REPUBLIC OF TÜRKİYE
ALTINBAŞ UNIVERSITY
Institute of Graduate Studies
Electrical and Computer Engineering

A SECURE 5G DATA COMMUNICATION PROTOCOL BASED NOVEL AI TECHNIQUES

Ahmed Naif Hadi ALJBOURI

Master’s Thesis

Supervisor

Asst. Prof. Dr. Sefer KURNAZ

Istanbul, 2023

A SECURE 5G DATA COMMUNICATION PROTOCOL BASED NOVEL AI TECHNIQUES

Ahmed Naif Hadi ALJBOURI

Electrical And Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled A SECURE 5G DATA COMMUNICATION PROTOCOL BASED NOVEL AI TECHNIQUES prepared AHMED NAIF HADI ALJBOURI and submitted on 14/04/2023 has been **accepted unanimously** for the degree of Master of Science in Electrical and Computer Engineering.

_____

Asst. Prof. Dr..Sefer KURNAZ

Supervisor

Thesis Défense Committee Members:

Asst .Prof. Dr. Sefer KURNAZ	Department of Computer Engineering and Natural Sciences, Altinbaş University	_____
Asst. Prof. Dr. OĞUZ KARAN	Department of Computer Engineering and Natural Sciences, Altinbaş University	_____
Asst. prof. Dr. Zeynep ALTAN	Department of Engineering and Architecture Beykent University	_____

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

Submission date of the thesis to Institute of Graduate Studies: ____/____/____

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Ahmed ALJBOURI

Signature

DEDICATION

First, I would like to thank Allah Almighty for his grace in helping me and guiding me on the path of knowledge in this study

ABSTRACT

A SECURE 5G DATA COMMUNICATION PROTOCOL BASED NOVEL AI TECHNIQUES

ALJBOURI, Ahmed

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Sefer KURNAZ

Date: April / 2023

Pages: 58

With the development of the Internet and communication technology, the number of smart devices that can connect to the Internet is increasing day by day. These devices do not have the same security mechanisms that computers and servers have. Today, 5G networks are becoming increasingly popular for monitoring different types of critical environments. Due to the large number of 5G devices, the security of these networks and devices is an important concern. In this study, new method-based machine learning techniques presented to detect attacks in the 5G data communication environments. The proposed method combined unsupervised and supervised techniques to detect attacks. The data first analyzed using Boltzmann machines which high level features extracted. Then, the extracted features wired to the AdaBoost that classify the extracted features to the normal and abnormal classes.

Keywords: AdaBoost, AI, SVM, MLP, Communication Protocol.

TABLE OF CONTENTS

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

SVM : Support vector machine

NN : Neural Network

RBF : Radial Basis Function

ANN : Artificial-Neural-Network

1. INTRODUCTION

5G is the fifth generation of wireless cellular technology, offering faster upload and download speeds, more stable connections, and improved capacity than previous networks. 5G will be much faster and more reliable than the 4G networks that are prevalent today, and could change the way we use the internet to access apps, social networks, and information. Technologies that require fast and reliable data connections, such as self-driving cars, advanced gaming applications, and media streaming, will greatly benefit from 5G connectivity [1].

With the advent of new technologies such as artificial intelligence, the Internet of Things (IoT) and automation, the need for Internet access is driving a huge increase in the amount of data generated. Data creation is growing rapidly and is expected to grow by hundreds of zettabytes over the next decade. The existing mobile infrastructure is not designed for such an information load and needs to be upgraded. At the same time, with its high speed, huge capacity, and low latency, 5G can help support and scale various applications such as cloud-connected traffic management, drone delivery, video chat, and console gaming on the go. From global payments and emergency response to distance learning and a mobile workforce, the benefits and applications of 5G are endless. 5G. It can change business, the global economy and people's lives [2].

With the development of technology, the number of devices connected to the Internet also increases. There are many definitions in the literature on the Internet of Things. With the definition made by ITU, IoT is a technology where any object can be connected to each other at anytime, anywhere. British scientist Kevin Ashton used the term "Internet of Things" for the first time in 1999 in a presentation prepared for the company Procter & Gamble. In this presentation, the benefits of RFID technology are explained and the RFID effects of the starting point of IoT are accepted. Then, in 2005, the first report of ITU on the subject was published. IoT would perceptually connect objects in the world by bringing together technological developments such as item identification, sensor and wireless sensor networks, embedded systems and nanotechnology. IBM's CEO, Samuel J. Palmisano, suggested the concept of Smart Planet. At the CASAGRAS (2010) conference held in October 2009, the need for integration with the developing internet, the use of internet protocol in the sense of being a stable, stable and scalable communication technology, and the use of IPv6 were

emphasized [3]. IoT technology has made significant advances since its first introduction. In IoT technology, every object can be detected via detection methods such as RFID, NFC, sensors, and WiFi, Wimax, Zigbee, Bluetooth, infrared, etc. With wireless communication techniques, information about objects can be obtained. Thanks to this, many conditions of objects such as temperature, pressure, light, distance can be observed through sensors, and objects can be decided according to the conditions. Objects can store or share this information and states. In addition, remote control and monitoring of all electronic devices is possible thanks to IoT. IoT systems have added communication capabilities to physical devices to benefit their users. The market share of IoT applications continues to increase as it is seen that the use of this communication in production provides benefits in terms of cost. Internet of Things technology is used in many areas such as environment and infrastructure monitoring, industrial applications, energy management, medical services, building automation, transportation. The increasing number of devices in the Internet of Things and the expansion of the usage area also bring some threats. Cyber-attacks, unauthorized modification and destruction of information systems are defined as any threat to the integrity of information. Intrusion Detection Systems, on the other hand, are systems that identify malicious network activity to help ensure computer security. As the number of devices connecting to the Internet increases, the number of threats also increases. The increasing number of threats makes computer security a highly important task. Anti-virus systems, firewalls, etc. to combat cyber-attacks. Various hardware and software solutions are being developed. One of these solutions is Intrusion Detection Systems (Çakır, 2019). The purpose of intrusion detection; to detect unauthorized and misuse of computer systems by both insiders and outsiders. In general, IDSs use statistical, anomaly, and rule-based abuse models to detect intrusions.

Mankind has needed to make various inventions since ancient times in order to make his life easier and more livable. These inventions have come about as a result of the inferences made by human beings from the period in which they lived. An example of this situation is a community that goes hunting to meet the food needs of its tribe. The community will be inexperienced on their first hunt and there will likely be many casualties from the hunting team. But the survivors will be better prepared for the next hunt. For example, they will develop cutting tools, invent clothing that protects their bodies, and act more organized while hunting. This situation has led to the fact that human beings are intelligent beings and to

learn various lessons from the positive or negative situations in which they live and to improve themselves. Human beings, who learn and can teach what they have learned to other people, have experienced throughout history and have formed a common memory with the accumulation of these experiences. This common memory has made a great contribution to the development of civilization. For example, the primitive hunter who wanted to meet his tribe's need for food and protection has turned into a sophisticated soldier with a machine gun. Primitive people, who found thick and durable clothes to protect themselves from the attacks of wild animals and other tribes, today meet this need in steel vests and so on. meets the equipment. Civilization has experienced a great development after events that significantly affected the future of the world, such as the French Revolution, the Renaissance and the industrial revolution. These developments have increased production in addition to fields such as science, culture and art, and thus, the country's economies have improved significantly. Before these developments, people who had the desire to think and invent something, took advantage of this environment of freedom and prosperity and embodied their thoughts and made inventions that laid the foundations of today's world. The rush of machine production and digitalization from the early 1900s to the present has significantly increased the welfare level of humanity compared to previous centuries. During this period, various mass production machines were invented to speed up production, and computers in this period were influential in every field in a short time. Especially after the invention of the computer, digitalization has accelerated significantly and the computer industry has come closer to the other sectors. In this direction, although it was initially made for computer production, army needs and accounting works, then the computer became important in terms of entertainment and information as well as these works. Digitization has reduced production and labor costs in all sectors and made the capital sufficient for the whole world. So much so that production was slow and expensive in the times when handwork was important, but with digitalization, the production speed has increased and this has reduced costs. Thus, this process accelerated over time and was adopted by the capital owners. Today, however, this process evolves to a different point. To artificial intelligence. All these developments have revealed new technologies and as a result, data piles have emerged at a rate that human beings cannot cope with. Especially with the development of communication technologies, the emergence of the Internet and the emergence of pocket computers called smart phones, there is a data production that the world history has not seen before. These data are

undoubtedly insignificant for people, but for capital owners, each of these data means money. In this direction, artificial learning algorithms that work faster than the human brain are needed to solve this data. These algorithms are faster, more accurate and less expensive than the human brain. These algorithms are based on biological learning processes, and they learn through the experiences they get from the data. At this point, it is obvious that artificial intelligence learning will provide good services to humanity, but the following situations should not be ignored. Artificial intelligence can learn from good data as well as from bad quality content with malicious intent. This situation may reveal concepts such as good artificial intelligence and bad artificial intelligence, and it is necessary to foresee that various damages can be caused to humanity by artificial intelligence. In this sense, while humanity uses the positive aspects of artificial intelligence, it should also find ways to protect itself from the negative aspects and show the necessary care about this issue [5].

2. OVERVIEW

2.1 INTERACTIVE WEB

The concept of Web 2.0 was first introduced by Tim o'Reilly at a conference in 2004. Tim o'Reilly explains Web 2.0 as follows: "Web 2.0 is a concept that emerged after the advent of the Internet to understand the developments in computer technology and computerized manufacturing. To develop this notion12, more people need to use Web 2.0 and develop applications." Web 2.0, also known as wisdom web, human-centered web, participatory and literacy web, has an interactive structure because it allows two-way communication. Although Web 2.0 users do not encounter any restrictions on the Internet, the page manager in Web 1.0 is now users. With Web 2.0, it has become easier to keep content and pages up-to-date. Moreover, Web 2.0 prioritizes collective intelligence as it makes it possible to work collectively over the Internet. There are some terms that emerged with Web 2.0; are words like wiki1, mashup2, tagging3. Again, social media pages emerged in this period [6]. There are also aspects of Web 2.0 that negatively affect people. For example, when you make a post on the interactive web, this share may be copied by malicious people, or an editor who prepares content for the web may be lynched by people for the content he has prepared.

2.2 THE WEB OF ALGORITHMS

Web 3.0 was invented by the inventor of the web, Tim Berners-Lee, but in 2006 a journalist named John Markoff suggested the name 'Web 3.0'. Web 3.0 connects existing datasets to create new data and creates new datasets. The main purpose of Web 3.0 is to make the data shared on the web readable not only by humans but also by machines and to get rid of the limited structure of web 2.0. Web 3.0 elevates users' experience to the next level. It is also known as the 'semantic web' because of its ability to analyze user data. The biggest difference with Web 2.0 is that it is algorithm and machine based. Companies such as Google, Microsoft, Facebook and Amazon are investing in the web 4.0 era by using deep learning and machine learning algorithms. For example, Microsoft company's 'cortana' assistant, Google company's 'Google Assistant' and Apple's 'siri' are developing software based on artificial intelligence learning, and these assistants can do more and more successful jobs. As long as you don't turn off the computer and your smartphone doesn't run out of battery,

the assistants listen to the environment and learn from the sounds they detect. In this way, they offer people a more personalized web experience. In addition, assistants, who monitor the health status of people with the data collected from smart watches, can call for help in various accident or disaster situations. Personalization for Web 3.0 has an important point, that the interests of users are directly followed and analyzed by the third version of the web. In short, Web 3.0 semantically combines disjointed web pages and then helps users by drawing conclusions from the collected data. Communication, which has become two-way in Web 2.0, has turned into a process that includes machines and algorithms with web 3.0. Web 3.0 has opportunities for users, but constant surveillance on the web is its worst part. Web 4.0, also known as the symbiotic web, refers to the interaction between humans and machines in a state of symbiosis. Mind control becomes possible in the Web 4.0 era. In other words, machines will have the ability to read the content on the web and to decide on the next steps to be taken. With Web 4.0, the web will work like an operating system. Companies such as Google, Microsoft, Facebook and Amazon are investing in the web 4.0 era by using deep learning and machine learning algorithms. For example, Microsoft company's 'cortana' assistant, Google company's 'Google Assistant' and Apple's 'siri' are developing software based on artificial intelligence learning, and these assistants can do more and more successful jobs. As long as you don't turn off the computer and your smartphone doesn't run out of battery, the assistants listen to the environment and learn from the sounds they detect. In this way, they offer people a more personalized web experience. In addition, assistants, who monitor the health status of people with the data collected from smart watches, can call for help in various accident or disaster situations [7].

2.3 SURVEY

Ning et al. [8] analyzed insider attacks on MANETs using the AODV routing protocol. Studies involving attackers abusing redirect messages are simulated using the Ns-2 simulation tool. The experimental results show that attacks can easily disrupt the network, hijack routes, isolate non-aggressive nodes from the network, or consume network resources. The systematic analysis of attacks and the applicability of this analysis scheme to other protocols contribute to the literature. Studies that perform black hole attack analysis [9] also develop the AODV routing protocol. Jain et al. [10] analyze black hole attacks using the

AODV routing protocol. In line with their experimental results, they developed the AODV routing protocol and proposed a new protocol called SAODV. In this protocol, the assumption is that the first incoming RREP message comes from the attacking node. For this reason, a method has been developed that ignores the first incoming message. The performances of AODV and SAODV protocols under black hole attack have been compared and it has been shown that SAODV is more secure against attack. Dokurer et al. [11] not only analyzes black hole attacks, but also offers solutions for black hole attacks by improving AODV. Using the Ns-2 simulation tool, 100 different scenarios with 20 nodes were created. In these scenarios, a new method is proposed by making performance evaluations under a black hole attack and without an attack. In the proposed method, the source node assumes that the first incoming RREP control message is sent by the attacker node, and waits for a second RREP message, and if the message comes, data transfer is provided over the new route. If there are RREP messages sent third, this route is also preferred. However, no difference in performance was observed between the second route 8 and the third route. With the proposed modified AODV protocol method, it is shown that packet losses are reduced by 19% in a network under attack. Using this method when the network is not under attack increases packet loss by 4%. According to the results, with the newly developed method, the network performance under attack is better. Lu et al. [12] simulated black hole attacks against MANETs using the AODV protocol. A safer and more effective protocol has been presented by developing the AODV protocol. In this protocol, when the source node receives an RREP message, it sends an acknowledgment message to the destination node in the opposite direction of RREP. This verification message sent contains a random number generated by the source node. Verification messages reaching the source node are checked if they contain the same random number and if they contain the same number, validation takes place. If the package does not contain the same number, a second package containing the same number is expected to arrive. A secure message is then transmitted to the source node and the source node checks the random numbers in this packet. Here, too, the method that two packages contain the same number is used. When two or more packets containing the same number are received, the route is judged to be reliable and data transmission is performed. With the new protocol presented, AODV security vulnerabilities were tried to be closed and it was made more secure against black hole attack. As a result of the analysis, it is shown that the newly developed protocol is more secure than the AODV protocol.

