within the EU, which involves the analysis and classification of AI systems applicable across diverse applications. Based on the level of risk they pose to users, the extent of regulatory measures applied will be determined. The regulatory proposal, also, aims to provide clear guidelines for AI developers, users, and deployers while also minimizing administrative and financial burdens, especially for small and medium-sized businesses [24]. Once approved, these regulations will mark the world's first comprehensive rules governing AI [25]. The AI Act, therefore, introduces a graduated approach to regulation, tailoring rules based on the risk levels associated with different artificial intelligence (AI) systems, and the extent of obligations for both providers and users varies according to the perceived risk. AI systems are considered of unacceptable risk when they constitute a threat to people, therefore, they will be prohibited as stated in Title II (Article 5) of the proposed Act [26]. For example: the introduction, deployment or utilization of AI systems employing harmful manipulative "subliminal techniques"; the application of AI systems targeting specific vulnerable groups (whether dealing with physical or mental disabilities); the use of AI systems by public authorities or on their behalf for social scoring purposes; and the deployment of "real-time" remote biometric identification systems in publicly accessible spaces for law enforcement objectives, with only a few exceptions allowed (such as post remote biometric identification systems) [25]. High risk is intended, instead, when AI sys-

tems have the potential to negatively impact safety or fundamental rights, as regulated by Title III (Article 6) of the Act [26]. These will be classified into two categories: AI systems integrated into products governed by the EU's product safety legislation (comprehending toys, aviation, cars, medical devices, and lifts); and AI systems falling within eight specific domains necessitating registration in an EU database. These areas, specifically, include biometric identification and categorization of natural persons, management and operation of critical infrastructure, education and vocational training, employment, worker management, and access to self-employment, access to and enjoyment of essential private services and public services and benefits, law enforcement, migration, asylum and border control management, and assistance in legal interpretation and application of the law. With the clarification that all these AI systems will undergo assessment before entering the market and will continue to be evaluated throughout their lifecycle [25]. AI systems categorized as having limited risk are expected to meet basic transparency standards, ensuring users to make informed choices. Following interaction with these applications, users have the option to decide whether to proceed. Users should be informed when engaging with AI, particularly in cases involving the generation or manipulation of image, audio, or video content, such as deepfakes [26]. Finally, for AI systems categorized as presenting low or minimal risk, no additional legal obligations are imposed. These systems can be developed and utilized within the EU without specific requirements, with the exception of the codes of conduct established by the Act. Essentially, these guidelines are designed to encourage providers of non-high-risk AI systems to voluntarily follow the mandatory rules and standards applicable to high-risk AI systems [26]. Due to the rapid evolution of artificial intelligence (AI) technology, the proposed regulations take a "future-proof" approach: the rules are, in fact, designed to adapt to change in technology over time [24]. The main objective is to ensure that AI applications maintain trustworthiness even after they have been introduced to the market, and achieving this requires continuous quality and risk management by the providers of AI technology.

5

Conclusion

This thesis has delved into the intersection of Artificial Intelligence (AI) and Telecommunication Networks, shedding light on key insights such as technological complications, security threats and regulatory obligations. This research mainly aimed to develop the challenges and the potential solutions in incorporating AI into the realm of telecommunication networks. Through a comprehensive exploration of Artificial Intelligence, which includes its subsets Machine Learning (ML) and Deep Learning (DL), we have navigated the complexities and arrived at a complete understanding of the subject and the evolution of these technologies. The process of examining AI, Machine Learning (ML), and Deep Learning (DL) including Neural Networks (NN) architectures revealed their intricate relationships and connections, and while they may appear similar, they are distinct and have different roles and systems. Specifically, NN architectures have gained prominence and importance in the field of AI and are being used for a variety of applications. Exploring the intersection of AI and telecommunications revealed a landscape rich in both opportunities and obstacles. Discovering the benefits, ranging from improved efficiency and performance to security and flexibility, was enlightening. However, dealing with lots of data and making sure that computer programs have enough power is a realistic challenge. The development of solutions, such as federated learning and mobile edge computing provided a view of how intelligent networks can change the systems. As we looked deeper into the world of cybersecurity risks and the regulations that need to be followed, we faced the reality of the downsides of this big technological advancement. The key focus is, in fact, realising the importance of applying strong regulatory frameworks, making sure that privacy is respected and ensuring the ethical use of AI: at this point, it became clear that technology and law have to work together to build a strong and lasting system.

