
REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

# CLASSIFICATION OF TRAFFIC AND IOT DEVICES THROUGH ARTIFICIAL NEURAL NETWORK AND FEATURE SELECTION

**Ahmed Thamer Hameed HAMEED**

Master's Thesis

Supervisor

Prof. Dr. Osman Nuri UÇAN

Istanbul, 2022

# CLASSIFICATION OF TRAFFIC AND IOT DEVICES THROUGH ARTIFICIAL NEURAL NETWORK AND FEATURE SELECTION

**Ahmed Thamer Hameed HAMEED**

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled CLASSIFICATION OF TRAFFIC AND IOT DEVICES THROUGH ARTIFICIAL INTELLIGENCE AND FEATURE SELECTION prepared by AHMED THAMER HAMEED HAMEED and submitted on 9/12/2022 has been **accepted unanimously** for the degree of Master of Science in Information Technologies.

Prof. Dr. Osman Nuri UÇAN

The Supervisor

Thesis Defense Committee Members:

| | | |
|---|---|---|
| Prof. Dr. Osman Nuri UÇAN | Department of Electrical Electronics Engineering, Altınbaş University | _____ |
| Asst. Prof. Dr. Oğuz ATA | Department of Software Engineering, Altınbaş University | _____ |
| Prof. Dr. Hasan Hüseyin BALIK | Department of Computer Engineering, Yıldız Technical 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 Thamer Hameed HAMEED

signuture

# DEDICATION

I dedicate and promise this study effort to my supervisor, who has been instrumental in helping me through the entire research process, as well as my family, who has always been supportive during my difficult times.

# PREFACE

First and foremost, I want to express my gratitude to my supervisor, Prof. Dr. Osman Nuri UAN, for guiding and assisting me during the composition of this dissertation. Every week, I met with my supervisor, Prof. Dr. Osman Nuri UAN, to discuss my progress, challenges, and suggestions. This helped me enormously in grasping the reasoning behind the research. It helped me understand the technological requirements for this study project.

# ABSTRACT

## CLASSIFICATION OF TRAFFIC AND IOT DEVICES THROUGH ARTIFICIAL NEURAL NETWORK AND FEATURE SELECTION

Hameed, Ahmed Thamer Hameed

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

Supervisor: Prof. Dr. Osman Nuri UÇAN

Date: December /2022

Pages: 46

With the expansion of technology spread at the present time, devices have become highly connected to the Internet. A lot of data is collected from the daily use of devices connected to the Internet. The connection of devices to the Internet makes them important sources of data. The most important characteristic of IoT data is its abundance and its large sizes, and for the purpose of building technical systems based on that data, these data should be classified for ease of study and analysis. Close interconnection and interconnected networks make traffic abundantly distributed within the data of IoT devices. AI algorithms and tools can be used to analyze and classify data and provide detailed reports on the state of data for the Internet of Things.

Providing a lot of neural networks, especially the artificial neural network, makes choosing features and knowing important data a very easy job by using hidden nested layers that classify the data and choose its features. A neural network trained on real data for most of the devices connected to the Internet was used for the purpose of classification. The system was operated with high accuracy compared to most of the studies presented in this field. Our proposed study can be considered as an important functional tool for using artificial neural networks to extract and classify the features of IoT data.

**Keywords:** IoT, AI, neural networks, analyze, classify.

# TABLE OF CONTENTS

**Pages**

# LIST OF TABLES

**Pages**

# LIST OF FIGURES

**Pages**

# ABBREVIATIONS

| | | |
|---|---|---|
| NN | : | Neural Network |
| IOT | : | Internet Of Things |
| KNN | : | k-Nearest Neighbors |
| SVM | : | Support Vector Machine |
| AMC | | Automatic Modification Classification |
| FFM | : | Fast Fourier Method |
| CNN | : | Conventional Neural Network |

# 1. INTRODUCTION

The existence of data requires the provision of the correct management of it. With the increase in the number of devices connected to the Internet, we have recently found a large number of Internet of Things data and an increase in its use in most areas. The connection of devices to the Internet requires protection and security for the data collected from them. The challenges are the early and timely detection of technical problems in the devices. Knowing the state of the Internet of Things device, its mechanism of action and its behavior helps a lot in knowing the purpose of using its data. Traffic classification methods are used in IoT data for the purpose of studying the data completely. The traffic classification process extracts unique keywords that include the associated devices, their domains, and their mechanism of action to determine which package belongs to the device. Feature extraction is automated within keyword classification and using artificial intelligence algorithms.

Network traffic can be categorized in several ways to create a suitable tool that can be used in a computer network [1]. Previously, it relied on machine learning algorithms to choose different forms of network traffic. Traffic classification can be used to identify problems and constraints and also to improve the quality of network service. Network data flows procedures limit device traffic. These movements can be categorized to identify smart home devices connected to the Internet, personal computers, medical and engineering devices. The performance resulting from using supervised or machine learning algorithms differs greatly in determining the goal of building smart data classification systems.

The existence of a network of devices connected to the Internet and they are in the case of sharing with each other to generate a huge amount of data that need a central location to store that data. The devices in the Internet of Things operate on the principle of the necessity of connecting to a network and therefore are vulnerable to attacks [2]. The data, like other private data, needs management tasks, security measures, and the need for prior detection of the intrusion process. The main source of data is providing quality of service, performance monitoring, resource provisioning, traffic engineering and traffic classification.

## 1.1. RESEARCH MOTIVATION

With the development of technology, many institutions and companies have transferred data management to the modern method due to the ineffectiveness of traditional management schemes. Providing a classification based on machine learning systems is one of the important methods in building systems for extracting features in the data of devices connected to the Internet. Machine learning-based traffic classification systems based on artificial neural networks have been widely used in the literature [3]. We compare the results obtained from building our model from the previous literature. An artificial neural network has been developed and built that contains internal layers that are hidden by variables that make it work with high accuracy in classifying traffic for Internet of things data. Machine learning-based models are not good at classifying Internet traffic due to incorrect selection of features. In order to obtain high-fidelity performance, data is preprocessed and stream features extracted using an artificial neural network.

Another reason derives from big data itself, despite the significant concern that this data generates, but it is the cost that we desire. In reality, we create many methods and gadgets to collect this information. As a result, we are looking for the best ways to govern and profit from this data.

## 1.2. PROBLEM DEFINITION

The data of IoT devices is one of the most important resources of artificial intelligence systems. Artificial intelligence systems include all information technology systems that provide functions and services that help facilitate tasks. AI can analyze questions, discover data in them, and then automatically make a decision based on extensive data sets. There are many branches of artificial intelligence, including machine learning. Computers have the ability to learn from data after entering it into special algorithms to perform that function.

The close link between smart systems and data, including IoT data, helps build models to identify hidden patterns (correlations) without the need for actual programming. The presence of unclassified data for Internet of Things devices does not allow it to be used in similar systems, which prompts researchers to find the most important ways to create tools that help classify it and extract features from it.

## 1.3. AIM AND OBJECTIVES

Big Data Analytics (BDA) generally refers to large amounts of data that can be generated, processed, and increasingly used by digital tools and information systems to perform predictive, descriptive, and didactic analysis [4]. The design and implementation of similar systems need prior steps of data processing, accuracy and high efficiency.

Providing a tool for classifying IoT devices data is useful to know the current situation and to rely on IoT correctly. Groups of decision makers can identify research and development initiatives by studying and analyzing data and extracting features from it. The data classification process helps data management institutions in maintaining the confidentiality, accessibility and integrity of their data. For unstructured data in particular, data classification reduces the vulnerability of sensitive information. Having systems to classify data and extract features from it after collecting IoT data [5] helps simplify the task of data deployment, storage, sensor selection, processing power, and security.

## 1.4. THESIS ORGANIZATION

The chapters of this thesis are arranged to cover the most important information on the subject and a full scientific study regarding the previously presented literature in this field. In the second chapter, we review the research and literature that dealt with the subject of the Internet of Things and ways to deal with it, as well as review the literature whose results have been improved. In the third chapter, we present a comprehensive theoretical study of all the tools used in our research, the methodology of the proposed method, and the general outline. In the fourth chapter, the experiences and results obtained are studied, and then the discussion and recommendations for future work.