Deshmukh et al. [13] propose a secure routing system based on AODV for MANETs. This system was created to detect black hole attack and affected routes in the early stages of route discovery. In the simulated method with Ns-2, a validity value is added to the RREP message and stored in the route table of all nodes on the active route. The RREP validity value will be blank, as the attacker node will send a response when looking at the table. In this case, the RREP message from an attacking node will be dropped. This validity value to be included in RREP will not increase the overhead since it is only a single bit. When the newly developed protocol is used, AODV performances in an attacked network and in a non-attack environment are almost equal. This shows that the proposed method is reliable. In another study [14], the effect of black hole attacks is evaluated on networks using different routing protocols, AODV and OLSR. Network performance under black hole attack and without attack was evaluated according to packet deliver ratio (PDR) and network throughput metrics. According to the results, AODV 9 protocol without attack shows better network performance because it has higher PDR values than OLSR protocol. However, if there is no attack on the network and the number of nodes is less, OLSR routing protocol is recommended for higher efficiency. Despite this proposition, the authors stated that the AODV protocol has a better approach than the OLSR protocol due to its ability to report connection errors and repeat operations. While there is a lot of work in the black hole attack, there is less in the rushing attack [15]. In [16], a rushing attack on arbitrary ad hoc networks is described. This attack prevents the source node from finding routes longer than 2 hops to the destination node during route discovery. It is explained that all the proposed secure optional ad hoc routing protocols are vulnerable to attack. For this reason, a new route-finding RAP (Rushing Attack Protection) protocol has been proposed to prevent the attack. In the proposed method, a randomly selected RREQ message is transmitted among the RREQ messages received in route discovery. In order to prevent the attacker from transmitting more than one RREQ message, a list of nodes it passes into is placed in each RREQ packet. A two-way neighbor authentication is then performed for each connection. Finally, in order to authenticate the node list, each node is requested to verify the RREQ message it transmits. Although the node brings more overhead than classical route finding methods, it is shown to be successful in routing and packet delivery by finding routes that other protocols cannot find. Studies showing the effect of flood attack, which is a denial-of-service attack in MANETs, are given in [17]. Bandyopadhyay et al. [18] investigated the

effect of flood attack on AODV routing protocol in MANETs. A flood attack is a type of attack that aims to initiate a denial of service on the network by a malicious node. It is an attack that aims to reduce network performance by sending large numbers of RREQ packets or unnecessarily large data packets to existing or non-existent nodes in the network. The effects of flood attacks were simulated by sending 8 RREQ packets per second to invalid destination addresses using the Ns-3 simulation tool. In the study, the percentage of packet loss, routing load and bandwidth requirement of the attacks in the network were evaluated. As the fake RREQ messages created as a result of the evaluations increase the routing overhead, the attack increases these values and causes the data packets to drop, thus increasing the packet loss percentage. In addition, when the number of attacking nodes is increased, the average bandwidth usage and packet loss percentage increase. Flood attack has been observed to reduce MANET performance. In [19], the effects of flood attacks on MANET are investigated using the Ns-2 simulation tool. A simulation, which is called the overflow frequency at different rates, is increased from 0 packets per second to 100 packets per second. Packet delivery rate and packet delay times were measured by increasing the number of attacker nodes from 0 to 5. It was observed that the packet delay time first increased and then decreased. The reason for this is presented as the network congestion in the flood attack and the dropping of the packets on the long paths as the remaining packets on the short paths. In addition, when the overflow frequency is increased to 100 packets per second, the PDR drops to 27%. When the attacker rate is increased to 5, PDR becomes 17% when 20 packets are sent per second. The increase in the number of attacking nodes and the increase in sent packets significantly accelerate the PDR speed. Although there are many studies on MANET attack analysis and protocol development in the literature, these studies in FANETs have recently begun to be investigated. When drones are part of a group, such as in ad hoc networks, they can be a potential target of attackers to damage the device or access the data it contains. The impact of such threats targeting the privacy, security and physical integrity of UAVs [20] can seriously affect both the mission of the UAV and its network. Also, multi-UAV communication is exposed to additional threats to secure communication mechanisms. FANETs have a higher level of node mobility and therefore change the network topology more frequently than traditional MANETs. Therefore, studies [21] have been conducted discussing the unique features and challenges of FANETs. Bekmezci et al. [22] discuss the security requirements of FANETs and potential threats to

these high-level networks. In addition, the authors present well-known ad-hoc network attacks and discuss security solutions for such attacks against FANETs. In addition, a new AODV-based routing and power-controlled protocol for FANETs is presented in another study [23] for FANETs. MDRMA (Multi-data rate mobility aware, MDRMA) protocol consists of two algorithms, MDRMA-Forwarding and MDRMA-Power Control. MDRMA-Routing enables the routing layer and MAC sublayer to exchange information to ensure fast and efficient transmission of data. In addition, the transmission power must be controlled for a fast data transmission. The MDRMA-Power Control algorithm, on the other hand, controls the transmission power for the data flows in each connection along the route established between the destination and the source node. The performance of the proposed routing protocol is analyzed in the Ns-3 simulator based on packet delivery rate, overhead, and end-to-end latency metrics. As a result of the study, it has been observed that this new protocol for FANET reduces disconnections in the network by creating fast data transmission and stable paths. Studies for FANETs are limited. Most of the studies were done using the Ns-2 simulation tool. Two-dimensional mobility models were used as the mobility model. These mobility models cannot reflect the 3D movements of FANETs. For this reason, it is difficult to apply simulations to real systems. In addition, velocities are set to a maximum of 30 m/s. This corresponds to a speed of 108 km per hour (108 km/h), which is usually much lower than the speed of UAVs. In [24], different types of attacks against MANETs are analyzed. However, the majority of studies focused on a single attack type. Most of the studies done involve the black hole and flood attack. In this study, the Ns-3 simulation tool was used for more realistic tests and the 3D Gauss-Markov mobility model was chosen. Due to the high real UAV speeds, the speed was determined as 720 km/h (720 km/h). In addition to these, the effects of 4 different attack types on the networks created with 13 different topologies were analyzed by changing the attacker rates between 5% and 25%. Intrusion detection studies were carried out for MANETs. In [25], an intrusion detection system with an artificial neural network using information visualization was developed due to the inadequacy of traditional methods in MANET protection. In addition, the stamping method was used to prevent the maps derived using visualization from being modified by the attacker. This method is created by placing a unique message into the original information. The proposed method showed high performance in different conditions (various traffic conditions, mobility models and visualization measures). In this study, a dataset was created with the

attributes of the MAC layer. This dataset was trained with artificial neural networks and evaluated according to attack detection rate and false alarm rate metrics. And the detection rate is 80% on average, and the false alarm rate is around 20%. Unlike our study, artificial neural networks were applied to the data collected from the visualization and the MAC layer. In our study, 30 different features are used for the ANN input layer. Having a lot of features plays an important role in these algorithms to give accurate results. In addition, our study makes two different classifications based on user and network, and the detection rate is 94% on average and the false positive rate is 9% on average. In another study [26], a wormhole attack was performed on MANETs using Matlab 2019b. In MANET, data was collected using the characteristics and speeds of the nodes. While marking the data, the attacking nodes were specified and a total of 8 attributes were used. Intrusion detection was performed by classifying these data with various machine learning methods (nearest neighbor, support vector machine, decision tree, linear discrimination analysis, naive bayes and convolutional neural networks). The general purpose of this study is to detect attacking nodes among normal nodes. Among the methods used, the decision tree made the best classification with 92.6% sensitivity (detection rate). Laqtib et al. [27] stated that MANETs cannot be protected from new threats with traditional encryption methods and therefore deep learning methods should be applied in intrusion detection systems. They have worked to adapt MANETs to volatile environments and to detect intrusions with high accuracy. In their study, they applied various deep learning methods to the NSL-KDD dataset published in 1999 and obtained high accuracy rates. There are several recommended security solutions for FANETs in the literature [28][29]. Another solution proposed by Bhatia et al. [30] consists of monitoring, detection and isolation steps to identify malicious nodes that triggered the sybil attack. In the monitoring step, the activities of the nodes in the network are observed in a central controller. This central system memorizes the identities of the nodes participating in the network, and if the node changes its identity, the central system runs an authentication mechanism. When the identities differ, the node is marked as an attacker and the node is isolated from the network. In addition to all these studies, artificial intelligence and machine learning studies are limited in the literature, as there is no published dataset for FANETs. In a study using traditional methods [31], a hybrid method is proposed for FANET intrusion detection. The proposed method consists of two steps. First, spectral analysis is used to generate a specific traffic signature that provides basic information about the type of intrusion on the network.

Second, the output of the first step and the controller/observer-based prediction step evaluates the level of attack observed in the network. It has been observed from the results that the attacks were detected quickly and accurately. The delay in detection is negligible, and intrusion traffic and normal traffic can be accurately separated. Other studies [32] are attack detection using machine learning and artificial intelligence algorithms. Ouiazzane et al. [33] proposed a new intrusion detection model for UAV fleets. In this study, a system with a multi-agent paradigm has been developed that constantly listens to the network and gives a warning when anomaly is detected. The machine learning method was used to teach the normal reference profile, and if a deviation from the model is found, it is perceived as an attack. First of all, all traffic circulating the UAV fleet is caught by the network listening tool. The captured traffic is sent to the filtering tool to check whether there is a match from the database containing all known attacks and signatures. If there is a match, an alarm is given. If the packet is normal, no action is taken. However, if the 14 packets are not recognized by the signature database, they are stored in the Hadoop Distributed File System. Then, by selecting the feature, deviations from the normal traffic model are performed using machine learning technique. If an attack is detected, a warning is given. The new model is saved in the database regardless of the result. Decision tree algorithm is used as machine learning method and it detects known attacks in real time. The accuracy value of the system is shown as 100%. Semi-supervised machine learning techniques are used to detect unknown attacks. If there are unknown attacks that the decision tree algorithm cannot detect, the network traffic is sent to the next step for the semi-supervised machine learning process. As these steps could not be tried in the study, test results were not given. In [34], each UAV has a recurrent neural network module that tries to detect the attack. At the same time, there is a base station with a central module that confirms the attack and notifies other nodes. When the recurrent neural network receives an input, it passes through four layers in order. Finally, output from the last layer is produced to make the proposition. The algorithm directs the parameters extracted from the traffic to the prediction function to detect the anomaly. When there is an abnormal behavior, the system informs the controller to run the decision mechanism. Datasets such as KDDCup 99, NSL-KDD, UNSW-NB15, Kyoto, CICIDS2017 and TON_IoT were used for the training and testing stages of the system. As stated in the study, there is no dataset of FANETs that can be processed in the literature. For this reason, data sets collected from networks such as MANET and VANET are used. This prevents the

results from being realistic. FANETs have a different structure from other traditional networks in terms of characteristics, and attack parameters collected from other networks can be considered invalid for FANET. As it is known, in this thesis study, artificial intelligence algorithms were applied on this dataset by creating a special dataset as a result of detailed attack analysis on FANETs. In order for the results to be realistic, a mobility model suitable for the three-dimensional movement of FANETs was chosen. In addition, two different attack detections were performed with the attacking nodes and the affected nodes, and the classification process was performed successfully. In summary, although routing attacks against AODV have been extensively studied in the literature, the different characteristics of FANETs, such as having higher-speed nodes and 3-dimensional movement, require a new analysis of attacks against these high-mobility networks. The absence of such analysis also negatively impacts the development of security 15 solutions for FANETs. In addition, there is no FANET-specific dataset for artificial intelligence applications. In addition to the detailed attack analysis on FANETs, a dataset to be studied in order to detect the attacks was also created. By using this dataset, attack detection was performed with artificial neural networks instead of traditional intrusion detection systems.

2.4 ATTACKS ON 5G

2.4.1 Denial of Service Attacks Message

security in IoT networks is provided by end-to-end encryption techniques. However, in IoT, there are actions that can disrupt the operation, occupy resources and consume the energy of objects. The most important of these attacks are DoS/DDoS attacks (Raza et al., 2013). DoS attacks are the malicious use of existing resources in a communication system, making components unable to communicate. If the attack is made from more than one point, it is called DDoS attacks. In these attacks, communication is prevented by exploiting the vulnerabilities in the TCP/IP protocol structure and consuming resources. There are some applications that perform DoS and DDoS attacks. LOIC attack was used in this study. The LOIC attack tool is an open source denial of service attack application written in C#. DoS and DDoS attacks are carried out with LOIC. Communication is stopped by sending TCP, UDP or HTTP packets to the target server. After entering the URL or IP address, “Lock on” is selected. Port, method and thread settings can be made in options. “Charging my laser” is

selected to start the attack. In this study, a port and protocolbased filter was applied with Wireshark to view the attack. Data is saved in CSV format after filtering [35].

2.4.2 Slow Denial of Service Attacks

The DoS Slow Internet of Things (SlowITe) attack is a new denial of service threat targeting the MQTT protocol. This type of threat belongs to the category of Slow DoS attacks that use minimal bandwidth and attack resources to target network services that do denial of service. Since slow DoS attacks can only target TCP-based protocols, MQTT services are slow targets because they operate over TCP (Vaccari et al., 2020). More specifically, the purpose of SlowITe is to initiate multiple connections to the server to simultaneously support all available connections that the MQTT proxy can handle. The DoS condition is reached when the attacker has established all available connections (application layer). At this point, the attacker's goal is to maintain the DoS state with the help of the MQTT proxy as long as possible with the lowest possible bandwidth. For example, other SDAs like Slowloris work over HTTP by sending a GET request to run to the server. However, as mentioned above, a CONNECT package is required to authenticate the client to the MQTT proxy. Then the attack must send these packets to the server. SlowITe initiates communication with mediators through the MQTT-based CONNECT package (Vaccari et al., 2020). The impact on the server in terms of bandwidth, CPU or memory is minimal because it is associated with intercepting all connections from the server using the lazy method. This is a common feature of slow DoS attacks. This is a hard-to-detect attack that should not be ignored. Kali Linux and Windows operating systems use slow DoS attack tools. In this study, we preferred the Kali Linux operating system and used RUDY and Slowloris attack tools [36].