This thesis should represent a moment of reflection, not a definitive endpoint because bringing together technological processes, legal structures and the push for responsible AI, exposes gaps that require attention and need further investigations. Nevertheless, given that the utilisation of AI in wireless communications is still in its early stages, several challenges remain open. Addressing

issues like the need for infrastructure updates to support the implementation of machine learning based paradigms, is crucial to foster future research. The lesson learned is that the fusion of AI and telecommunications is not a one-time achievement but rather an ongoing evolution that requires careful work, and recognizing this leads to the understanding that every technological leap carries responsibilities, emphasising the imperative role of ethical considerations.

References

[1] E. Burns, "Artificial Intelligence (AI)," *TechTarget*, 2023.

[2] I. Asimov, *I, ROBOT*. Oscar Mondadori, 2003.

[3] A. Schroer, "Artificial Intelligence," *Builtin*, 2023.

[4] S. Barbarossa and A. Zanella, *Machine Learning and 5G/6G Networks: Interplay and Synergies*, 1st. CNIT-Consorzio Nazionale Interuniversitario per le telecomunicazioni, 2021.

[5] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Introduction to Machine Learning, Neural Networks, and Deep Learning," *Translational Vision Science & Technology*, vol. 9, no. 2, Feb. 2020. eprint: `https://arvojournals.org/arvo/content\_public/journal/tvst/938366/i2164-2591-226-2-2007.pdf`.

[6] P. P. Shinde and S. Shah, "A Review of Machine Learning and Deep Learning Applications," pp. 1–6, 2018.

[7] I. Data and A. Team, "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the difference?" *IBM*, 2023.

[8] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wire-less networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.

[9] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine Learning for 6G Wireless Networks: Carrying Forward Enhanced Bandwidth, Massive Access, and Ultrareliable/Low-Latency Service," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 122–134, 2020.

[10] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.

[11] V. Berggren, R. Inam, and L. Mokrushin, "Artificial intelligence in next-generation connected systems," *Ericsson White Paper*, Sep. 2021.

[12] Z. Lemaster, T. Kinnman, and A. Laya, "How is intelligence transforming telecom? Five benefits that reveal the full value of AI," *Ericsson Blog*, Mar. 2023.

[13] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for mimo detection," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1702–1715, 2020.

[14] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive mimo for hybrid precoding," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 3027–3032, 2019.

[15] C. Fan, B. Li, C. Zhao, W. Guo, and Y.-C. Liang, "Learning-based spectrum sharing and spatial reuse in mm-wave ultradense networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 4954–4968, 2018.

[16] M. Abbasi, A. Shahraki, M. Jalil Piran, and A. Taherkordi, "Deep reinforcement learning for qos provisioning at the mac layer: A survey," *Engineering Applications of Artificial Intelligence*, vol. 102, p. 104 234, 2021.

[17] L. Baier, F. Jöhren, and S. Seebacher, "Challenges in the deployment and operation of machine learning in practice," May 2019.

[18] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2022–2035, 2020.

[19] E. Rodríguez, B. Otero, N. Gutiérrez, and R. Canal, "A survey of deep learning techniques for cybersecurity in mobile networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1920–1955, 2021.

[20] Y. Xin, L. Kong, Z. Liu, *et al.*, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35 365–35 381, 2018.

[21] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.

[22] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Springer Link*, vol. 22, 2017.

[23] C. Anirban, A. Manaar, D. Vishal, C. Anupam, and M. Debdeep, "Adversarial attacks and defences: A survey," 2018. arXiv: 1810.00069 [cs.LG].

[24] "Regulatory framework proposal on artificial intelligence." (Oct. 2023), [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

[25] "Eu ai act: First regulation on artificial intelligence." (Jun. 2023), [Online]. Available: https://www.europarl.europa.eu/news/en/headlines/society/20230601ST093804/eu-ai-act-first-regulation-on-artificial-intelligence.

[26] M. Tambiama, "Artificial intelligence act," *EPRS|European Parliamentary Research Service*, Jun. 2023.