# 2. LITERATURE REVIEW

## 2.1. TRAFFIC AND IOT DEVICES IN PAST LITERATURES

Recently, the majority of researchers have focused on IoT data. Antônio J et al. presented a thorough study [6] on recognizing IoT devices and events based on the length of their encrypted data packets. The objective of this study was to investigate the potential privacy issues that could come from deploying machine learning algorithms in IoT devices. Traffic classification has numerous moving aspects, making it difficult to learn. The creation of data buses is highlighted as a complex way for collecting TCP protocol characteristics. Strongly recommended is the establishment of a technique to classify encrypted packet traffic and IoT device behavior and provide statistics on both. As part of this research, methods for classifying data from connected and unconnected IoT devices are being developed. Using five distinct classifiers and two distinct traffic data-based realistic test scenarios, the data was collected. Among the most popular methods for traffic classification were models employing k-Nearest Neighbors (k-NN), decision trees, random forests, support vector machines (SVM), and majority voting. Using normal hypothesis-testing procedures, the results were examined. Among the five tested classifiers, the decision tree exhibited the smallest delay for picking devices. These results were accurate.

In their study on IoT data, Hamid Tahaei and colleagues [7] noted that connecting to various IoT gadgets is becoming a regular practice due to the integration and proliferation of numerous products. After the proliferation of IoT devices and the diversification of IoT traffic patterns, traffic categorization techniques have become indispensable. Despite the fact that several studies, surveys, and research publications have successfully classified network traffic, the idiosyncrasies of traffic in IoT devices and devices that do not support IoT have led to the immaturity of traffic classification in these devices. When discussing the numerous sorts of Internet of Things-related traffic, the researchers focused on the conventional models. This article examines the existing classifications of traffic in IoT networks, reviews the categorization of IoT traffic from the perspectives of commercial, economic, and realistic service, and discusses the challenges and issues in this field.

Ahmad Azab et al. [8] covered the topic of classifying network traffic. Learn how to tie network traffic and applications, protocols, or groups to the underlying service in order to aid in legal interception, ensure service efficiency and quality, reduce application bottlenecks, and overcome obstacles in recognizing potentially dangerous network behaviors. Port-based identification, which depends on in-depth packet

inspection, statistical properties in addition to machine learning, and deep learning algorithms are among these categorization methodologies. In addition, we discuss their usefulness and prospective uses, as well as any limitations associated with them. In addition to addressing present challenges, the article speculates on the future of the field.

SaadatIzadi et al. [9] authored an essay on network traffic classification. According to them, traffic has become increasingly difficult to navigate in recent years. A plethora of recent studies have addressed the difficulties of investigating deep learning approaches. These sorts of study encounter obstacles, such as the need for enormous amounts of data for training and the requirement of expert feature extraction. Deep learning strategies are essential due to the difficulties inherent in effective classification, namely the difficulty of identifying valuable characteristics. Using methods of deep learning, we can automatically and effectively extract these attributes. In order to solve these issues, this study details efforts to merge convolutional neural networks (CNN), ant heuristic algorithms (ALO), and self-organizing maps (SOM) in order to create models capable of accurately classifying and detecting traffic's multiple categories. This approach identified encrypted communication and discriminated VPN traffic from other traffic types. Using the ISCX VPN-non-VPN dataset, the model demonstrated a high degree of accuracy and hence passed the examination.

Sangita Roy et al. [10] proposed a new methodology for traffic classification. They demonstrated that current encryption makes it more difficult to identify a certain program or category of network traffic. One of the challenges this movement must overcome is data encryption via a virtual private network (VPN). Maximizing the efficacy of deep learning gave classification solutions for traffic. The duration of these solutions, which can range between 15 and 60 seconds, is one of its distinguishing characteristics. This is because they either analyze a large volume of data flows or rely on manually built characteristics to expedite classification of the flows. This study begins by proposing a novel algorithm for categorizing flows that minimizes packet sizes and interlayer arrival times. The model that calculates the hidden state derivative of the data flow using access timings and a neural network (Neural ODE). By comparing their results to a solution that did not rely on ODE but utilized the same data, the researchers highlighted the advantages and high likelihood of success of this method. These results permit the classification of streams based on the mapping of 20 or 30 consecutive packets from any location along the stream in a single direction. This method minimizes the quantity of data exchanged between the sample site and the analyzer since it does not rely on the matching of two data flows. Using this method, you can accurately

classify traffic in a matter of seconds. Researchers also developed methods for combining this method with classification solutions characterized by very high accuracy and slow speed, enabling classifications with both high speed and accuracy.

Automatic Modulation Classification (AMC) was recognized by JieYin and colleagues [11] as a key technology for identifying unknown signals in IoT devices and cognitive radios. Automatic Modification Classification has recently included deep learning to improve the classification process (AMC). Large model sizes and high computational complexity, which are distinguishing characteristics of heterogeneous (IOT) systems, make it difficult to develop and use DL-based AMC solutions. Because of this, the ShuffleNet idea prompted the development of SuffleCNN, a lightweight convolutional neural network (CNN). The ShuffleCNN-based AMC approach is recommended for all Internet of Things (IoT) and electrical physical systems employing orthogonal frequency division. Fast Fourier analysis was presented as a method for efficiently classifying and operating with orthogonal frequency division signals as well as detecting an over-processing structure (FFT). The simulation results demonstrated that the ShuffleAMC method reduces performance declines when compared to CNN-based AMC alternatives. Among them are the approach's low computational complexity and compact size.

Rajiv Yadav and coworkers [12] developed a new strategy to handle the security difficulties and threats facing wireless sensor networks in view of the growing use of Internet of Things applications and the resulting increase in network complexity. Using the recently developed Fast Link-Based Feature Selection (FCBFS) in conjunction with XG-Boost on the reference dataset permits the prediction of security incidents, risks, and threats. As a pre-classification stage, this technique can assist in wireless and cluster-based sensor networks by highlighting the most essential characteristics. IoT and WSN applications leverage machine learning-based classifiers for effective anti-threat, intrusion, and intrusion detection. The random forest, decision tree, additional tree, XG-Boost, and Naive Bayes are among examples. In addition to the score of F, the effectiveness of a recall, the simplicity with which information can be recalled, and the accuracy with which it can be categorized were all examined. The XG-boost model method achieved superior levels of classical accuracy, retrieval success rates, and overall performance when compared to previous methods in this discipline. It is evident from these results that the suggested method has several advantages over current filtering techniques, including a smaller number of necessary features, higher recognition rates, and more exact classifications. In addition to reducing sensor power consumption during intrusion detection and assault, the suggested solution

prolongs the network's operational life.

Ticiana Capris and others [13] call attention to the CORONA disease, which has killed millions of people worldwide and brought a temporary stop to human life. Since the beginning of this outbreak, numerous individuals have pondered what can be done to stop the development of this ailment (COVID-19). People use masks, hand sanitizers, and avoid touching doors to prevent the spread of viruses. The authors of this research propose using the Convolutional Neural Network (CNN) model to assess an individual's facility-use eligibility and health status. Researchers utilized IoT sensors to confirm that guests were washing their hands after using hand sanitizer. A trained model was also used to determine whether an individual was concealing their identity. These results indicate the concept's great effectiveness.

The focus of the study by Yoshinobu Yamashita et al. [14] was on Internet traffic from mobile devices. Currently, mobile devices provide a greater proportion of Internet traffic than Internet of Things devices. Multiple simultaneous connection requests can cause incompatibilities. Due to the argument, the cell network is congested and tense. The investigation of this issue led to the development of an initial viable solution. It has been examined how data-generating gadgets affect both time and space. The key concept of this study is that the communication delays between devices, which were previously used to generate a temporal discharge, can also generate a spatial discharge. The evaluation time was chosen to assess the dependability of network and data transfer. As a technique of controlling the connection's timing, the offloading of spatial and temporal traffic from the transmission of IoT devices has been suggested. The novel strategy relocated a portion of peak traffic to off-peak hours. The assessment comparison approach was able to limit traffic growth while still meeting the on-time access requirement for IoT device data production.