2.4.3 Rudy

RUDY is categorized as a low and slow attack. It focuses on making a few extended requests rather than keeping a server busy with high volumes of fast requests. A successful RUDY attack renders the victim's web server unavailable for legitimate traffic. This tool splits the payload into packets as small as 1 byte each and sends these packets to the server at random intervals of about 10 seconds each. The tool continues to send data indefinitely. Since the behavior of the attack is similar to that of a user with a slow connection speed submitting

form data, the web server keeps the connection open to accept packets (Chaddha, 2018). It is mandatory to write “-t” or “--target” here and it is written as IP address/Web address. The optional “-d” or “-- delay” value is 5 by default, and the “-n” or “—number Of Connections” value is 500. If the port value is not entered, the value 1883 is given with "--torPort" because it will target port 9050 by default. “-o” or “--torpor” is used to enter the port value. Detailed information about the installation of the attack, usage examples and options are available on Github (Chadda, 2018). The attacks were instantly monitored and recorded with the Wireshark tool. Data saved in CSV format

2.4.4 Message Queuing Telemetry Transport Slowite

MQTT_SLOWITE is a free and open-source tool available on GitHub. Slow DoS Attack is carried out using this tool. This program, developed at the University of Pisa, is written in Python. Eclipse Paho is written to broadcast message, subscribe and perform Slow DoS attack on MQTT broker using MQTT Python client library (Carli et al., 2020). Since the source codes are written in Python, Python software packages have been installed.

2.4.5 Brute Force Attacks

Brute Force Attacks are one of the most preferred cyber-attacks to crack the passwords and passwords stored in the target system. Registered passwords and passwords include user information, credit card information, social media account information, etc. can happen. In Brute Force Attack, it results in the attacker gaining full access to the files requested by accessing the administrator authority on the target system. The attacker can also control tools such as camera, mouse or keyboard with full access. If the administrator account is protected with a password, this password is cracked with a Brute Force Attack (Kara, 2019). As in this study, a user name and password list containing letters and special characters is prepared. With this list, attack operations are carried out to reach the password.

2.4.6 Hydra

In the Kali Linux operating system, the installed Hydra tool is used to perform Brute Force Attacks. This tool supports many protocols. It can also be downloaded and run on Github. The commonly used “rockyou” password list was used in this study. In this study, the

attacker performs tests to login to Mosquitto broker with the created password list. The test process continues until the password is found. If the correct password match is found, the attack stops automatically. The attacker and broker machine runs on Kali Linux operating system. The attack is launched as FTP or IMAP. FTP, as File Transfer Protocol, is a protocol used for file transfer. It provides file transfer between two computers connected to the Internet. IMAP, on the other hand, provides e-mail access from anywhere and from any device. In this study, FTP attacks were made because there was no action related to e-mail.

2.4.7 Message Queuing Telemetry Transport

MQTTSA is open-source software written in Python. Providing an open-source client implementation of the MQTT messaging protocol, this tool is written with the pyshark4 libraries, which provide an interface to the command line tool of Paho3 and Wireshark5. There is also the functionality to detect potential vulnerabilities in MQTT brokers by automatically launching a series of attack patterns to reveal known vulnerabilities. With the MQTTSA tool, data separation and infiltration, credential retrieval, broker fingerprinting, Brute Force authentication, data tampering, DoS attacks and reports can be created. In this study, MQTTSA tool was downloaded from Github and run-on Kali Linux operating system.

2.4.8 Malicious Message Queuing Data

A malicious IoT device periodically sends large amounts of malicious MQTT data in terms of ports, networks, or other resources to take over all of the server's resources. This attack tries to saturate resources using a single connection instead of creating multiple connections. This attack was created in this case using a module inside the IoT-Flock tool. Accessed via IoT-Flock Github, the project is integrated into Creator 5 in the Linux operating system.

2.4.9 Message Queuing Telemetry Transport

2.4.10 Set Dataset

In this section, information about the MQTT set dataset, operations performed on the data, and the flow diagram of the study are explained in detail. The MQTT set dataset was created using IoT-Flock, a network traffic generation tool that can emulate networks based on

MQTT and CoAP protocols. They saved the data as PCAP file in the MQTT network traffic they created. The dataset is represented by 11,915,716 network packets and a total size of 1,093,676,216 bytes. In this study, Botnet attack 59 was added in addition to MQTT set data in order to enrich the MQTT set dataset, which includes 6 different classes, 3 attacks, 1 normal, 1 corrupt and 1 legitimate. The obtained data was added to the training and test data as a new class. In addition, data collected from the network for DoS, Brute Force, Slow DoS, corrupted data, malicious data and legitimate data has been added to the MQTT set data [40].

3. MATERIAL AND METHODS

3.1 THE CONCEPT OF ARTIFICIAL INTELLIGENCE

Although the concept of artificial intelligence is thought to be a new concept today, the history of artificial intelligence is almost aligned with the history of the computer. However, before talking about the history of artificial intelligence, it is useful to explain artificial intelligence and the concepts that make it up. According to one definition, artificial intelligence is explained as the science of machinery and engineering that does all its work using its intelligence. According to another definition, artificial intelligence is explained as transferring the abilities of human intelligence such as reasoning, producing solutions to problems and learning to machines. At this point, it becomes important not to confuse concepts such as intelligence and reason. In this sense, the concept of intelligence represents being able to relate between two situations, thinking about these situations, judging situations and reaching a conclusion. The concept of mind, on the other hand, is a phenomenon that develops and emerges by being influenced by the environment and the conditions of the society throughout a person's life. The biggest difference between mind and intelligence is that intelligence is unchangeable throughout human life, whereas mind always shows a variable feature. Based on these definitions, it becomes possible to reach the following conclusion. Systems using artificial intelligence can make people's work much easier in daily life, but they are smart in only one job. However, people can do many things thanks to their minds. For example, an artificial intelligence-supported car specializes only in driving a car, but thanks to the human mind, it can both drive and deal with other jobs. At this point, after giving introductory information about artificial intelligence, its historical process is important in terms of easier understanding of the situation. It is expressed as the first artificial intelligence studies that some people try to give soul to the sculptures made in those times. As can be seen in the table above, the artificial intelligence-supported systems we use today are actually almost the same as the first systems, it is seen that the difference between today's artificial intelligence systems and old artificial intelligence systems is quite small. This is due to the fact that artificial intelligence technology, which emerged in the 1900s, attracted a lot of attention by curious circles and then its development stopped due to the fact that artificial intelligence could not give what was expected due to technological

inadequacies. The periods when the development of artificial intelligence technology stops and is not in demand is called “artificial intelligence winter”. Although the interest in artificial intelligence has decreased from time to time, the importance of this technology has been understood again as the 2000s approached. At this point, one of the reasons why artificial intelligence is needed is due to the development of information and communication technologies and the fact that the control and processing of the resulting data has become too intense for humans to do. Another Need is artificial intelligence, although its capacities were limited at the first stage, since it is a helping system to people in most jobs, artificial intelligence was needed again. Especially in 1997, when the computer named "Deep Blue" defeated the world chess champion Garry Kasparov in the game of chess, people from many circles became interested in this field again. Moreover, the development continued without any artificial intelligence winter in the twenty-five years that have passed since then. So much so that today, artificial intelligence technology is used in cars, homes, mobile phones and even vacuum cleaners, although it is limited. At this point, it would be more accurate to examine artificial intelligence in two stages according to its functions and periods. These periods are divided into two as “limited artificial intelligence” and “artificial general intelligence”. Limited artificial intelligence, also known as "weak artificial intelligence", is a system where a computer does a predetermined job within the framework of determined rules without taking initiative. For example, assistants on our phones or systems in autonomous cars are examples of limited artificial intelligence (Reese, 2020: 77). In this context, the working principles of artificial intelligence are as follows. The first of these principles is “classical artificial intelligence”. Classical artificial intelligence is a situation that has existed since the earliest times. Classical artificial intelligence analyzes the situations in detail and makes recommendations by reaching conclusions. The second principle is “expert systems”. Expert systems analyze the data they collect from the samples and make suggestions by revealing a new variable. The last principle is “machine learning”. Although machine learning is a widely used system today, it collects the data of people and makes in-depth analysis. As a result of this analysis, the machine knows many things and is in a position to help people present rational results (Reese, 2020: 78). Artificial general intelligence, which is the second stage of artificial intelligence, is as smart as humans but does not need any programming. He analyzes the task requested from him on his own and acts on his own for the result. The fact that artificial general intelligence acts on its own

causes it to have an extremely complex structure. Therefore, there is no system with artificial general intelligence in the world yet (Reese, 2020: 195). At this point, the ability of humanity to reach artificial general intelligence is a separate issue that science needs to work on. So much so that artificial general intelligence will be the most intelligent being in the known universe, including humanity, and the road to it is quite troublesome. In this regard, Nick Bostrom has collected the ways to artificial intelligence under five different titles in his book "Super Intelligence". According to Bostrom (2021), these ways are explained as follows:

a. Artificial Intelligence: According to the first title, the way to reach artificial general intelligence is that artificial intelligence follows a path similar to human evolution. At this point, Alan Turing's concept of "child machine" becomes important. According to this concept, it is more logical and effortless to make a child machine instead of making a smart machine from scratch. The child machine, which is built at the beginning, will learn various information in the following stages, and the development of machine intelligence will be ensured by transferring what they have learned to the next machine child generations. However, since this situation is completely related to technological developments, it may cause us to reach artificial general intelligence slower than thought, but another point that should not be forgotten is that human evolution and machine evolution do not have to go the same way. It took billions of years for human beings to evolve as a whole, but since the machine is deprived of many things that humans have, it is thought that it may evolve faster, but it should be kept in mind that machine evolution may still take generations.
b. Whole Brain Emulation: It is mentioned that it is possible to achieve artificial general intelligence by imitating the human brain in this way. This way passes through following some stages. In the first step, a detailed copy of the human brain is made. At this point, the brains of deceased people need to be stabilized in order not to be damaged after death. Afterwards, the human brain is divided into various parts and these parts are scanned by a machine. 36 In the second step, the scanned digital neural network data is stored by a computer. In the third stage, the stored data is tried to be converted into a virtual brain by a computer capable of performing the operation. If the result of this process is successful, the human brain continues to live in the digital environment and it becomes easy to make and develop any add-ons to the brain that turns into software. At this point, the scanning process, the communication network established for the scanned human brain and the

computer to detect that brain in the first stage, and simulating the brain in the computer environment depend on the development of various technologies. Although there is no brain emulation developed until now, the development of technology makes this possible. At this point, it should be noted that the emulations made are likely to be successful or unsuccessful depending on the situation [42].

c. Biological Cognition: Another path to artificial general intelligence is through the development of biological brains through genetic engineering. This path consists of an attempt to create intelligent individuals by editing the genetics of embryos in the mother's womb before the baby is born. In this way, although human beings can get rid of various disadvantages, it is thought that it will not work much in developed societies, but it is thought that it can have a positive effect in societies experiencing various deprivations. However, individuals born through genetic engineering are likely to be smarter, healthier and live longer than normal people. However, this route has some disadvantages. The first of these is the time factor. Considering that every genetically modified embryo is useless, this means that the level of intelligence that will emerge will be low. This means that reaching the target may take more time than usual. Again, there may be a possibility that some closed societies may reject this situation by citing religious or moral reasons. Families, on the other hand, may oppose this path for reasons such as having children naturally or not accepting children for adoption. However, at this point, it is thought that if any of the families accept to have a child in this way and the result is successful, other families will also follow this path.
d. Brain-Computer Interfaces: Another way to artificial general intelligence is through brain-computer interfaces. This work is carried out with some devices mounted directly on the human brain. These devices are devices that are tasked with increasing the capacity of the human brain. There are currently such devices used to eliminate health problems such as Parkinson's or paralysis, but this method may cause some problems that healthy people who do not have health problems cannot afford. Problems that may develop in this method are as follows; It is thought that the device to be placed in the brain may cause some negative consequences such as various infections, infection-related inflammations, cerebral hemorrhage and mental retardation.
e. Networks and Organizations: This method is perhaps one of the most familiar situations for human beings. With this method, it is thought that it will be easier to reach artificial

general intelligence by connecting human brains to each other through a network and creating a common intelligence. Bostrom said that reaching artificial general intelligence can only be achieved by following the above ways, and he also stated that not all of these ways have to be successful. At the same time, he stated that it is correct to divide the artificial general intelligence that humanity has reached into three. These are: "fast superintelligence" as a system that does everything that human intelligence can do faster than it, "collective superintelligence" formed by the combination of more than one discrete intelligence, and the last one is "qualified superintelligence" that shows human qualities in terms of intelligence but is smarter than a human. states that he is “super-intelligent”. As a result, although artificial intelligence studies started in the old times, it has come to the present day by experiencing pauses from time to time. However, today, artificial intelligence studies have gained speed and investments in this field have increased considerably. Thanks to this interest in artificial intelligence, it is a fact that sooner or later artificial general intelligence will be achieved. However, the path to artificial general intelligence seems to be quite complex and difficult, and it is an undeniable fact that humanity must calculate the situations that will happen to it when it reaches artificial general intelligence [43].

f. Machine learning is a sub-field of computer science that first developed in 1959 with study of artificial intelligence for computer learning and pattern recognition. Machine learning is a discipline that explores the study and generation of algorithms that are capable of learning with structured functions and making predictions from data. These algorithms work by building models that predict and make decisions on data based on input patterns rather than precisely following firmware instructions. Specifically, machine learning automatically detects patterns in data and uses hidden patterns (such as additional data collection plans) to develop a set of tools to predict or study future data as technical.

3.2 ARTIFICIAL LEARNING AND BIG DATA

Artificial Intelligence is the sum of all tasks for the creation of intelligent agents/programs that are thought to exhibit higher cognitive functions or autonomous behaviors such as perception, learning, concepts, reasoning, reasoning, problem solving, communication,

reasoning, decision making. It is embedded in human intelligence. specialization. Artificial intelligence is inspired by the brain. Cognitive scientists and neuroscientists who want to understand how the brain works, model artificial neural networks and conduct simulation studies. Artificial intelligence is a part of computer science, and like other branches of technology, its purpose is to create useful systems. Computers were once called "electronic brains," but computers and the brain work differently. Computers often have more than one processor, but the brain is made up of many computational units (neurons) working in parallel. The exact details are not known, but the processor is said to be much simpler and slower than a typical computer processor. A neuron in the brain is connected to tens of thousands of other neurons, called synapses, all working in parallel. In a computer, the processor is active; memory is separate and passive. However, in the brain, both processing and memory are believed to be distributed together on the network; Processing is done by neurons, and memory occurs at synapses between neurons. Artificial intelligence needs a high amount of data in order to perform a healthy learning. In this sense, artificial intelligence needs big data for the realization of artificial learning. Considering the amount of data that humanity is exposed to and produced today, processing this data is well above the capacity of human beings. Artificial learning comes to the fore in order to process or analyze the accumulated and growing data pile [44]. However, before talking about machine learning, it is more important to talk about the concept of big data, which is the data pool that feeds machine learning algorithms. The word data is of Latin origin and used in forms such as “dare, datum”. The word data is used as "data" in English (etimolojiturkce.com). The word used as "data" in Turkish is defined as the word "information" in its simple form (TDK, ty). In this direction, data is also defined as raw evidence according to another definition [45]. Big data, on the other hand, can be explained as the whole of the data produced to facilitate the analysis of people and the environment and the prediction of their subsequent behavior. The concept was first used by NASA in 1997. A problem arose when all the storage units available at that time were full, and this problem was called the "big data problem". Although this situation is not seen as a problem today, it still continues. So much so that everything that can turn into data, including all the behaviors done in digital environments and our personal preferences, is in the field of big data. In this sense, it is necessary to mention five basic features that define big data. The first feature is “data size”. Data size briefly describes the amount of data produced by the data generating device. The

second feature of big data is “speed”. Speed refers to the rate at which the data generated by the device is processed. The third feature of big data is “diversity”. Diversity is the feature that states that the 39 data produced can be different types (sound, text, image) rather than just one type. The fourth feature of big data is “validation”. The validation feature is the feature that means that the data used is tested whether it is safe or not. The last feature of big data is the "value" feature. The value feature is the feature that helps to bring the previous transactions to a conclusion. In addition to all these, although big data contains many negative situations, it offers various opportunities to people and therefore it is desired to be perceived as a positive situation. Moreover, companies such as Facebook and Google store all kinds of data of their users and make commercial profits from this data. They operate more like data merchants than social media companies. In this sense, big data is used by capital owners as a tool to monitor and control users. At this point, machine learning significantly benefits from big data, but large amounts of data complicate the work of algorithms and experts. For this reason, the data mining method, which is the process of selecting and using only important and meaningful data among the dense data load, is applied. With data mining, it is aimed to eliminate not only the meaningful data but also the errors in the system. Besides, data mining works on two types of models, these models are; The "predictive" model based on the estimation of the results and the "descriptive" model that reaches the result by establishing a relationship between the data. It is known that artificial learning can be performed as a result of collecting data by data mining method. In this sense, although artificial learning is related to the creation of information from data by extracting information, it has a common denominator with statistics, computer science and artificial intelligence. In addition, artificial learning is called statistical learning and predictive analytical learning (Müller and Guido, 2016: 1). Machine learning is basically one of the sub-fields of artificial intelligence, but artificial learning has been so successful today that its name is mentioned more than artificial intelligence. The reason for this is that the possibilities offered by artificial intelligence to humanity since its emergence are less than the possibilities offered by artificial learning. So much so that it had to be programmed by an expert in order to make artificial intelligence do good works, but artificial learning is a successful algorithm by making inferences from its experiences, although it is based on data. In other words, it can be said that it is more successful according to the frequency of use and the experiences that the algorithm makes from the mistakes made, and thus learning

algorithms work more accurately. At this point, it seems that it is an undeniable fact that artificial learning will bring a new breath to artificial intelligence, which often pauses, and will carry artificial intelligence into the future. In this sense, artificial intelligence is experiencing its most intelligent period with artificial learning because the learning function requires intelligence. An entity can only adapt to its environment with intelligence and learn from its mistakes with intelligence. In addition to all these, artificial learning eliminates some of the disadvantages that occur during the algorithm development phase. These are:

a. Algorithms written in the past were experts in a single job, so inputs were created according to a single job. In the slightest change in the inputs, the software had to change radically, but this situation disappears in machine learning.
b. For algorithms written in the past, it was necessary to be an expert in that field. This situation caused problems in works such as face recognition technology because algorithms needed deep thoughts of people at that time, but artificial learning eliminated this problem. Machine learning is widely used today because it works successfully. Especially in the Internet environment, machine learning algorithms are widely used by capital owners to measure the behavior of users, to offer suggestions to users and to manipulate users about a subject they want. For example, Google personalizes the ads it will show to users through learning algorithms. Again, online shopping sites recommend products to users using learning algorithms, and this directly affects the purchasing behavior of users. Learning algorithms are also used in assistants on smart phones, assistants help users more by learning from people's data. It performs its functions with the help of learning algorithms in vehicles with autopilot, which are being used today. At this point, some learning models are used to realize the learning process. These; There are three different forms of supervised learning, unsupervised learning and reinforcement learning.

The chronological progress of the concepts of artificial intelligence, machine learning and deep learning can be seen in the graph given in figure 3.1.

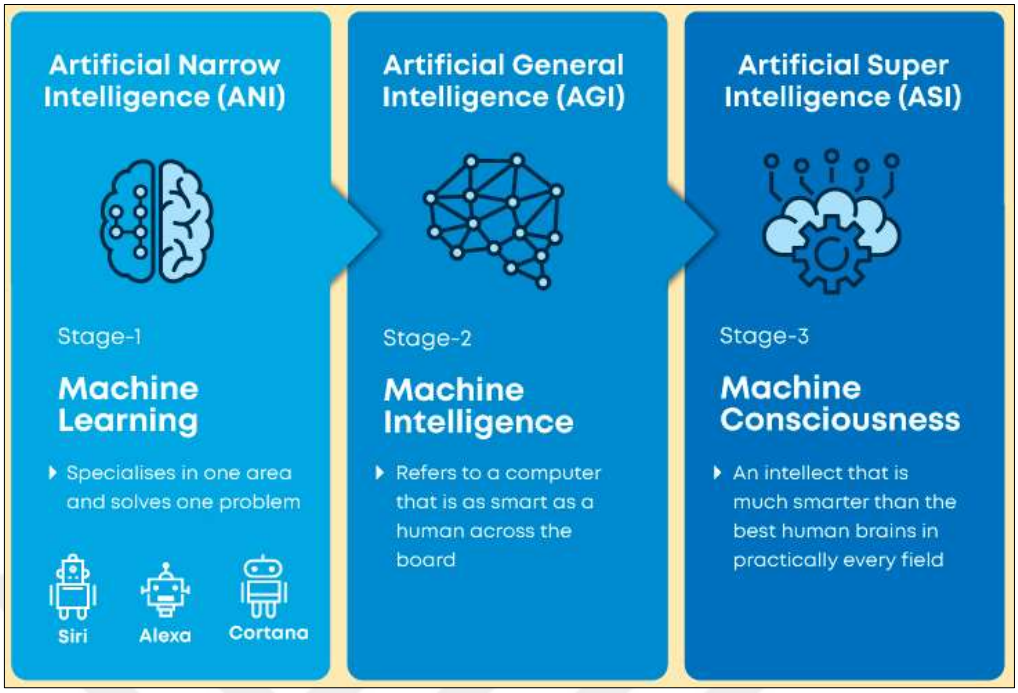

Figure 3.1: Artificial Intelligence.

3.2.1 Supervised Learning

Follows a process in which the inputs and outputs to be made in the data sets created in the supervised learning process are known. At this point, the inputs and outputs in the data set should be created correctly and homogeneously. Otherwise, the result may not be successful. The reason for giving inputs and outputs in supervised learning is to find errors during learning and to reduce these errors to zero according to the given result to create an error-free learning process [48]. As seen in Figure 2.2, the supervised learning process includes two stages, namely "classification" and "regression". Classification is an important issue in the supervised learning process. By choosing the appropriate technique for the model to be made, the classification of the elements in the data set with an unknown class is made. Therefore, the elements in the data set stand apart from each other, not nested. In this way, it becomes easier to do things such as making a diagnosis or the spam status of incoming e-mails. On the contrary, if the elements are not discrete but continuous and nested, it is in the domain of regression. The purpose of regression analysis is to measure the effect of inputs on outputs. With the regression analysis, conditions such as weather forecast or life expectancy are measured [49]. Artificial intelligence or categorizing data according to

algorithms. Classification in machine learning is the process of classifying objects based on a pre-classified set of training data. As in Fig. 2.10, attempts are made to tightly group the data in certain regions using various mathematical algorithms. Classification is a supervised learning algorithm. Classification algorithms use the classification of training data to calculate the probability that a new element will fall into a given category.

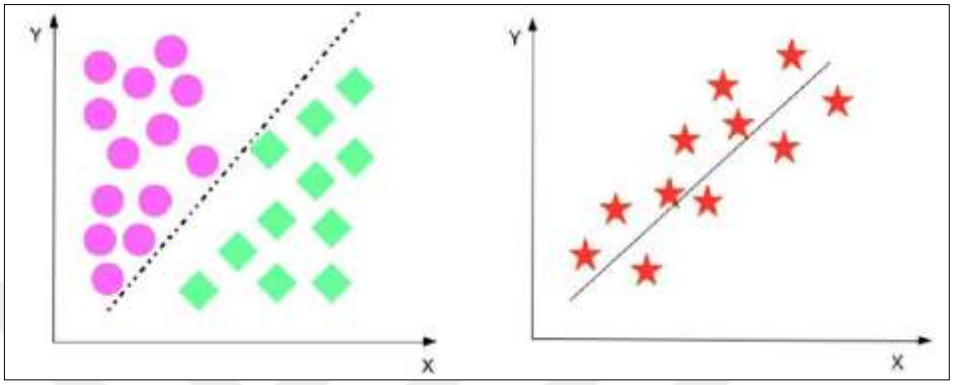

Figure 3.2: Supervised Learning.

In supervised learning, can predict a set of data and a desired outcome based on that data. Supervised learning allows the machine to extract features (relationships between inputs and outputs) from this information and return the data and outputs from that data to the machine. In this way, the device gets to know the relationship between the data. The expected learning tasks are divided into two categories: 'regression' and 'classification'. Regression problems attempt to predict the outcome of a set of assumptions. In classification problems, letters of placement are assigned to individual categories. Try to evaluate the results of each task. Targeted Learning According to the logic that loads the training and test sets into the system, it generates the necessary labels for each data in the data by creating a relationship between the input data set, test data, and output data. The main goal is to make effective predictions for data sets with unknown outcomes, based on software designed for data sets with known outcomes. In supervised learning tasks, you start with a data set containing training examples with an exact match label. For example, when learning the order of handwritten numbers, the learning algorithm uses thousands of handwritten pictures, each picture representing the number 21. Then the algorithm learns the relationship between the picture and the corresponding number and uses the learned relationship to select a new (unwritten) picture that it has not seen device before. In supervised learning, the machine attempts to learn from

scratch the relationship between money and education by running the recorded training data through a learning algorithm. This learned function can be used to estimate an unknown amount such as Y. That is, we can estimate Y by applying the generated model to unpublished test data.

Supervised learning algorithms are used to build models from data sets with desired inputs and outputs. The data used is training data and consists of training examples with one or more inputs and outputs. The methods used are known as control screens. When dealing with a symbol problem, we start with a set of instructional examples that include the appropriate characters. An example is training a supervised learning algorithm that takes thousands of pictures of numbers handwritten in real letters with a given number for each picture it represents. This allows the algorithm to sort handwritten numbers, learn the relationship between the pictures and those numbers, and use that relationship to organize new, unwritten pictures that were not previously displayed on the device. Planning was discussed and the learning component identified. Classification algorithms are used when the output is limited to a finite set of values. Regression algorithms are used when the output can be numeric. The most popular supervised learning algorithms are DVM, NB, RF, DT, KNN algorithms and optimization algorithms [50].

3.2.2 Unsupervised Learning

In unsupervised learning, there is no supervisor as in supervised learning. Therefore, it is expected that the system itself will reach the result by giving only input data to the data set. At this point, since the result from the system is important, what the result means is looked at and classified accordingly. Another purpose of unsupervised learning is to reveal the relationship between the elements in the data set whose status is not clear. Unsupervised learning consists of two stages as “clustering” and “dimension reduction” as seen in Figure 2.3. As the name suggests, clustering is done to collect similar items in the data set in the same area. Supervised learning algorithms are used to build models from data sets with desired inputs and outputs. The data used is training data and consists of training examples with one or more inputs and outputs. The methods used are known as control screens. When dealing with a symbol problem, we start with a set of instructional examples that include the appropriate characters. An example is training a supervised learning algorithm that takes

thousands of pictures of numbers handwritten in real letters with a given number for each picture it represents. This allows the algorithm to sort handwritten numbers, learn the relationship between the pictures and those numbers, and use that relationship to organize new, unwritten pictures that were not previously displayed on the device. Planning was discussed and the learning component identified. Classification algorithms are used when the output is limited to a finite set of values. Regression algorithms are used when the output can be numeric.

This method is used in data mining. Some methods are used for clustering. Unsupervised learning is the approach used when we have little or no idea of what the desired output from our data looks like. Here the data collected in the dataset has no labels. A model can be created from data whose effects of variables are unknown. We can create various models and structures by clustering the data based on the relationships between the variables. When the algorithm learns, it is not intervened. These resulting models are useful when there is a need to categorize items or find a relationship between them. They can also help detect anomalies and flaws in data. Since the dataset is not labeled, there is no way to get the way the data is sorted. They may be less accurate as the input data is unknown and labeled by the people who built the machine. The information obtained by the algorithm may not always correspond to the desired output class. These:

a. Distance-based techniques
b. Techniques based on finding the center point of clusters
c. Statistical techniques based on distribution of elements. Another stage of unsupervised learning is dimension reduction. Size reduction is the step that simplifies the data set by removing meaningless data in systems that are dense in terms of data. If size reduction is not done, the process becomes more difficult and performance drops are experienced in the system. During this process, only useful information for the system is preserved. Some techniques used for size reduction are:
d. Principal component analysis
e. Kernel PCA.

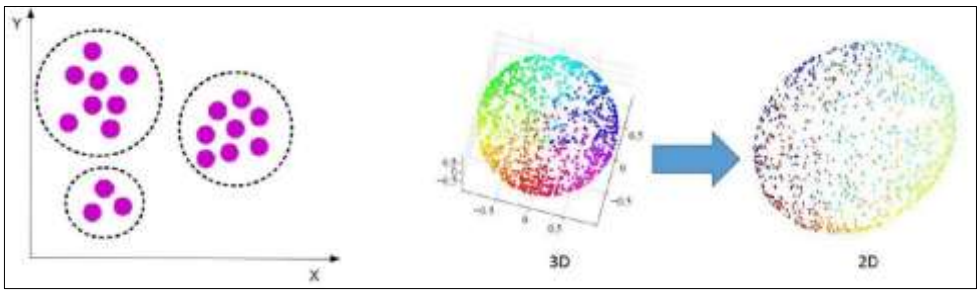

Figure 3.3: Clustering.