FadiAl-Turjman, a data scientist, undertook research to enhance and replicate the infrastructure of IoT devices [15]. Due to the growing number of Internet-connected devices, it addresses the issue of data overload in the same manner as previous research. For the expansion of the reach of future communication systems, mobile and stationary Internet-connected devices are equally viable possibilities. The proposed recruitment technique included the repetition of information distribution. By using this new strategy, priority will be given to IRs in high-demand regions. The basic objectives are to decrease network traffic, reduce the average traffic load per publication, and keep latency to a maximum. The method also examines the behavior of the IoT model in real-world traffic circumstances. In other words, you are acquiring a new vocabulary to explain your surroundings.

Lopez et al. presented the findings of a study in which they constructed a scaling-boosting neural network and used it to anticipate IoT traffic. The ability to forecast network traffic is a crucial operational indicator for any data network. The combination of supervised regression with a novel deep learning architecture has been proposed. The model's structure is an additive network of "learning blocks" built in uniform layers. The final structure was trained using random gradient descent from scratch. This study considers the modernized network architecture that utilizes the surviving networks. The method was designed to assist the training of complex models and incorporates an outstanding convergence behavior. The given prediction outcomes outperform gradient enhancement model outcomes. The building blocks are the essential units of any learning or neural network architecture. Utilizing historical data, the model is used to estimate future network traffic in order to obtain empirical data on the challenging prediction problem. It is demonstrated that the system's efficiency and precision are increasing, hence enhancing the model's performance..

## 2.2. RELATED WORKS

Given the topic's significance, numerous studies have previously been undertaken in this field. All academic research is evaluated using algorithmic measures of precision. We examine the relevant literature in order to develop our procedures and increase our level of precision. Using machine learning, Dhanke JyotiAtul and others [17] addressed the problem of safeguarding Internet of Things (IoT) systems against intrusion. The Physical Cyber System can give a variety of services and amenities to its consumers due to its capacity to connect to many types of electronic equipment (CPS). Among the greatest issues this system faces are threats to the integrity of these systems, anomalies, and service disruptions. These issues can only be remedied by implementing a system that is both dependable and effective. The Energy Aware Smart Home (EASH) framework was developed by this study following a comprehensive analysis of all potential system concerns. This article analyzes communication difficulties and potential network threats. Machine learning is one such technique that assists in identifying the problematic sources of the communication model. The effectiveness and accuracy of this work's execution were reviewed. Up to 85% success was obtained, and it is anticipated that this percentage will increase as technology advances.

According to research by Jamila Manan et al. [18], intrusion detection is implemented using next-

generation networks. The Next Generation Network, also known as the IP Multimedia Subsystem (IMS), offers a range of hypermedia services to its customers, including but not limited to text, images, audio, and video. Increasing demand for IMS-provided services renders these systems increasingly vulnerable to assault. The introduction of an intrusion detection system (IDS) was a critical step in the evolution of network security in the face of attacks that are becoming increasingly sophisticated and ubiquitous. The security of users' private information and the secrecy of their communications is one of the most critical components of IMS networks. It is also largely the result of the SIP protocol, which provides many precautions against electronic attacks. This study's objective was to evaluate this technique within the context of the registration procedure and to identify its key flaws. In this study, DDoS attacks against IMS servers were undertaken via the SIP protocol, specifically via the REGISTER and IDS SNORT messages.

This work by Ensieh ModiriDovom et al. [19] used the fuzzy pattern tree to classify harmful attacks and applications. It is evident that IoT networks, often known as the "Internet of Things," have emerged due to the rapid surge in IoT-related hardware and software development (big data). As a result of this computational effort, a variety of new approaches and difficulties have emerged. Internet of Things advanced computing services include decentralizing, distributing, and transferring accounts to IoT nodes. Significant progress has been made in network security with the availability of tools that can consistently detect potentially hazardous digital activities. The feasibility of translating live software into a vector space and employing quick ambiguous tree topologies to recognize and categorize intrusions is studied in the proposed paper. Using the fast blur pattern tree, accuracy rates in excess of 90% were achieved. ShailendraRathore et al. [20] investigated the problem of the security threat provided by the hazards of electronic attacks on networks and how feature extraction and fuzzy classification were utilized to develop a more effective approach for detecting and classifying dangerous applications. Currently, the development of strategies for spotting centralized assaults in IoT systems is an important topic of research and development. It was proposed that a central hub be established to monitor the network for threats. This system examines network data using a machine learning algorithm and labels it as "attack" or "normal" based on the level of threat it poses to the network. Due to the Internet of Things' particular requirements for scalability, widespread availability, limited resources, and low latency, these strategies were unable to produce the intended outcomes. For categorization, supervised machine learning approaches require a substantial amount of labeled data. Using the fog computing paradigm and the semi-

supervised Fuzzy C Method (ESFCM) based on ELM, an effective system for detecting attacks and threats has been designed. Fog computing's many benefits include its ability to prevent network-peripheral intrusions and its support for distributed intrusion detection systems. The ESFCM technique employs the Extreme Learning Machine (ELM) technology in order to achieve robust generalization at a faster rate of attack detection. Utilizing the semi-supervised fuzzy technique, the issue of disaggregated data is addressed. The suggested system was evaluated using the NSL-KDD dataset, and the results indicate that it outperforms the industry benchmark for attack detection. The accuracy rate was 86.53 percent, and it took less than 11 milliseconds to identify an invasion.

Anita Patience Namanya et al. [21] detailed other ways for identifying malware in the Internet of Things. The prevalence of malicious files is a factor in the current epidemic of cyberattacks. The approach of similarity segmentation was utilized to classify samples for malware access testing. The malicious files are categorized into families based on their commonalities, and a signature is subsequently generated for each family. This study examines the four most common hashing algorithms used to analyze and identify malicious executable and portable (PE) files. Each of the four methods produces decent results, but each has significant shortcomings when it comes to detecting fraud. In order to generate measurable data with a high rate of malware detection, this study studied how to combine these numerous approaches. Based on the hashing discoveries, the researchers provided a new approach for registering malicious software. Studies using accuracy-based evaluation criteria found promising results with actual detection rates for malicious files reaching over 90%, suggesting that this approach is worth further investigation..

# 3. METHODOLOGY

## 3.1. OUTLINES

In this chapter, we define the tools and algorithms used in the design and construction of our proposed model. The system architecture is built to include analysis, classification, and feature extraction from IoT device data. The construction of the system was based on the construction of an artificial neural network with special inner layers. Activation functions are used to reach the optimum accuracy compared to the rest of the literature. Through the activation function we get the desired final value, and the signal passes through it to perform processing by way of an exponential relationship.

We will discuss the neural network and its types used in building smart systems, then the Internet of Things and how to obtain its data. The data set used in our proposed system is displayed, we perform analysis and prior operations to initialize the data. The steps for building the model are indicated in the model parts diagram.

## 3.2. NEURAL NETWORK NN

A virtual group of neurons that are interconnected by default. The main function of these interconnected groups is to imitate biological cells in terms of work or electronic chips that are designed to mimic the work of neurons called electronic structures [22]. Neurons are designed through algorithms implemented by computer programs to reach the desired goal of their work, and mathematical models are used to process information. These cells are cells that handle simple tasks, and the connections that are made between these different cells determine the general behavior of the cells and determine the indicators of these cells. It was proposed to create this network of virtual neurons to be a network that imitates the neural network of human brain cells, which is similar to a biological electrical network to process the information received into the human brain [23]. The idea of creating this network was proposed by Donald Hebb, who confirmed that the synapse has a key role in processing processes, and this point was the first spark in the process of communication and artificial neural networks. Artificial neural networks include cells or processors, and they are called nodes. These nodes communicate with each other to form a network of nodes, and these connections have a set of values called weights. The role of weights in this

network is to determine the resulting values of each element to be processed based on the values entered for that element.

Mathematical neural networks are simulating the way the brain works with mathematical attempts and software solutions [24]. When the way in which the brain and brain cells learn and its ability to remember, learn, distinguish between things and make decisions was reached, the idea of creating a simulation process for this neural network was established. The interconnection between biological neurons through the presence of billions of interconnected cells that are intertwined in a complex manner with nerve appendages, allows and facilitates the process of storing information, images, sounds and signals transmitted by the five senses, and this facilitates the learning process by repetition and error [25].

Artificial intelligence is produced through the use of mathematical models to simulate the work of biological networks, and one of the types of artificial intelligence is neural networks. Through many examples, it was found that there is a neural network that was formed from neurons to train, develop, grow, learn and make the right decisions.