Unsupervised learning algorithms are given only datasets that contain input, and then the algorithm finds the clustering of data points. Therefore, algorithms learn from unlabeled, uncategorized, or unclassified test data. Unsupervised learning algorithms detect commonalities in data based on the absence or presence of commonalities in each new piece of data rather than a feedback approach. Grouping, summarizing and finding the basic structure of unlabeled data in a useful way is the function of unsupervised learning. Unsupervised learning is generally used for clustering because it groups the data.

3.2.3 Reinforced Learning

Reinforcement learning has some similarities with other types of learning, but in this type of learning, the system receives only one input. There are also system control mechanisms that speed up from time to time. In this type of training, the system of rewards and punishments that provides the training is used to the fullest. This is how the system solves problems. The relationship between the model and the environment is also important for this type of learning. Positive and successful relationships between role models and the environment are reinforced with rewards and bad behavior is punished. This method is also used in many other fields, especially in artificial intelligence in games. Reinforcement Learning there is a third, less commonly used type of machine learning called reinforcement learning. This can help them learn to behave or not to behave when sometimes signs of reward or punishment are given. I want to build a machine that can learn to play chess. Let's say the pieces on the board are cameras that watch the players and they can choose targets to determine moves to win the game. In this case, the supervised student is not used for two reasons. First, it is very expensive for users to go through many games and prepare tutorials that show the best move for each situation on the board. Second, in most cases, there are no best practices. The quality

of your moves depends on your next move. The only real feedback is when you win the game and lose at the end of the game. The decision maker, called the arbiter, is placed in the environment. The board is the client's game environment. At each instant, the environment is in a certain state. That is to say the position of the stones on the board. The decision maker has several possible actions. Correct movement of pieces on the board. When an action is selected and executed, the state changes. In any external situation, an agent can initiate an action, and this action can change the situation and reward or not. Solving a problem requires a series of actions, and feedback is received in the form of a reward. What makes it difficult to learn is that you get very few rewards and games are usually won or lost after a long series of moves after completing a full round. Rewards define the purpose of the activity and are necessary if training is required. Agents learn the optimal set of tasks to solve. By "best" here we mean the sequence of actions that provides the maximum reward in the shortest possible time. It is a reinforcement learning environment. Reinforcement learning differs from other learning methods in several ways. Unlike teacher-assisted learning, this is called “critical learning”. Critics, unlike teachers, tell us not what to do, but how well we have done in the past. Critics don't care. Critical comments are lacking and delayed. After performing several actions to receive a reward, it is proposed to evaluate the individual actions performed in the past, find out which actions led to the reward, and then return the reward and keep it. Initially, the machine's algorithms perform better as the amount of data increases. First, it is very expensive for users to go through many games and prepare tutorials that show the best move for each situation on the board. Second, in most cases, there are no best practices. The quality of your moves depends on your next move. The only real feedback is when you win the game and lose at the end of the game. The decision maker, called the arbiter, is placed in the environment. The board is the client's game environment. At each instant, the environment is in a certain state. That is to say the position of the stones on the board. The decision maker has several possible actions. Correct movement of pieces on the board. When an action is selected and executed, the state changes. In any external situation, an agent can initiate an action, and this action can change the situation and reward or not. Solving a problem requires a series of actions, and feedback is received in the form of a reward. What makes it difficult to learn is that you get very few rewards and games are usually won or lost after a long series of moves after completing a full round. Rewards define the purpose of the activity and are necessary if training is required. Agents learn the optimal

set of tasks to solve. However, the condition persists beyond a certain level of data size. No performance changes due to critical settings [50].

3.2.4 Artificial Neural Networks and Deep Learning

The working principle of artificial neural networks is based on biological neural networks. In other words, artificial neural networks learn by imitating neurons in living brains. Although the history of ANNs dates back to the 1940s, today it has come to the fore especially with the development of technology. However, it can be said that artificial neural networks still have a long way to go. This is due to the difficulty of studying living brains. So much so that the brain is almost in a closed box, so it becomes difficult to examine it. The brains of dead creatures are also not worth examining. Another difficulty in this regard arises from the fact that the brain has a very complex structure. So much so that the number of neurons in the brain and the connections between them is thought to be equal to the number of stars in the galaxy. Thanks to this complex structure of the brain, it is known that many operations can be performed per second. Moreover, as the neurons in the brain learn and process, developments occur in the brain. These layers are as follows; The first layer of the artificial neural network is the input layer where the inputs are made. External data is entered into this layer and the data is sent to the next layer by this layer. There is a section called "weights" between the input layer and the hidden layer. In this section, the order of importance of the data is determined. If the weight value given to the data is large, it means it is important, if the weight value is small, it is less important. After the weights of the data are determined, they are sent to the hidden layer where the data is collected and processed. In this layer, the states of the data are recalculated, except for the collection of data. The data processed in the hidden layer is sent to the output layer, but there is a section called the "activation function" between the two layers. Thanks to this section, the outputs are recalculated and arranged according to the inputs. Afterwards, the data is sent to the output layer, and thus learning is achieved. It is known that the data coming out of the output layer can be connected to another artificial neural network from there, or if the learning process has failed, a new learning process can be started again [51]. Deep learning occurs by increasing the number of hidden layers of artificial neural networks. Using deep learning artificial neural networks, complex tasks such as face recognition and medical fields are

made possible thanks to the increasing layer and developing computer technology (Learning, 2021: 49). In addition to all these, artificial neural networks are divided into two as single layer neural networks and multilayer neural networks. These layers can be explained as follows:

a. Single-Layer Neural Networks: Single-layer neural networks consist of only input and output layers. These types of neural networks are not useful for complex tasks. Threshold values are important for such layers. The threshold value is determined by the founder of the system and this is important in determining the quality of the classification process [52].
b. Multilayer Neural Networks: Multilayer neural networks, which emerged in the 1980s, contain hidden layers between the input and output layers, unlike single-layer neural networks. Before multi-layer neural networks, complex problems could not be solved, but with multi-layer neural networks, complex problems can now be solved. The purpose of using multilayer neural networks is known as minimizing the error during learning. In addition, deep learning has emerged thanks to this network system.

In general, machine learning systems learn by processing data consisting of a single layer, while deep learning systems learn by analyzing using very large data sets and multiple layers. In Deep Learning, the machine learns how to learn by processing large amounts of data in different layers. While performing normal machine learning application, the data loaded into the model is given together with the desired properties of that data. The model also produces a function by learning which feature affects the result and how much. In deep learning, features are not determined by the user; The model decides which features to look at and then finds how much weight it will give to which data in the function. In order for this learning to take place, powerful computer hardware that can process big data is needed. In machine learning, data analysis that was previously done with the CPU (Central Processing Unit) is now done tens or even hundreds of times faster and more securely with the much faster GPU (Graphics Processing Unit).

a. Input Layer: By examining the pixels of an image, it is perceived that there are dogs, sticks and nature in the picture.
b. Intermediate (Hidden) Layer: Usually when training the neural network, different weights are given to the use of tags/information, and these are known as "Learned Weights". In this layer, the relationship between these data is analyzed with the help of previous inputs.

c. Output Layer: The final layer gives you an estimate of the data you enter into your network. Estimated output in our example.

3.3 ROUTING PROTOCOLS

Routing Protocols Ad hoc networks provide a temporary network connection between devices without the need for wireless access points or any router. Since they do not need a fixed infrastructure, they can be configured in a short time and made ready for use quickly. The absence of any infrastructure needs also reduces costs. In addition, since all nodes in the network cooperate in packet transmissions, there is no need for special nodes to perform the transmission. Nodes send data directly over each other using the multi-hop method without a central access point. Ad hoc networks have a variable topology because there is no fixed access point. In order to adapt to the dynamic topology, traditional routing protocols are replaced by adaptive routing protocols that can adapt to the dynamic structure. In order for FANETs to perform their tasks, the nodes need to communicate both with each other and with the ground base station. While establishing this communication, routing protocols designed to provide real-time data transmission, reduce processor and energy costs, do not affect control messages and network traffic, and adapt to the constantly changing topology [53] are needed. Considering all FANET features, a protocol should be designed from scratch or the recommended routing protocols for MANETs should be adapted to FANETs.
Proactive routing protocols

Proactive routing protocol is a table-based protocol, and each node contains the final route information of all other nodes in the network, regardless of whether there is data to be sent. In order to update the latest route information, each node sends control packets to each other at certain intervals. With the proactive routing protocol, the selection of the shortest path between the source and destination node is easier since all nodes have the latest route information, so the packet delivery time can be reduced. However, when proactive routing protocols are used in highly dynamic networks such as FANETs, continuous packet exchange occurs and bandwidth is consumed. As this situation causes congestion in the network, connection interruptions occur. In addition, since the nodes will constantly listen to the environment, it creates a negative situation in terms of energy. Examples of proactive

routing protocols are Dynamic sequence distance vector routing protocol 29 (dynamic sequence distance vector- dsdv) and Optimized link state routing protocol.

a. Dynamic Sequence Distance Vector (DSDV) DSDV [54], a table-based routing protocol, is based on the Bellman-Ford algorithm. The sequence number parameter is used to prevent routing loops and network congestion after topology changes. Routes with higher sequence numbers are designated as current routes. In addition, a damping parameter is included in the packages, which allows to report short-term topology changes. For any short-term topology change, the DSDV routing protocol is based on the method that each node transmits its routing table to neighboring nodes at certain time intervals and the parameters are recalculated.
b. Optimized Link State Routing Protocol (OLSR) OLSR creates a global information of all existing links between nodes with the link state routing strategy method [55]. This protocol broadcasts 2 messages on the network to update the topology information. The first are Hello messages and are sent to check neighbor connection status. The other is the topology control message and is broadcast to share the change information in the network. In order to reduce the overhead on the network, when the source node wants to communicate with the destination node, it determines another node as Multipoint Relays (MLR) as an intermediary. Each node creates a list of neighbors 1 hop away to be able to retransmit messages. This list is made by means of topology control and periodic exchange of hello messages. After creating a neighbor list, each node chooses a node as the CNR. Only these CNR nodes can generate link state information and forward data packets to other CNRs. As shown in figure 3.9, S is the source and D is the destination node.

3.3.1 Reactive Routing Protocols

In the reactive routing protocol, also called the on-demand routing protocol, a route is created when packet transmission to the destination node is desired. The route search is started with the route request packets broadcast on the network. Then the destination node sends the path response packet using the shortest path. In this case, instead of saving all routes, each node saves only the current route used and all tables in the network do not need to be refreshed. The main advantage of using this technique is the bandwidth efficiency. However, since the

routes are not periodically shared with all nodes in the network, the route discovery process, etc., when sending data. network delays for reasons. Therefore, reactive routing protocols start sending data later than proactive routing protocols.

a. Dynamic source routing protocol (DSR) DSR, which is a reactive routing protocol, is designed for MANETs. It allows the network to be self-organized, maintained and configured. In order to establish communication between the nodes, first the route discovery process begins. The shortest path 31 to the destination is determined. The route response packet is sent from the destination node to the source node and the data sending process is started. In this protocol, it is insufficient for networks with a large and constantly changing topology, since each packet contains all the addresses of the transmitted nodes. The route search is started with the route request packets broadcast on the network. Then the destination node sends the path response packet using the shortest path. In this case, instead of saving all routes, each node saves only the current route used and all tables in the network do not need to be refreshed. The main advantage of using this technique is the bandwidth efficiency. However, since the routes are not periodically shared with all nodes in the network, the route discovery process, etc., when sending data. network delays for reasons. Therefore, reactive routing protocols start sending data later than proactive routing protocols.
b. Ad Hoc On Demand Distance Vector (AODV): When data transfer is requested, the source node checks the routing table for the existence of a route for the destination node before starting the route discovery. The routing table generally records the destination address, next hop address, destination sequence number, and hop count. This information is checked and the route formation is checked. If a route has been created before, the source node sends data packets over this route. If no route has been created before, the source node broadcasts route request (RREQ) packets. This process is illustrated with a flowchart in Figure 3.10. When any node with a valid route to the destination or the destination node receives an RREQ packet, it sends a route response (RREP) packet to the source node (a unicast RREP packet is sent). RREP packets contain information about the uniformly ascending sequence number and the minimum number of hops held by the source nodes. The source node chooses the newest and shortest route with the minimum number of hops and the maximum sequence number. After the route is created, data transfer is initiated between the destination node and the source node. Route error (RERR)

packets are sent to notify other nodes of disconnections between nodes, and data packets are dropped if local maintenance is not enabled. When local maintenance is enabled, the node with 32 connection failures tries to establish a new route by sending an RREQ packet back to the destination node [56].

c. Route discovery and data sending for AODV routing protocol: When RREQ message is broadcast for route discovery in AODV protocol or forwarded to neighbors, there is a high probability that a node will receive the same message more than once. It is necessary to prevent the nodes from constantly transmitting bear messages to each other and the formation of an endless loop. Therefore, when the node receives an RREQ, it looks to see if it has received any RREQs containing the same source IP address and RREQ ID (RREQ ID) during the route discovery parameter period. If such an RREQ message has been received before, the packet is discarded, if not, it is transferred to another node. If no RREP message is received in response to the sent RREQ messages, the RREQ message is sent repeatedly until the determined maximum RREQ limit. In order to prevent RREQ packets from looping endlessly, the maximum hop amount or lifetime values in the packet are controlled. A packet with a value higher than the specified limit values is ignored by the node. Unicast RREP message responses are sent to RREQ messages sent from the source node to the destination. There are two options here. RREP can be sent by the destination node or by the intermediate node indicating that it has an active route to the destination node. The route search is started with the route request packets broadcast on the network. Then the destination node sends the path response packet using the shortest path. In this case, instead of saving all routes, each node saves only the current route used and all tables in the network do not need to be refreshed. The main advantage of using this technique is the bandwidth efficiency. However, since the routes are not periodically shared with all nodes in the network, the route discovery process, etc., when sending data. network delays for reasons. Therefore, reactive routing protocols start sending data later than proactive routing protocols. After a node responds with RREP, it discards the RREQ message it receives. 33 When the RREP message sent by the destination node or the node reporting the active route to the destination reaches the source node, data transmission is started by selecting the newest and shortest route with the minimum number of hops and the maximum sequence number in the packet.

d. Repair process for AODV routing protocol: local link in AODV, hello messages, feedback messages, link layer level checks, etc. can be controlled by various methods such as in this thesis, periodic hello messages are used for local connection control. Each node in the network is constantly checking the next active hop and its connections to its neighbors transmitting hello messages. If a node does not receive periodic hello messages from a previously neighbor node, it is assumed that the connection to that neighbor is broken. The node with the disconnection reports the disconnect by sending a RERR message to the source node. If there are many 34 nodes using the connection, the RERR message will be broadcast, otherwise unicast. When any node receives a RERR message, it checks whether the sending node is on the active route to the destination. If it is on the active route, it marks these routes from its route table as invalid and sends the RERR message towards the source. This process continues until the RERR is received by the source. If this route is still needed when the source node receives the message, route discovery is restarted. AODV enables nodes to quickly establish routes to new destinations. It can respond to changes in network topology and connection errors without delay. Therefore, AODV is more suitable for use in FANETs compared to other routing protocols. Studies such as [72-73] have been done to adapt AODV to FANETs [57].