### 3.2.1. NN Basics

Each neural network can be conceptualized as a hierarchical structure composed of layers of artificial cells concealed between the input layer and the outer layer or the inner layer and the outer layer. There is complete communication in both directions between all neurons in the present layer and those in the layers below and above.

An early discovery was that neurons execute signal collection; for example, if a neuron has two connections and each connection transmits a signal, the result is the result of the signals in the normal collection [26]. If a cell gets two signals, for instance, each would be multiplied by the neuron's coefficient before being added together. It was constructed with the understanding that the engineering techniques used to generate artificial cells cannot duplicate the complexity of genuine neurons.

Each synapse between neurons has a unique value, known as a weighting, that represents the significance of the connection. This is a quick description of the operations neural networks execute. After integrating all of the inputs and applying a transformation function that changes depending on the kind of neuron, the result is taken as the neuron's output and transmitted to the neurons in the next layer. [27].

### 3.2.2. NN Layers

A neural network consists of 3 types of layers. The first layer is the input layer which is the raw data of the neural network. Te second is hidden Layers, this layer is the middle layer between the input and output layer and the place where all calculations are performed [28]. the last layers is output Layer that produces the result for a given input as showed in Figure 3.1.

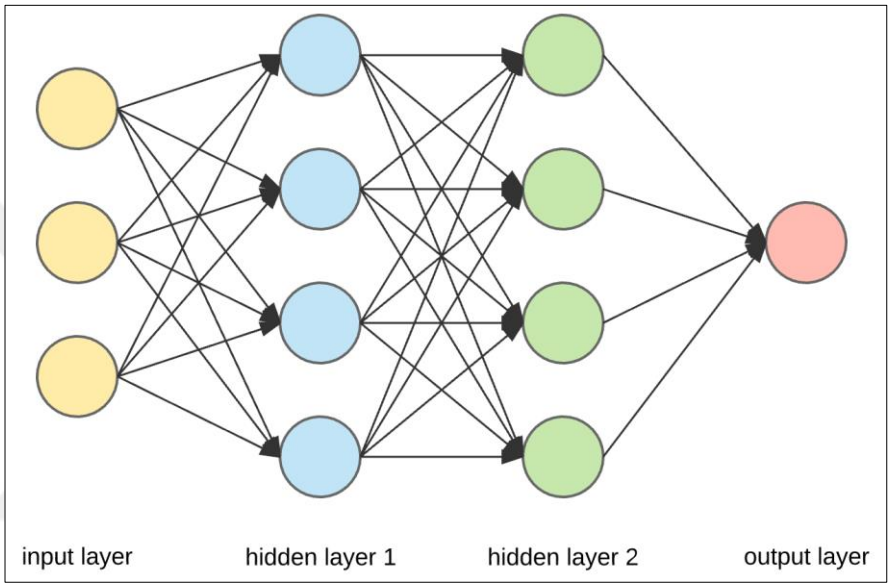


**Figure 3.1:** NN layers

The three yellow circles in the Figure 1 represents the input layer and is usually denoted as the X vector. There are four blue circles and four green circles representing the hidden layers. These circles represent "activation" nodes and are usually denoted as W or θ [29]. The red circle is the output layer or the expected value. The expected value here means values in the case of multiple output classes/types.

Each node is connected to each node of the next layer and each connection has a certain weight. The weight can be thought of as the effect of that node on the node of the next layer as shown in Figure 3.2.

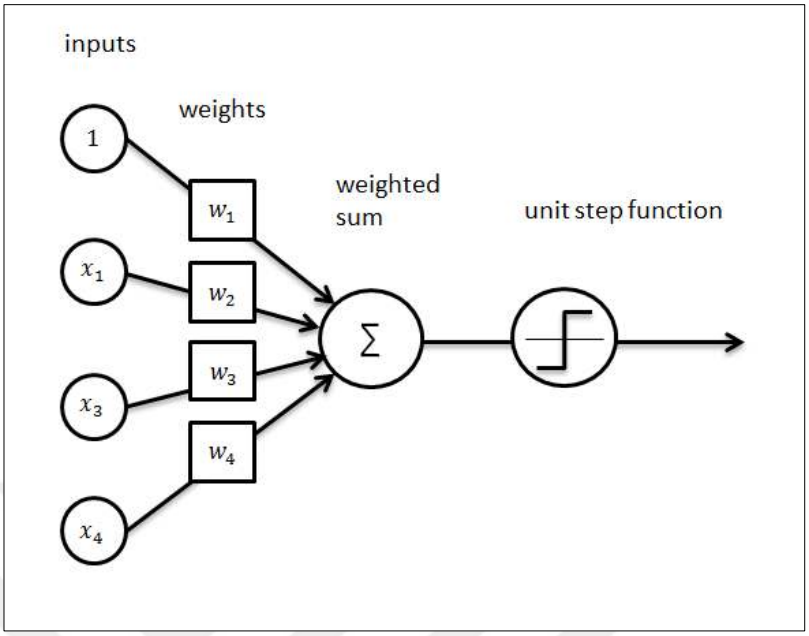


**Figure 3.2:** Nodes of NN

### 3.2.3. NN Main Types

There are many neural networks that each have specific functions and capabilities that distinguish them in order to perform the tasks required of them. Networks differ from each other in general in the number of layers, data direction and speed, and we will review the most important and most used neural networks at the moment, Feed Forward, convolutional, recurrent, Radial Basis Function, Self-Organizing Map and Modular neural network [30].

#### 3.2.3.1. Feed forward neural networks

Feed-forward networks are one of the simplest neural networks and the least in size, but it is characterized by its simplicity and speed, but because of this its accuracy is not the best among other neural networks, and it is called forward because the data goes in one direction, and the backpropagation algorithm is not used in it [31].

It may contain one hidden layer or it may not contain any hidden layer, and we use in this network the sigmoid function, and it is used to identify and distinguish sounds specifically, in computer vision, and self-driven cars and may be used in simple classification.

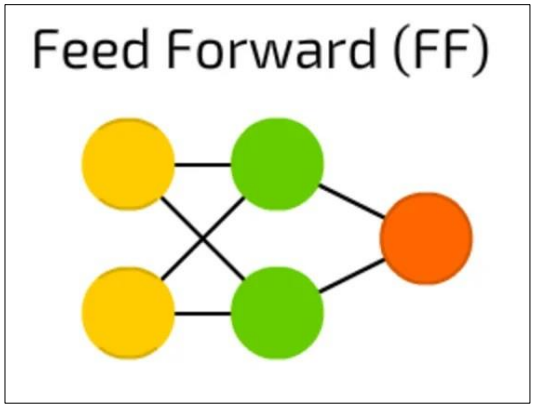


**Figure 3.3:** Feed Forward Neural Networks

It is worth noting that there is an update of this neural network to what is known as the Multi Layers Feed Forward Neural Networks (MLFFNN) [32] as showed in Figure 3.3.

### 3.2.3.2. Convolutional neural network (cnn)

Convolutional neural networks are the most widely used neural networks, which we certainly use at least one of its applications every day in our daily lives, and they are very famous for their use in image and video recognition [33], but nevertheless they are used in many suggestion algorithms, such as the Netflix algorithm as showed in Figure 3.4.

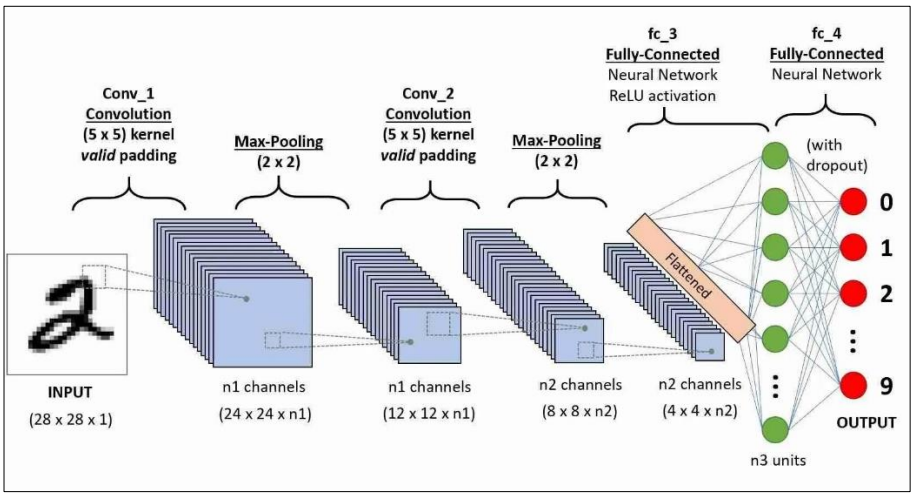


**Figure 3.4:** Convolutional Neural Network

These neural networks are known for being less parameter dependent than other networks, but on the other hand, they are slow and complex and difficult to design and modify.

Despite the complexity of this algorithm, but simply that this algorithm is used in image and video analysis, in face recognition, computer vision, voice recognition, and in a large number of medical uses [34].

### 3.2.3.3.Recurrent neural networks (rnn)

These recursive networks are one of the most used neural networks, and they are the most used in the process of prediction and reinforcement learning [35], if they rely on input feedback or ratings for you to deal with, and are used in applications such as Facebook friend suggestions, completion of writing, translation, or voice-to-text translation [36]. The Figure 3.5 shows the principle of RNN.

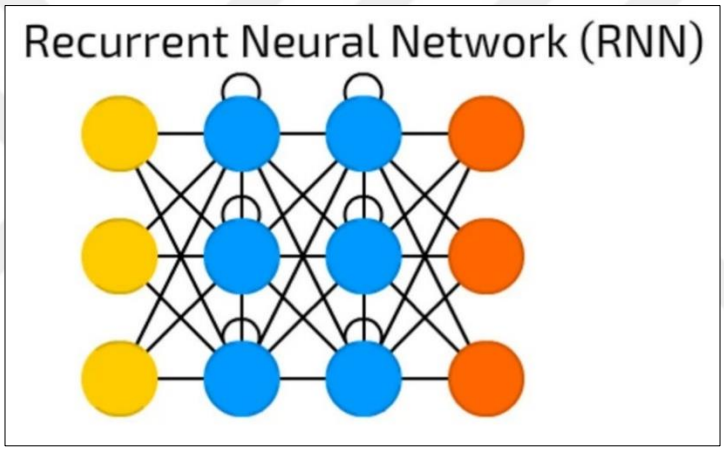


**Figure 3.5:** Recurrent Neural Networks

### 3.2.3.4.Radial basis function neural networks

These networks are very similar to feed-forward networks, but instead of a sigmoid function they use a radial basis function (RBF), which is based on the distance between the center of the circle and points, and is used for classification, chronological functions, and machine control [37].

### 3.2.3.5.Self-organizing map neural network

These networks are also called Cohonen networks after the Finnish scientist who invented them, and these networks receive a lot of input, and then divide and distribute them as Clustering, and therefore they are used in unsupervised learning, and it takes place in three steps: construction, training and introduction [38].

To explain the idea and if we have the characteristics of a million people and we want to number them, group them, or cluster them, using this neural network; it does all the work and produces the appropriate output [39]. This network is used in navigation, mapping, water and petroleum exploration, and big data analysis.

#### 3.2.3.6.Modular neural network (MNN)

This network has a very innovative idea which is to combine more than one unit of neural networks or combine a number of neural networks, which work in parallel with each other and eventually intertwine in order to aggregate the processed values and from which a single output is issued.

These parallel networks may be different, one handles text, one handles audio, and one handles video [40]. These networks are very fast compared to the data they process, are used in financial and economic analyses, and have many uses in biology [41].

## 3.3. ARTIFICIAL NEURAL NETWORK ANN

A neural network can be defined as a group of algorithms that are designed to be similar in their work to the work of the human brain, trying to imitate techniques and form relationships between different data. There are many uses for these networks [42][43][44], including regression, classification, face recognition, and many other applications that fall within the field of neural networks. Since the neural network is trying to imitate the mechanics of the human brain, there are some differences and similarities.

By making a comparison between them, the biological neural network shows parallel processing, while the artificial neural network performs and executes the processing in a sequential manner, and the biological neural network is the slowest of the artificial network, the first being in milliseconds, and the second being in nanoseconds. [45]. There is a problem with this process's lack of independence in solving problems, even though it results in high speeds, because the electronic computer sends a large number of data in digital form between each of its parts, but the brain works in some parts in a digital form and in other parts in an analog way, and the brain does not You send redundant data while your computer lacks this feature[46].

### 3.4. INTERNET OF THINGS

IoT is a network of interconnected computing devices that allows for the automatic collection and wireless transmission of data.

Almost device with an on/off switch can be part of the Internet of Things, not just personal computers and mobile phones. Something connected to the Internet may be anything from a person with a pacemaker to a camera that broadcasts live footage of marine life in coastal waters to a car with sensors that detect danger and sound an alarm. Everything that can be assigned an IP address and communicates data over a network can be categorized as Internet of Things devices.

i. The sensors and gadgets that make up an IoT system are linked to a cloud service over some sort of network [47]. Data is sent to the cloud, where software examines it to see if any changes need to be made (to the sensors or devices) or whether an alert should be sounded.

ii. There are four essential parts to any completely functional IoT system. Sensors and devices, as well as communication and data processing. Now, let's have a look at each part in detail:

iii. It is equipped with sensors and other technologies to learn more about its environment. The device may be equipped with a wide variety of sensors, such as the global positioning system (GPS), camera, accelerometer, and more found in a modern smartphone. The sensor or sensors, in essence, collect data from the surrounding environment for some predetermined goal.

iv. When the data collection is complete, the device must send it to the cloud service. ii. Wi-Fi, Bluetooth, satellite, Low Energy Wide Area Networks (LPWAN), and even an Ethernet connection to the Internet are all viable options for this purpose. Which communication protocol is used depends on the IoT application being used.

v. After data has been uploaded to a cloud service, the software can process it and decide what to do with it. Notifying people or making instantaneous, automatic adjustments to sensors or the device are two examples. The user interface, however, is helpful for those times when human input is required.

vi. User Interface (IV): The user interface allows the user to provide input as necessary or when testing the system. Any user-initiated actions made in the opposite direction are relayed by the system. A user must navigate between the interface and the cloud service and then return to the sensors or devices in order to make the required change.

IoT applications determine the networking, connectivity, and communication protocols used by various web-based devices. The internet of things (IoT) increasingly makes use of AI and machine learning to streamline and accelerate data collection.

### 3.4.1. IOT Technologies

There are five technologies used for the Internet of Things [m9], radio frequency identification (RFID), wireless sensor networks (WSN), middleware, cloud computing, IoT application software.

#### 3.4.1.1. Radio frequency identification (RFID)

RFID is a technology that uses radio waves to exchange information between devices. RFID uses RFID readers and RFID antennas to communicate.RFID tags are small chips that store data about the object to which they are attached. It can be read by an RFID reader, which uses radio waves to collect data from the tag. This data is then sent to an RFID antenna, which transmits it to a computer or other device [49].

In the context of the Internet of Things, RFID can be used to collect data about things around us. For example, an RFID tag attached to a product can be used to track its location as it moves through the supply chain. Or, an RFID tag can be used on a piece of equipment to keep track of its use and maintenance history. RFID can also be used to initiate events. RFID is a key technology in the Internet of Things [50] because it allows devices to communicate with each other without any human interaction. This makes it ideal for data collection and event triggering in intelligent systems.

RFID and IoT can work together to create smarter systems. For example, an RFID tag on a product can be used to trigger a notification when it is being scanned by an RFID reader. This notification can be sent to a smartphone or other device so that the person carrying the product can be alerted [51].

RFID tags can also be used to track the location of objects in real time. This information can be used to create a map of things around us. For example, an RFID tag on a package can be used to track its location

as it moves through the delivery system. This information can be used to improve the delivery process and reduce costs.

RFID and the Internet of Things can also be used together to create smarter buildings. For example, an RFID tag could be used on a person to unlock a door when they approach it. Or an RFID tag can be used on a piece of equipment to monitor its efficiency. This information can be used to improve device performance and reduce downtime.

RFID technology and the Internet of Things can work together to create smarter cities. For example, an RFID tag on a vehicle can be used to initiate a garage door opening as the vehicle approaches. Or, an RFID tag can be used on a person to turn on the sidewalk lighting as they walk on it.