3.3.2 Security Attacks

a. Attacks Targeting the Network Layer FANETs are subject to different types of attacks. Attacks targeting the network layer are generally aimed at controlling network traffic, disrupting its functionality, and incorporating attacking nodes into the network. Attackers on the network can access packets, redirect them at will, and even disrupt the workings of routing protocols. These types of attacks are divided into passive and active attacks. Passive attacks are only aimed at eavesdropping on traffic on the network. Active attacks, on the other hand, are aimed at duplicating, changing, or downgrading data packets sent between nodes. In this study, various attacks were analyzed using the AODV routing protocol. The AODV protocol is a recommended protocol for FANETs in the literature. The most important reason for this is that the AODV protocol adapts to the constantly changing topology due to its high speed. At the same time, since it is a reactive protocol, it creates routes only when data transfer is desired, and this reduces the overhead. Because

FANETs have a highly mobile structure, disconnections are also common. The AODV protocol can respond quickly to connection errors. Therefore, it is suitable for use in FANETs. Each attack is described in detail below. In this attack scenario, as shown in Figure 4.1, the attacker node aims to attract network traffic by advertising to the source node that it has a better route to the destination [58]. Depending on the method to be applied, the attacker can either change the data packets that the node receives on it or perform an attack limited to just pulling on it. A sinkhole attack can cause disconnections as it increases the network overhead. It also increases energy consumption and shortens the lifetime of the network. In the attack scenario implemented in this study, when the attacker node receives an RREQ packet, it responds with a bogus RREP packet informing that it is one hop away from the target node. With this behavior, the attacker node reports that it has the shortest route to the target node and increases the probability of being selected. In addition, the attacking node increments the target sequence number and advertises itself as the most recent route to the target, thus ensuring that it is selected as the route to the target. When this route is selected, the attacker listens for all communication between the source and destination nodes, so it is called a sinkhole attack.

b. Dropping Attack In general, the attacking node must be activated during the route formation to initiate the packet drop attack. In this simple attack scenario, the attacker aims to drop the packets it receives. It can selectively downgrade packets sent to a specific destination. Or it may randomly drop some packets to reduce the detectability of the attack, but in this case the attack's effect is expected to be more limited. The attacker can drop routing control packets as well as data packets. In this case, active routes may not be created or notification of inactive routes may not occur in a timely manner. In such cases, network resources are consumed as the routing mechanism will be restarted, and situations may arise that may cause network congestion and delays. In this operation, if the attacker node is on the route between the source node and the destination node, it drops all the data packets it receives and causes the communication between the source node and the destination node to be interrupted. As shown in Figure 4.2, the attacker in the route between the source node and the destination node drops the incoming data packets on the node.
c. Black Hole Attack A black hole attack is a combined attack that performs cascading sinkhole and packet drop attacks. By promoting that it has the best route to the destination,

it first redirects the network traffic to itself, then performs other attacks such as modification and packet drop attacks to the network traffic it receives. In the simulations here, the sinkhole attack is performed as described above in the first stage of the attack, then only the packet drop attack is performed in the second stage of this attack. In other words, the attacker node first pulls the data packets and then drops these packets.

d. Ad hoc Network Flood Attack In this attack scenario, the attacker exploits control packets sent during the node route discovery process. There are many types of this attack depending on the application area. Multiple RREP or RREQ packets can be optionally sent during the attack. These packets can be sent to nodes in the network or to node addresses that are not present in the network. In this study, RREQ Flood Attack scenario was applied. Attacking nodes send multiple RREQ packets to selected nodes. This attack causes increased network traffic, consumption of network and node resources, disconnection between nodes, and interruption of data transmission. In simulations, a random target node is selected and 10 new RREQ messages are sent to discover the routes to that target node. The attack is repeated every 12 seconds for the duration of the simulation for another randomly selected target node.

3.4 PROPOSED METHOD

Data preprocessing is the first step in the data collection process before it can be used by any ML model. It is generally used to transform the raw data into a structure that the ML model can process and also helps to increase the quality of the model. Classification and outcome evaluation are the most important steps in the proposed model. It is highly effective on the performance of the dataset. Datasets can contain symbolic properties. But classifiers cannot handle these symbols. Pre-processing is therefore needed. All non-numeric or symbolic properties are removed or replaced from the dataset. Many methods can be used in data preprocessing. Symbols were removed on the data obtained from the smart home network, and non-numeric (nominal and string) values were converted to numeric values. These operations are carried out with the help of the functions in the filters→ unsupervised→ attribute in the Anaconda tool. In addition, normalization, standardization and feature selection methods were used in this study. Normalization is a preprocessing method often applied as part of data preparation for ML models. The purpose of normalization is to set the

values of numerical data in a dataset to a range (usually 0–1) using a common scale without distorting the differences in the true value ranges and losing information. In the Anaconda tool, values are set to the range of 0-1 with the “Normalize” function under filters → unsupervised → attribute.

Then, the data wired to the two Boltzmann Machines that are used to reduce the size of input data. The extracted features by the Boltzmann machines wired to the AdaBoost that used to classify the extracted features to the classes normal and abnormal. The flowchart in the

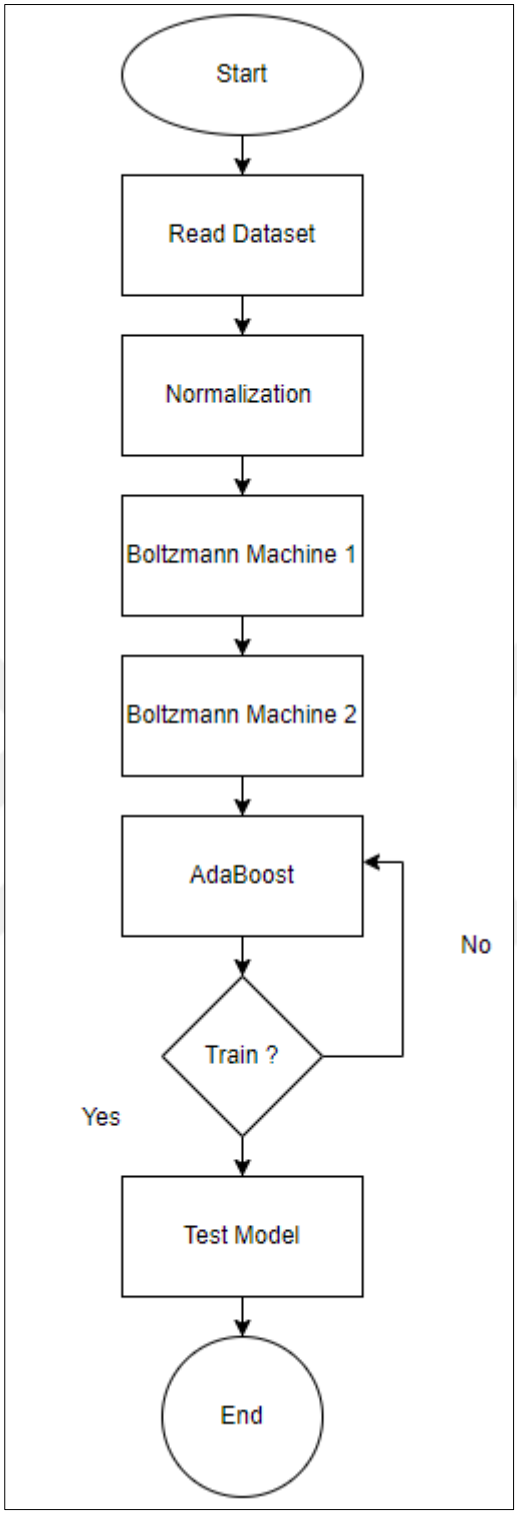

Figure 3.4: Proposed Method.

4. RESULTS

Machine learning techniques in Intrusion Detection Systems show high success for anomaly detection. In order to better detect new attacks, machine learning techniques are used in combination with data mining. In general, the success of an Intrusion Detection System depends on its ability to detect an attack effectively. Intrusion Detection Systems generate alarms that can warn system administrators in case of any attack. They do not take any preventive action. Intrusion prevention systems, on the other hand, have the ability to detect, isolate and block an attack. One of the stages that can affect the performance of the system in machine learning techniques is classification processes. According to the obtained features, the samples belonging to each class are located in a certain region. The separation process between classes is interpreted according to these positions of the samples belonging to each class. According to the findings, the appropriate classifier or multiple classifiers are selected. There are many ML techniques that can be used for data classification.

4.1.1 Decision Tree Results

A decision tree is a tree-based method in which a sequence of data splits is represented, starting at the root until it reaches a Boolean result at the farthest node. In the decision tree, which consists of trees, branches and nodes; The features to be classified are represented by a node, and the values that these nodes can take are represented by a branch. The nodes of the tree contain the decision variables and the leaves contain the predicted target values. Decision trees, which are frequently preferred in classification and clustering problems, are the most popular of supervised learning algorithms because they can be easily translated into human languages or programming languages. There are many algorithms developed based on decision trees. These algorithms are divided into different categories according to root, node and branching criteria. Commonly known decision tree algorithms are ID3, C4.5 and C5. C4.5 is one of the efficient and popular algorithms used to create a Decision Tree. There are many advantages to using a decision tree. Understanding and interpreting a problem is easy. In most of the alternative techniques, the data becomes available after a little processing. The preprocessing phase of the decision tree is shorter and simpler than other alternative methods. This method can be used to handle both numeric and class data. Most

machine learning algorithms are useful for numerical applications or classification problems. The results of this method presented in the figure 4.1.

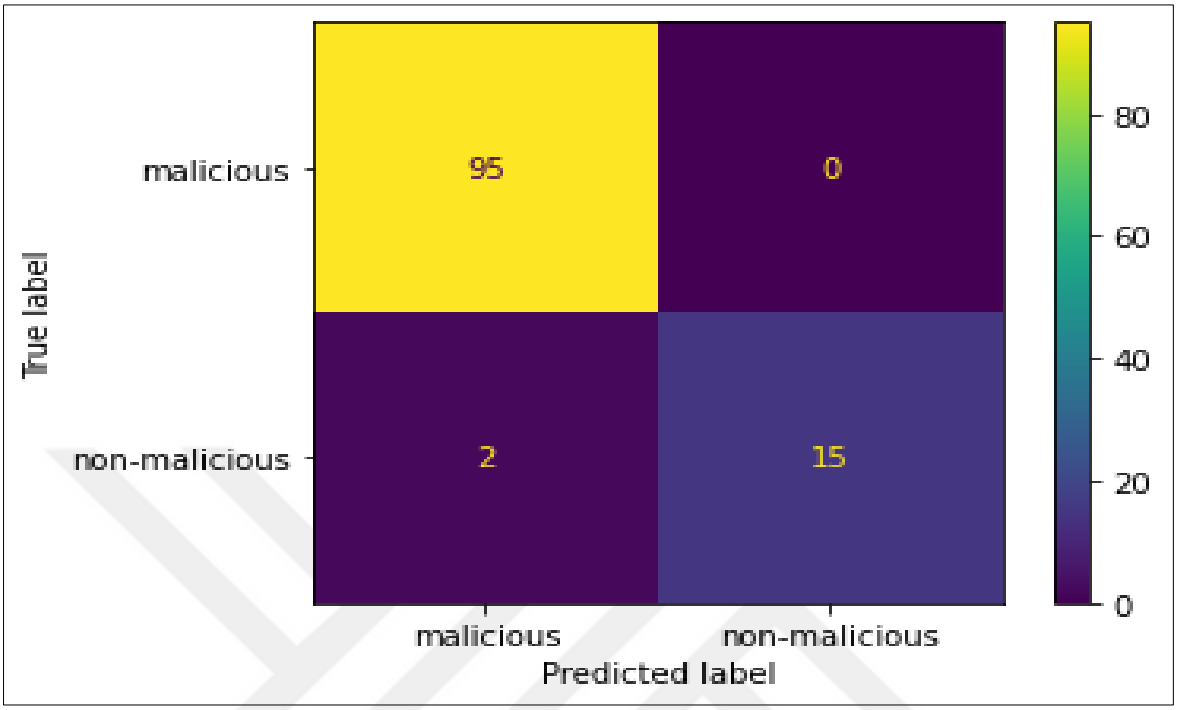

Figure 4.1: DT Results.

4.1.2 Random Forest Results

A Random Forest is a collection of independent decision trees based on random sampling from training data using the same distribution. The Random Forest consists of the number of trees to be created, the classification tree or the groups of the regression tree in accordance with its target. For this reason, it is one of the most preferred algorithms among the ensemble methods. The method is to generate ensembles by selecting a random subset from a large number of predictive trees (Breiman, 2001). It creates trees by applying Decision Tree Regression method on randomly selected sub-datasets. Generates a predictive value for all trees. It takes the average of the generated predictions and accepts it as its prediction value. In this method, more than one decision tree is created during the training of the dataset, and then the class of the input is decided by determining the majority vote, using the results of the classification of the decision trees during the estimation. There is no overfitting problem in random forests as in decision trees. The confusion matrix of RF presented in figure 4.2.