The possibilities of RFID and the Internet of Things are endless. By working together, these technologies can create smarter systems that make our lives easier and more efficient

#### 3.4.1.2.Wireless sensor networks (WSN);

In recent years, research into sensor networks has gained prominence [52] due to the immense opportunity they offer for developing applications that bridge the gap between the actual and virtual worlds. Large-scale sensor networks will make it feasible to collect data on physical phenomena that were previously unreachable via more conventional means.

WSN development is frequently recognized as one of the greatest technological achievements of the 21st century. Small, inexpensive sensors and the Internet enable contemporary breakthroughs in microelectromechanical systems (MEMS) and wireless communication technologies, which in turn provide previously inconceivable opportunities. For usage in a variety of civilian and military applications, including but not limited to environmental monitoring, battlefield monitoring, and industrial process monitoring.

Wireless Sensor Networks (WSNs) use self-configured wireless networks and infrastructure to transmit data to the main network, where it is analyzed and interpreted by humans [53] in order to monitor physical or environmental conditions such as temperature, sound, vibration, pressure, movement, or pollutants.

#### 3.4.1.3. Middleware

IoT middleware is an application and software that acts as an intermediary between the components of the Internet of Things.

The job of middleware is to connect different programs that are not designed to be a connected program. One of the main advantages of the Internet of Things is the ability to connect anything and the possibility to transfer data through networks. Middleware is an infrastructure element that facilitates the communication process for a large number of Internet of Things devices by providing a communication layer for sensors, in addition to providing communication to application layers that provide services through which effective communication between programs is carried out.

#### 3.4.1.4. Cloud computing

Cloud Computing, which is software and algorithms that are implemented through cloud processors to store, display and analyze data that has been sent through various devices supported by Internet of Things technology with the aim of identifying some statistics or predicting the performance of various operations through artificial intelligence techniques and analysis Big data to help decision makers reach to make sound decisions based on previously collected and analyzed data, which is called DDDM or Data Driven Decision making [54].

#### 3.4.1.5. IoT application software

Internet of Things applications are applications that provide the possibility of developing a large number of Internet of Things applications that are intended for users. When there is a physical connection between devices and networks, these applications provide an interactive connection between something and something and between people and something in the Internet of things, with reliability and strength at work. Ensuring that data is received and dealt with correctly and at appropriate times depends on providing these applications on Internet of Things devices. For example, when we design the Internet of Things to monitor the transportation of food products, these applications monitor the conditions that must be available to complete the transportation process in a healthy and correct manner, such as the availability of temperatures and humidity in order to preserve those products from damage, and the necessary measures are taken automatically to make this process sound . IoT applications provide the

necessary information to end users in a simple way that allows them to interact with that environment and provides them with the necessary information. The use of intelligence in designing Internet of Things applications is one of the most important factors that give the necessary strength to these applications to be able to make decisions quickly and without the interference of users in those decisions.

### 3.4.2. IoT Devices

Every day brings further proof of the astonishing reality that there are more connected devices than people on Earth. Due to the expansion of Internet networks to 5G, there are presently more than 20 billion smart devices for the Internet of Things, despite the fact that there are only approximately 7.62 billion people on Earth. As demonstrated by the graph below, the quantity of IoT devices has increased steadily. [55].

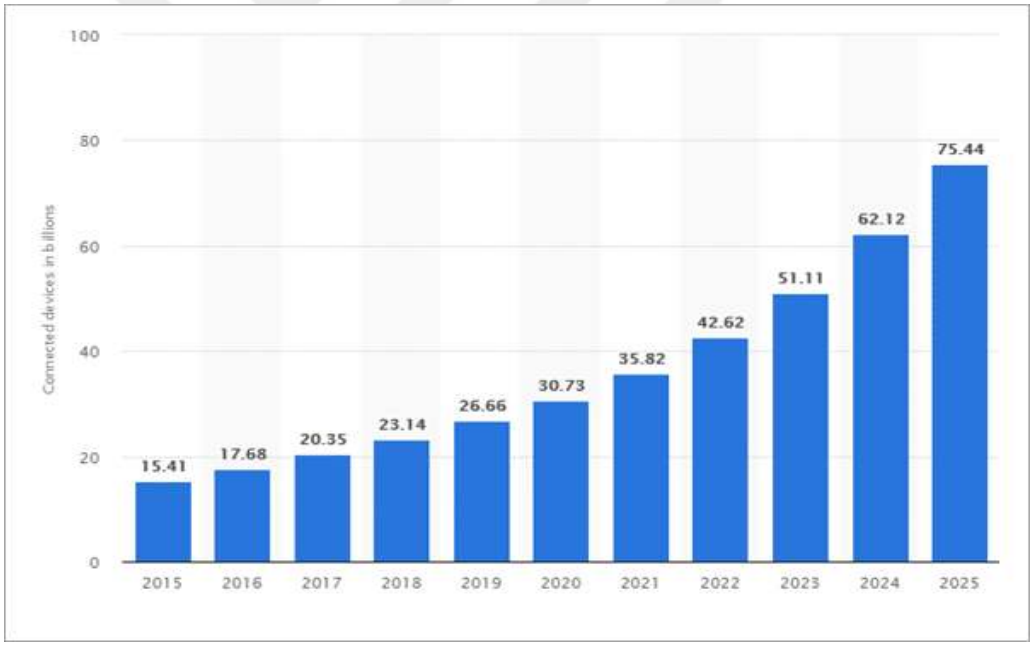


**Figure 3.6:** IoT devices statistics [57]

The production of Internet of Things devices is constantly increasing and dramatically. These products provide the necessary services in many areas of human life, and among the most important products of this technology are portable computers, smart phones, tools, watches and smart vehicles. We also see the use of these devices daily to facilitate daily life.

There are factors that make the ordinary device turn into a smart device for the Internet of Things.

First, the device's ability to connect to the Internet, and secondly, the provision of devices that carry sensor technology, functional applications, and technologies that provide network connections and drivers. The Internet of Things device is configured by merging these two features together. For example, watches were used to know the time and date in a simplified manner, but nowadays, smart watches provide multiple information, including human heartbeat, calculating the number of calories, and many other information that would not have been available without the presence of Internet of Things and smart devices.

IoT devices have a simple life cycle that begins with deployment, through monitoring, service, and management, continues with updates, and ends with shutdown at the end.

As we mentioned here, IoT devices have many benefits, including providing interaction between devices and other devices, providing the possibility of automation and quality control, which provides good control, in addition to saving money and time by reducing manual work, while increasing the efficiency of work in our daily life that makes it better [56].

Despite these advantages and benefits, these devices also have disadvantages. The lack of international compatibility of these devices, and the possibility of failure of the services provided by these devices due to the complexities that go into their formation, and that the emergence of these devices threatens the safety and privacy of individuals, as it may threaten the number of jobs to decrease to dispense with manual tasks due to the availability of automation in them. We see in the future that these devices began to control life in light of the rise in the pace of artificial intelligence techniques [58].

### 3.5. METHODOLOGY OF PROPOSED SYSTEM

In order to build the system, we relied on the basic steps in creating and analyzing data sets before implementing the system. In our proposed method, we analyze and classify the data of IoT devices and extract features from them. We perform all the steps by building a custom network from the ANN using special parameters to perform the functions in the most accurate form. First, real data sets for a group of devices are fetched and entered into the system. We first read the data and define the columns and their contents. After reviewing the contents of the data, we conduct a set of pre-processing on it to deliver it to the most appropriate form required. Data state sham charts are made in the preprocessing steps. Next,

we build an artificial neural network model and divide the ready data into one part to train the system and another to test it. Two models are built in order to evaluate our proposed method of calculating accuracy at the end of the process. Then it can easily perform classification and extract features for each device according to its data. Figure 3.7 represents the steps for implementing the model.

As shown in Figure 3.7, the data pre-processing steps consist of several stages. Two different models for evaluating performance are built using different parameters and with more than one activation function to conduct tests in several ways. A high performance accuracy was calculated compared to all relevant research.

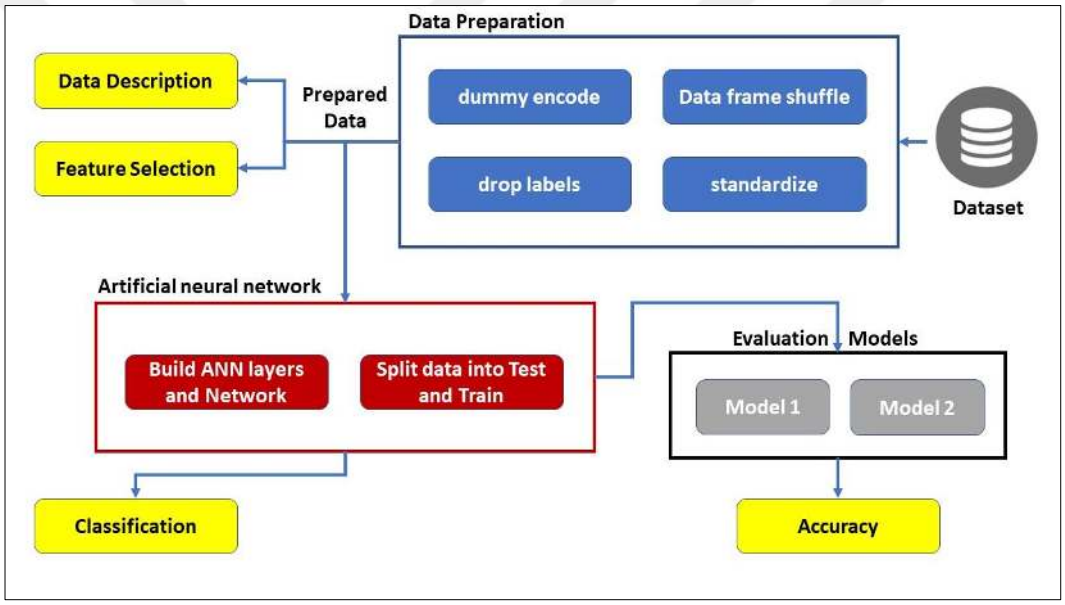


**Figure 3.7:** Implementing the proposed model

The importance of our proposed method can be seen at the end of the model after obtaining classified and usable data sets. A comparison is made of the accuracy of the system's work by applying it to all devices' data to provide a more effective and applicable tool later.

### 3.5.1. Dataset

It is based on a real data set for several devices connected to the Internet [59]. The dataset used in our research is public botnet datasets, particularly for the Internet of Things. It contains real traffic data; this data was collected based on 9 commercial IoT devices.

The dataset is Multivariate and Sequential, it contains 7062606 instances which obtained by computer attribute area. The dataset has 115 attributes collected during march 2018. The class distribution is shown in figure 3.8.

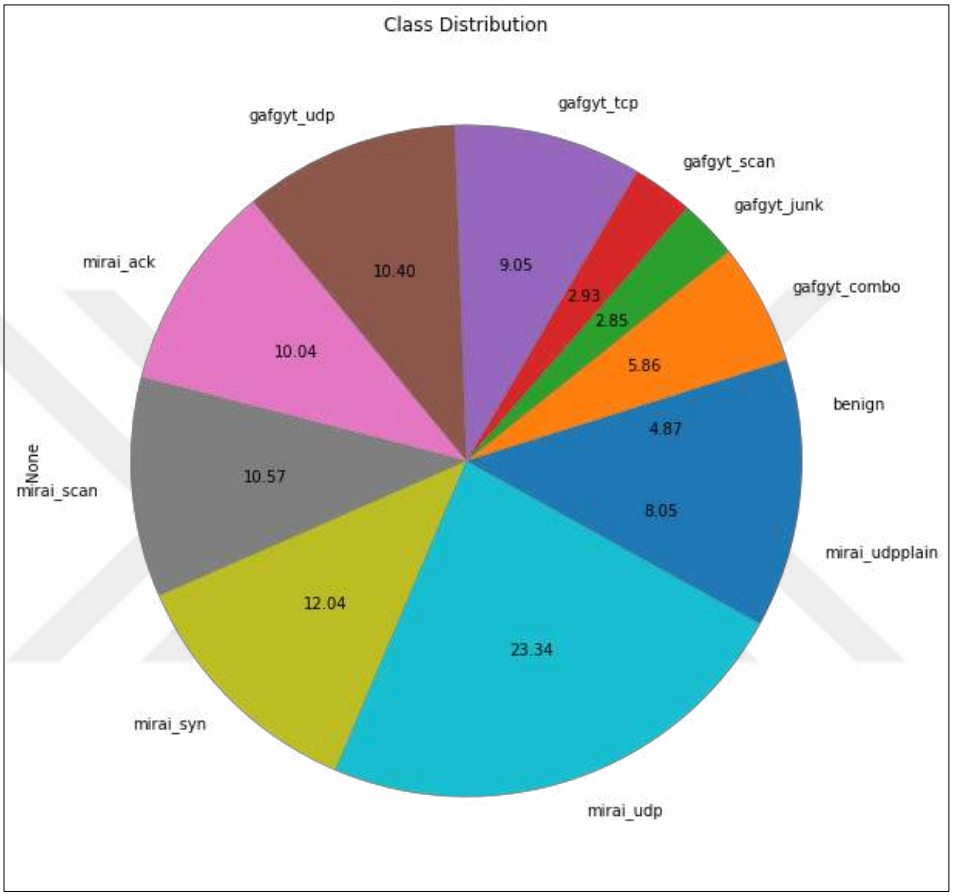


**Figure 3.8:** Dataset class distribution for sample device

### 3.5.2. Data Preparation

Data preparation is the process of preparing unprocessed information for analysis. Exploring and visualizing unstructured data is crucial, as is transforming it into a format that machine learning (ML) algorithms can comprehend.

In our process we are performing data preprocessing method with four steps. In the first step we are performing data frame shuffle. using the sample method. The sampling method allows its user to take a set of rows in a Pandas Data frame at random, not bound [60]. The importance of using it as a

preprocessing step is to facilitate the processing of big data simply and to frame the Pandas data completely and in random order. Dealing with more and more big data is relatively difficult and we have to randomly switch the original data set to reduce variance and ensure that the model will be well applied to the invisible data points [61]. The Figure 3.9 shows our data frame shuffle for the selected data.

```
#shuffle rows of data frame
sampler=np.random.permutation(len(data))
data=data.take(sampler)
data.head()
```

| | MI_dir_L5_weight | MI_dir_L5_mean | MI_dir_L5_variance | MI_dir_L3_weight | MI_dir_L3_mean | MI_dir_L3_variance | MI_dir_L1_weight | MI_dir_L1_mean | MI_d |
|---|---|---|---|---|---|---|---|---|---|
| **167885** | 56.448360 | 429.367371 | 45982.027637 | 97.204087 | 402.478772 | 51843.885830 | 311.535859 | 336.608335 | |
| **95048** | 1.000000 | 60.000000 | 0.000000 | 1.000000 | 60.000000 | 0.000000 | 1.000000 | 60.000000 | |
| **18811** | 57.299666 | 528.997745 | 11724.558590 | 92.969650 | 474.003878 | 33113.978503 | 333.352345 | 373.579146 | |
| **124285** | 70.903557 | 60.308473 | 6.573243 | 96.787338 | 60.267831 | 5.428306 | 236.711095 | 60.160333 | |
| **111507** | 31.161761 | 444.979477 | 46228.011874 | 69.717453 | 383.818471 | 58820.229681 | 284.719788 | 343.335534 | |

**Figure 3.9:** Data frame shuffle

In the second step we perform dummy encode labels. Dummy notation is used when features of related categorical variables are of importance in classification [62]. The optimal dummy notation method provides one way to use categorical predictor variables in different types of estimation models such as linear regression. In Figure 3.10 we perform dummy coding, the main reason to perform it as preparation step since it is a way to make the categorical variable into a series of dichotomous variables.

```
#dummy encode labels
labels_full=pd.get_dummies(data['type'], prefix='type')
labels_full.head()
```

| | type_benign | type_gafgyt_combo | type_gafgyt_junk | type_gafgyt_scan | type_gafgyt_tcp | type_gafgyt_udp | type_mirai_ack | type_mirai_scan | type_mirai_sy |
|---|---|---|---|---|---|---|---|---|---|
| **167885** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **95048** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | |
| **18811** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **124285** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| **111507** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | |