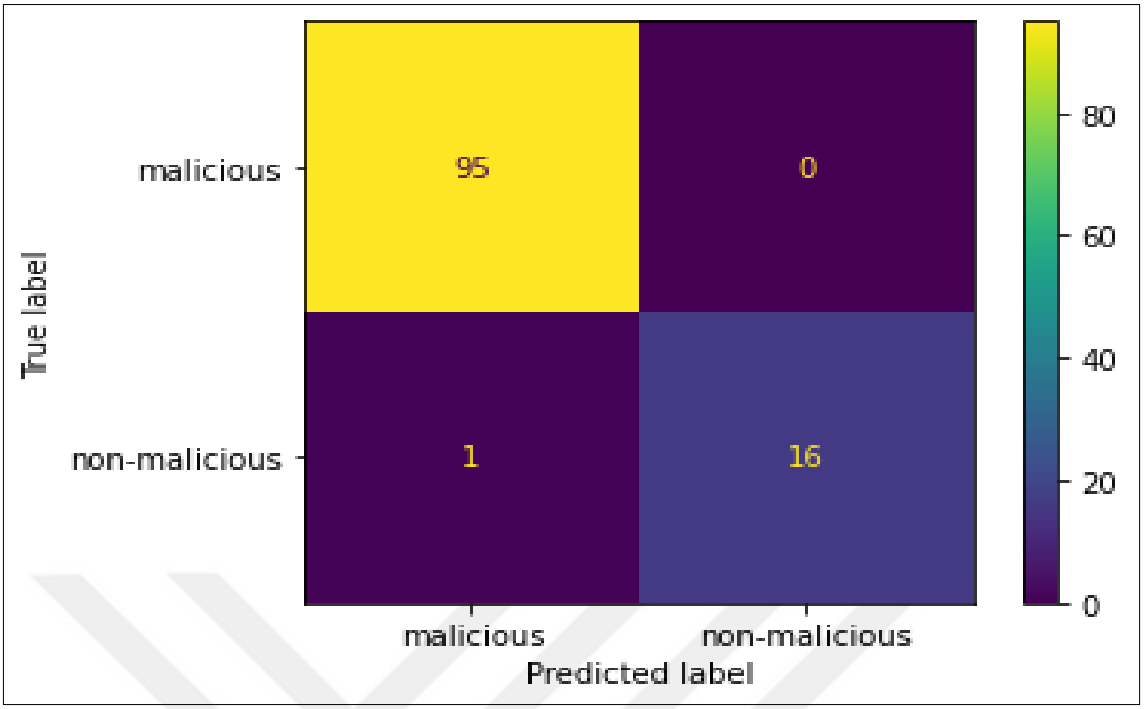

Figure 4.2: RF Results.

4.1.3 Neural Network Results

Artificial Neural Networks are a frequently preferred classification tool based on biological nerve cells and using neurons as building blocks. Here, artificial neurons are designed to be similar to biological neurons. The inputs given to the neurons are multiplied by the weights and summed in the summing unit. The sum mentioned here is given as an input to the transfer function. The output value in the transfer function is determined as the output value of the neuron. Many transfer functions are used in the ANN architecture. For example; While the sigmoid function is used for classification, the gaussian (normal) function can be used for function approximation. A single neuron is normally only used for linear problems. ANNs are intelligent systems that learn the relationships between input and output vectors. Because it is smart, it can establish a connection with the model learned by the new entries in the system. In Intrusion Detection Systems, ANN needs to be trained with system behavior traces or a data set with the same trend. Then, the flows specified as abnormal or normal are given to the ANN model. Learning with these data covers the process of training the user's movement or command by converging to the next movement or command. Figure 4.3 presented the structure of MLP.

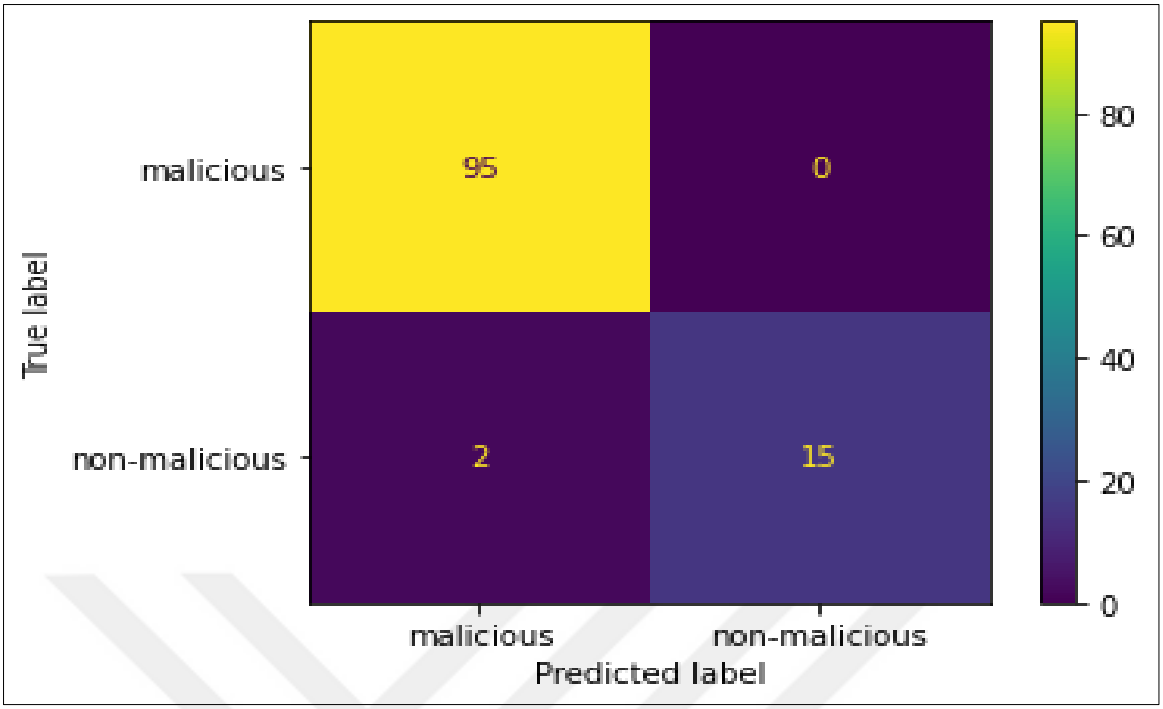

Figure 4.3: ANN Results.

4.1.4 MLP Results

MLP is one of the widely used machine learning techniques to increase the detection rate of intrusion detection systems. The training process for pattern classification problems of MLPs consists of two tasks; The first is to choose the architecture suitable for the problem, and the second is to adjust the connection weights of the network. MLP identifies typical characteristics of system users and identifies statistically significant deviations from established user behavior. In addition, the MLP neural network has flexible and extensible structures. Can construct a general knowledge model of the behavior of an environment. Requires estimation of neuron parameters to enable MLP to identify the relationship between the model and target output through training. The using MLP directly lead to some weakness such as in accuracy as shown below in Figure 4.4.

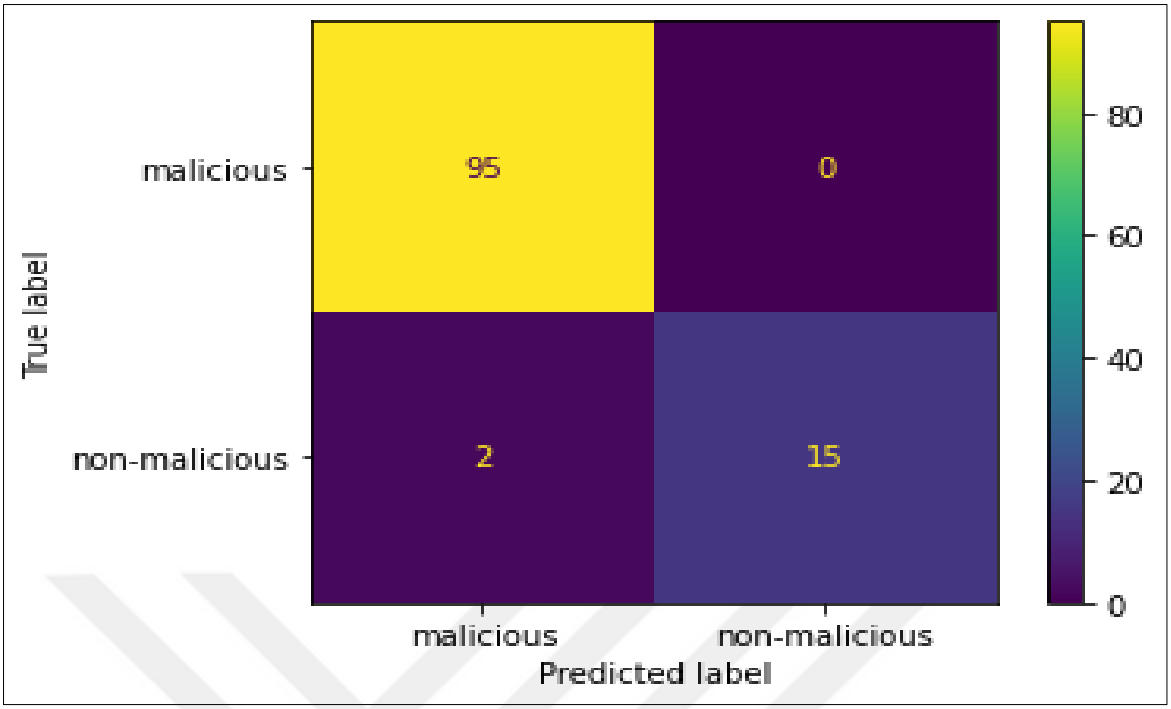

Figure 4.4: MLP

4.1.5 Proposed Method (Boltzmann Machine 1+ Boltzmann Machine 2 + AdaBoost) Results

Then, we applied the proposed method that combined the Boltzmann machine 1 with Boltzmann machine 2 and AdaBoost. Then, the results of this study presented in Figure 4.5.

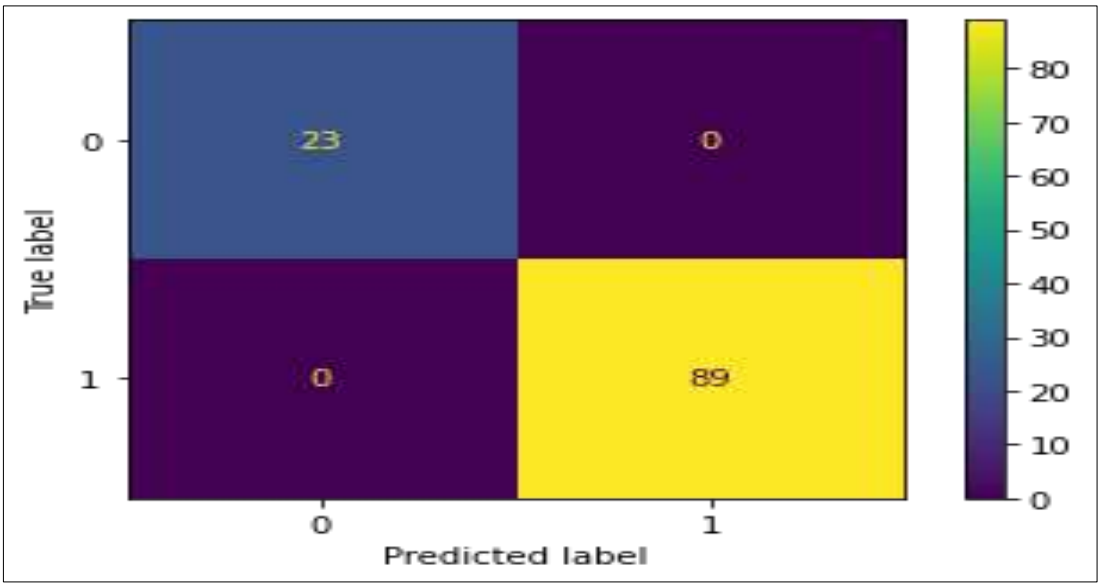

Figure 4.5: Boltzmann Machine 1+ Boltzmann Machine 2 + AdaBoost.

The Figure 4.6 represented the precision values of both techniques, Boltzmann machine 1+AdaBoost and Boltzmann machine 1+ Boltzmann machine 2 +AdaBoost.

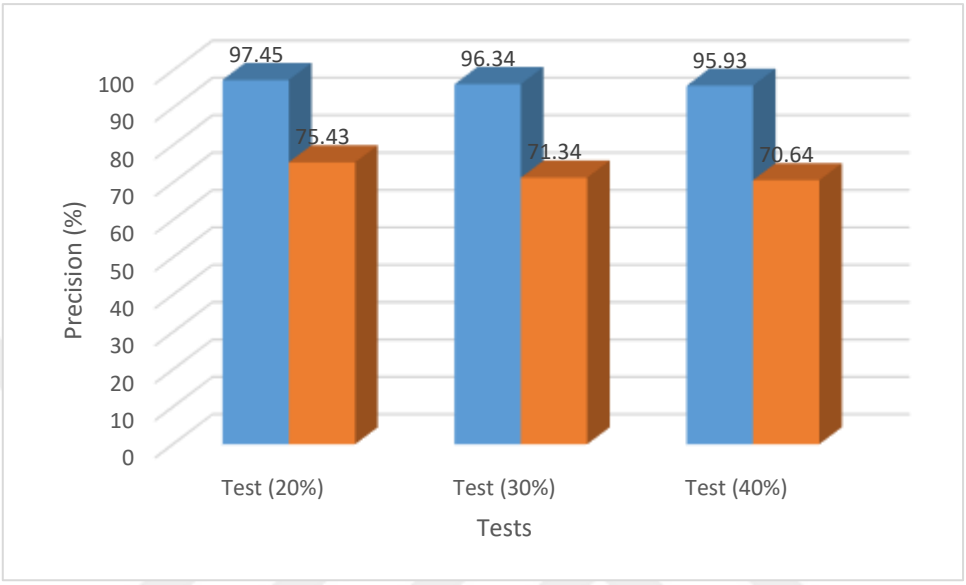

Figure 4.6: Compression of Precision Results of Two Techniques.

Furthermore, the Figure 4.7 represented the comparison between the proposed method and LSTM. By analysing the figure, we can calculate that the proposed method presented best Sensitivity results than LSTM.

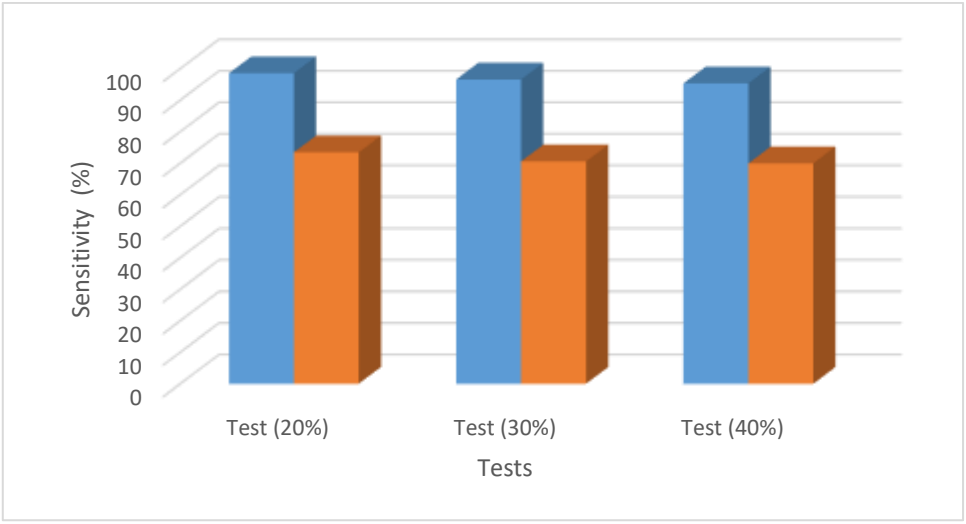

Figure 4.7: Compression of Sensitivity Results of Two Techniques.

5. CONCLUSIONS

In this thesis, an IoT home network prototype was created, which communicates with the MQTT protocol, which is frequently used in IoT. In the IoT home network, temperature, humidity and distance information were obtained from sensors with different network connections and recorded. Botnet, DoS, Brute Force and SlowDoS attacks were carried out on this network and the obtained data was added to the MQTTset dataset. It also integrates harmful, corrupted and legitimate data. Botnet attacks have been added to the MQTTset dataset and it has been expanded by adding data to other attacks. Data preprocessing and feature selection processes were performed on the MQTTset dataset. The MQTTset dataset is reduced to 21 features. The MQTTset dataset was evaluated with machine learning and deep learning algorithms using 5-fold cross validation. Separate tests were carried out with multi-class and two-class classification on the MQTTset dataset. In multiclass classification, AdaBoost gave the highest accuracy value, RF the highest precision value, RF the highest sensitivity value, and HT algorithms the highest F1 Score value. In the two-class classification, the RF algorithm gave the highest accuracy, precision, sensitivity and F1 Score values. MQTTset and IoT home network data were also evaluated according to the results of multi-class and two-class classification. HT, AdaBoost, RF algorithms gave the highest accuracy values in multi-class classification with MQTTset and IoT home network data, respectively. It has been seen that RF algorithms give the highest precision HT, sensitivity AdaBoost and F1 Score value. With MQTTset and IoT home network data, RF and AdaBoost algorithms give the highest accuracy values, HT and AdaBoost precision values, HT sensitivity values, and F1 Score values RF, HT, AdaBoost algorithms in two-class classification. It has been tested with 2- and 3-layer applications in deep learning methods. With the combination of data obtained from MQTTset and IoT home network, testing and analysis processes were carried out using the same methods and algorithms. In the MQTTset dataset, it was seen that only the precision value was higher in the deep learning results made with 3 layers than with 2 layers. Looking at the results in the MQTTset dataset in general, deep learning applications with 2 layers gave higher results. The highest values in deep learning applications with MQTTset and IoT 88 home network data were obtained from tests with 3 layers. In the framework of deep learning, it has been determined as a result of tests and analyzes that MQTTset and IoT home network data give higher

results. In addition, in this thesis study, tables were created as a result of literature research in the classification of attacks according to layers. In future studies, machine learning and deep learning methods for other protocols used in IoT can be tested as in this study and comparisons can be made on a protocol basis. Algorithms other than machine learning algorithms applied within the scope of this study can also be tested. In further studies, it is aimed to evaluate all types of attacks against communication protocols used in IoT networks, to create a new and more comprehensive dataset and to make it available for studies in this field.

REFERENCES

[1] Yu Liu, Ao Li, Xing-Ming Zhao, Minghui Wang, DeepTL-Ubi: “A novel deep transfer learning method for effectively predicting ubiquitination sites of multiple species”, Methods,2020.

[2] Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, Xuefeng Chen, Few-shot transfer learning for intelligent fault diagnosis of machine, Measurement, vol.166,2020.

[3] Shanshan Wang, Lei Zhang, Jingru Fu, “Adversarial transfer learning for cross-domain visual recognition”, Knowledge-Based Systems, vol.204, 2020.

[4] Xiyun Yang, Yanfeng Zhang, Wei Lv, Dong Wang, “Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier”, Renewable Energy,2020.

[5] Chuan Li, Shaohui Zhang, Yi Qin, Edgar Estupinan, “A systematic review of deep transfer learning for machinery fault diagnosis:, Neurocomputing, vol.407,pp. 121-135, 2020.

[6] Zhengjun Qiu, Shutao Zhao, Xuping Feng, Yong He, “Transfer learning method for plastic pollution evaluation in soil using NIR sensor”, Science of The Total Environment, vol.740, 2020.

[7] Santi Kumari Behera, Amiya Kumar Rath, Prabira Kumar Sethy, “Maturity status classification of papaya fruits based on machine learning and transfer learning approach”, Information Processing in Agriculture,2020.

[8] Seunghyeon Kim, Yung-Kyun Noh, Frank C. Park, “Efficient neural network compression via transfer learning for machine vision inspection,” Neurocomputing, vol.413, pp.294-304,2020.

[9] Xiang Li, Wei Zhang, Hui Ma, Zhong Luo, Xu Li, “Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks”, Neural Networks, vol.129, pp.313-322, 2020.

[10] Piyush Kant, Shahedul Haque Laskar, Jupitara Hazarika, Rupesh Mahamune, CWT Based Transfer Learning for Motor Imagery Classification for Brain computer Interfaces, Journal of Neuroscience Methods, vol.345,2020.

[11] Xinghua Li, and , et al , “Transfer learning based intrusion detection scheme for

Internet of vehicles", Information Sciences, vol.547, pp.119-135, 2021.

[12] Sheikh Saud, Basharat Jamil, Yogesh Upadhyay, Kashif Irshad, "Performance improvement of empirical models for estimation of global solar radiation in India:" A k-fold cross-validation approach, Sustainable Energy Technologies and Assessments, vol.40, 2021.

[13] Jie Wei, Hui Chen, "Determining the number of factors in approximate factor models by twice K-fold cross validation", Economics Letters, vol.191,pp.109-149, 2020.

[14] Tzu-Tsung Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation", Pattern Recognition, vol.48,2015.

[15] Gaoxia Jiang, Wenjian Wang, "Error estimation based on variance analysis of k-fold cross-validation", Pattern Recognition, vol.69, pp.94-106, 2017.

[16] Nima Shiri Harzevili, Sasan H. Alizadeh, Analysis and modeling conditional mutual dependency of metrics in software defect prediction using latent variables, Neurocomputing, vol.460, pp.309-330, 2021.

[17] Chao Ni, Xiang Chen, Fangfang Wu, Yuxiang Shen, Qing Gu, "An empirical study on pareto based multi-objective feature selection for software defect prediction," Journal of Systems and Software, Vol.152,pp.215-238, 2019.

[18] Ruchika Malhotra, Shine Kamal, "An empirical study to investigate oversampling methods for improving software defect prediction using imbalanced data," Neurocomputing, vol.343, pp.120-140, 2019.

[19] Yerima, S. Y., Alzaylaee, M. K., & Sezer, S. "Machine learning-based dynamic analysis of Android apps with improved code coverage". EURASIP Journal on Information Security, 2019.

[20] Siddiqui, M., Wang, M. C., & Lee, J. "Data mining methods for malware detection using instruction sequences". In Artificial Intelligence and Applications , 2008.

[21] Ahmad Karim, "Development of secure Internet of Vehicle Things (IoVT) for smart transportation system", Computers and Electrical Engineering, vol.102, 2022.

[22] A. Pektaş, M. Çavdar, T. Acarman, "Android malware classification by applying online machine learning", (ISCIS 2016) International Symposium on Computer and Information Sciences, Kraków, Poland, 2016.

[23] Karim AM, Kaya H, Alcan V, Sen B, Hadimlioglu IA. "New Optimized Deep Learning Application for COVID-19 Detection in Chest X-ray Images". *Symmetry*.

2022.

[24] S. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. Chan, “Cost-based modeling for fraud and intrusion detection: results from the jam project, in: DARPA Information Survivability Conference and Exposition,” DISCEX’00, Proc., vol. 2, 2000.

[25] M. Alkasassbeh, A. B. A. Hassanat, and G. Al-naymat, “Detecting Distributed Denial of Service Attacks Using Data Mining Techniques,” vol. 7, no. 1, pp. 436–445, 2016.

[26] A. Abraham, C. Grosan, and C. Martin-Vide, “Evolutionary design of intrusion detection programs,” Int. J. Netw. Secur., vol. 4, no. 3, pp. 328–339, 2007.

[27] A. M. Brues, “Genetic effects of the atom bomb,” J. Hered., vol. 38, no. 5, pp. 137–137, 1947.

[28] H. Liu, Y. Sun, and M. S. Kim, “Fine-grained DDoS detection scheme based on bidirectional count sketch,” Proc. - Int. Conf. Comput. Commun. Networks, ICCCN, 2011.

[29] C. Khammassi and S. Krichen, “A GA-LR wrapper approach for feature selection in network intrusion detection,” Comput. Secur., vol. 70, pp. 255–277, 2017.

[30] A. Abraham and J. Thomas, “Distributed intrusion detection systems: a computational intelligence approach,” Idea Gr. Inc. Publ. Usa,vol. 5, pp. 1–28, 2005.

[31] N. Sharma and S. Mukherjee, “A Novel Multi-Classifier Layered Approach to Improve Minority Attack Detection in IDS,” Procedia Technol., vol. 6, pp. 913–921, 2012.

[32] B. Škrbić and N. Durišić-Mladenović, “Principal component analysis for soil contamination with organochlorine compounds,” Chemosphere, vol. 68, no.11, pp. 2144–2152, 2007.

[33] N. Patani and R. Patel, “A Mechanism for Prevention of Flooding based DDoS Attack,” vol.13, no.1, pp. 101–111, 2017.

[34] L. K. Xr et al., “8VLQJ 6WDFNHG ’ HQRLVLQJ $ XWRHQFRGHU IRU WKH,” pp. 483–488, 2017.

[35] Y. Feng, R. Guo, D. Wang, and B. Zhang, “Research on the active DDoS filtering algorithm based on IP flow,” 5th Int. Conf. Nat. Comput. ICNC 2009, vol. 4, no. October, pp. 628–632, 2009.

[36] A. Meek, “DDoS attacks are getting much more powerful and the Pentagon is scrambling for solutions,” 2015.

[37] J. Mirkovic and P. Reiher, “A taxonomy of DDoS attack and DDoS defense mechanisms,” ACM SIGCOMM Comput. Commun. Rev., vol. 34, no. 2, p. 39, 2004.

[38] S. Taghavi Zargar, J. Joshi, D. Tipper, and S. Member, “A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks,” pp. 1–24, 2013.

[39] T. T. Oo and T. Phyu, “Analysis of DDoS Detection System based on Anomaly Detection System,” 2014.

[40] S. Sinha and M. Sharma, “Simulation and Analysis of DDoS Attacks by Specialized Simulator using Virtualization,” vol. 3, no. 2, pp. 2–4, 2014.

[41] Bojan Kolosnjaji, Apostolis Zarras, George Webster, and Claudia Eckert. “Deep learn- ing for classification of malware system call sequences”. In Australasian Joint Confer- ence on Artificial Intelligence, pp.137–149, 2016.

[42] B. Hang, R. Hu, and W. Shi, “An enhanced SYN cookie defence method for TCP DDoS attack,” J. Networks, vol. 6, no. 8, pp. 1206–1213, 2011.

[43] D. Wang, Z. Yufu, and J. Jie, “A multi-core based DDoS detection method,” Proc. - 2010 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2010, vol. 4, pp.115–118, 2010.

[44] N. Hoque, H. Kashyap, and D. K. Bhattacharyya, “Real-time DDoS attack detection using FPGA,” Comput. Commun., vol. 110, pp. 48–58, 2017.

[45] J. Singh, M. Sachdeva, and K. Kumar, “Detection of DDoS Attacks Using Source IP Based Entropy,” vol. 3, no.1, pp. 201–210, 2013.

[46] S. M. Lee, D. S. Kim, J. H. Lee, and J. S. Park, “Detection of DDoS attacks using optimized traffic matrix,” Comput. Math. with Appl., vol. 63, no.2, pp.501–510, 2012.

[47] H. Rahmani, N. Sahli, and F. Kamoun, “DDoS flooding attack detection scheme based on F-divergence,” Comput. Commun., vol.35, no.11, pp.1380–1391, 2012.

[48] Senirkentli, Güler B., Fatih Ekinci, Erkan Bostanci, Mehmet S. Güzel, Özlem Dağli, Ahmad M. Karim, and Alok Mishra. "Proton Therapy for Mandibula Plate Phantom" Healthcare 9,2021.

[49] V. P. Nigam and D. Graupe, “A neural-network-based detection of epilepsy,” Neurol. Res., vol.26, no.1, pp. 55–60, 2004.

[50] A. Abraham, U. States, and C. Grosan, “Genetic Systems Programming,” vol.13, no.

May, 2006.

[51] L. Deng and D. Yu, “Deep Learning: Methods and Applications,” Found. Trends® Signal Process., vol.7, no.3–4, pp.197–387, 2014.

[52] K. Ahmed, K. Nia, S. A. Khan, and A. Shaukat, “Identifying Be est Feature Subset for Ca ardiac Arrhy ythmia Classification,” pp. 494–499, 2015.

[53] A. Shenfield, D. Day, and A. Ayesh, “Intelligent intrusion detection systems using artificial neural networks,” ICT Express, vol. 4, no. 2, pp. 95–99, 2018.

[54] A. Kaushik, H. Gupta, and D. S. Latwal, “Impact of Feature Selection and Engineering in the Classification of Handwritten Text,” 2016 Int. Conf. Comput. Sustain. Glob. Dev., pp. 2598–2601, 2016.

[55] A. Gupta, A. T. Müller, B. J. H. Huisman, J. A. Fuchs, P. Schneider, and G. Schneider, “Generative Recurrent Networks for De Novo Drug Design,” Mol. Inform., vol. 37, no.1, 2018.

[56] S. Ibrahim, R. Djemal, and A. Alsuwailem, “Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis,” Biocybern. Biomed. Eng., vol. 38, no. 1, pp.16–26, 2018.

[57] N. Kohli, N. K. Verma, and A. Roy, “SVM based methods for arrhythmia classification in ECG,” 2010 Int. Conf. Comput. Commun. Technol. ICCCT-2010, pp. 486–490, 2010.

[58] V. Sze, Y.-H. Chen, T.-J. Yang, and J. Emer, “Efficient Processing of Deep Neural Networks,” A Tutorial and Survey, vol. 105, no. 12, pp. 2295–2329, 2017.