**Figure 3.10:** Dummy coding

In machine learning, data labeling is the process of identifying raw data by adding one or more useful and informative labels that help provide a context from which machine learning can learn. We perform third step of preparing the dataset as dropping labels from the part of training dataset. this step is performed to avoiding the whole data frame with target values and the function will return train data

frame and test data frame. The final step is represented by standardizing the numerical columns as shown in Figure 3.11.

```
#standardize numerical columns
def standardize(df,col):
    df[col]= (df[col]-df[col].mean())/df[col].std()

data_st=data.copy()
for i in (data_st.iloc[:,:-1].columns):
    standardize (data_st,i)

data_st.head()
```

| | MI_dir_L5_weight | MI_dir_L5_mean | MI_dir_L5_variance | MI_dir_L3_weight | MI_dir_L3_mean | MI_dir_L3_variance | MI_dir_L1_weight | MI_dir_L1_mean | MI_d |
|---|---|---|---|---|---|---|---|---|---|
| **167885** | -0.292727 | 2.218295 | 1.364498 | -0.189537 | 2.108314 | 1.449856 | -0.012457 | 1.656399 | |
| **95048** | -1.149135 | -0.614769 | -0.602271 | -1.135965 | -0.644789 | -0.624700 | -1.118579 | -0.679704 | |
| **18811** | -0.279578 | 2.982465 | -0.100782 | -0.231194 | 2.683287 | 0.700370 | 0.065253 | 1.968637 | |
| **124285** | -0.069464 | -0.612403 | -0.601990 | -0.193637 | -0.642636 | -0.624483 | -0.278981 | -0.678350 | |
| **111507** | -0.683282 | 2.338041 | 1.375019 | -0.459942 | 1.958308 | 1.729017 | -0.107975 | 1.713214 | |

**Figure 3.11:** Standardizing the numerical columns

When performing classification and feature extraction operations, it is important to standardize the variables. When the regression model contains polynomial terms or interaction terms, the standardize model should be applied. The important information and the relationship between the classification variables produces a lot of excessive linear relationships in the feature extraction operations, which increases the accuracy of the following operations. The benefits of standardizing dataset-specific variables are mentioned in [63], which should be followed with machine learning systems for classification and feature extraction. If the procedures are specific to the regression model and they contain polynomial terms and interaction conditions, then the unification of the variables is the most important steps for preparing the data, especially with the data of Internet of Things devices.

### 3.5.3. Classification

For classification we build and train a custom artificial neural network. The basis of the work of the artificial neural network is through the communication between nodes in all layers. The network is built using subtle hidden layers for the purpose of training data and developing a model that performs classification with high accuracy. A softmax function is used in building the model in order to optimize the process with each data training. Adam is used as a parameter to optimize within the neural network model.

### 3.5.4. Feature selection

The process of picking the most significant features from a given dataset is known as feature selection. In many circumstances, Feature Selection may also improve the performance of a machine learning model.

When working with a dataset with a large number of features, the significance of feature selection becomes clear. This sort of dataset is also known as a high dimensional dataset. With this high dimensionality comes a slew of issues, including: dramatically increasing the training duration of your machine learning model; making your model too sophisticated, which may lead to Overfitting.

Several elements in a high-dimensional feature set are frequently redundant, which means they are nothing more than extensions of the other key features. These redundant features do not serve to model training as effectively. So, clearly, there is a need to extract the most critical and relevant characteristics from a dataset in order to get the best predictive modeling performance.

Below is a table containing a set of features that the data set contains in Table 3.1.

**Table 3.1:** Dataset features sample

| Feature Name | Feature Description |
|---|---|
| MI_dir_L5_weight | MI dir Lambda 5 weight |
| MI_dir_L0.01_variance | MI dir Lambda 0.01 variance |
| H_L0.1_weight | traffic from this packet's host (IP) Lambda 0.1 weight |
| HH_L5_weight | traffic going from this packet's host (IP) to the packet's destination host Lambda 5 weight |
| HH_L5_pcc | traffic going from this packet's host (IP) to the packet's destination host Lambda 5 pcc |
| HH_L1_radius | traffic going from this packet's host (IP) to the packet'sdestination host Lambda 1 radius |
| HH_L0.01_radius | traffic going from this packet's host (IP) to the packet's destination host Lambda 0.01 radius |
| HH_L0.01_covariance | traffic going from this packet's host (IP) to the packet's destination host Lambda 0.01 covariance |
| HH_jit_L3_mean | Jitter of the traffic going from this packet's host Lambda 3 mean |
| HH_jit_L0.01_variance | Jitter of the traffic going from this packet's host Lambda 0.01 variance |
| HpHp_L0.01_pcc | traffic going from this packet's host+port (IP) to the packet's destination host+port Lambda 0.01 pcc |

### 3.5.5. Data input and proposed output

Our proposed system improves the performance of previous systems in this field that have been studied. Many systems have studied the data of devices in the scope of the Internet of Things for the purpose of analyzing, filtering and removing unwanted values, after which a classification is made for them. It can be said that the input is always the unclassified data received from IoT devices. And the outputs will be as extracting features from data, data classification, and providing the percentage of accuracy in the system based on more than one accuracy measurement model.

### 3.5.6. Evaluation models

Using evaluation metrics, the quality of a statistical or machine learning model can be measured. In every undertaking, it is vital to evaluate machine learning models and algorithms. There are numerous types of tests that can be performed to assess a model.

Our suggested technique requires the creation and evaluation of not one but two distinct rating models. Each model employs numerous factors in order to select and evaluate its accuracy. In this paragraph, we explain why our proposed method is preferable to the other evaluation methodologies previously examined for the accuracy measure Table 3.2

**Table 3.2:** Comparison with other literatures

| # | Used dataset | Used method | Accuracy |
|---|---|---|---|
| [17] | own | Multilayer perceptron (MLP), ANN and backpropagation algorithm (BP) | 85% |
| [18] | own | statistical algorithms | 92.76% |
| [19] | 1-Kaggle data<br>2- Vx-Heaven | fuzzy pattern tree | 88.76%<br>90.09% |
| [20] | NSL-KDD dataset | Semi-supervised Fuzzy | 86.53% |
| [21] | Own collected 104528 malicious files | hashing techniques | 91.00% |
| Our proposed method | Kaggle data for 9 devices | Custom ANN | Minimum_accuracy 92.84%<br>Maximum_accuracy 94.45% |

The results show that our proposed method has become more accurate and more efficient on all cases in the tests. The system was tested on the data of the nine devices within the dataset. Accuracy was recorded

using two models for each case, and the lowest and highest accuracy were compared with the results of previous research. The results of the evaluation of the other cases ranged between the highest and lowest value.

# 4. RESULTS AND CONCLUSIONS

## 4.1. EXPERIMENTS AND RESULTS

In this step, tests were conducted for the purpose of evaluating the performance of the system with all devices within the dataset. A dataset consists of a set of data for each device. Our proposed system was tested with nine cases in each case the accuracy measured for classifying single device data and extracting features in our artificial neural network.

i. Case1: In this case the system was tested using Danmini_Doorbell data. The information was read and the accuracy of the system was measured. The average accuracy rate for case 1 was 92.76% where the average response time is 2.19 ms.

ii. Case2: In this case the system was tested using Ecobee_Thermostat data. The information was read and the accuracy of the system was measured. The average accuracy rate for case 2 was 94.01% where the average response time is 3.55 ms.

iii. Case3: In this case the system was tested using Ennio_Doorbell data. The information was read and the accuracy of the system was measured.The average accuracy rate for case 3 was 93.30% where the average response time is 3.10 ms.

iv. Case4: In this case the system was tested using Philips_B120N10_Baby_Monitor data . The information was read and the accuracy of the system was measured. The average accuracy rate for case 4 was 93.09% where the average response time is 2.37 ms.

v. Case5: In this case the system was tested using Provision_PT_737E_Security_Camera data. The information was read and the accuracy of the system was measured.The average accuracy rate for case 5 was 94.22% where the average response time is 3.01 ms.

vi. Case6: In this case the system was tested using Provision_PT_838_Security_Camera data. The information was read and the accuracy of the system was measured. The average accuracy rate for case 6 was 93.87% where the average response time is 1.88 ms.

vii. Case7: In this case the system was tested using Samsung_SNH_1011_N_Webcam data. The information was read and the accuracy of the system was measured.The average accuracy rate for case 7 was 94.08% where the average response time is 2.00 ms.

viii. Case8: In this case the system was tested using Simple Home_ XCS7_ 1002_ WHT_ Security_ Camera data. The information was read and the accuracy of the system was measured.The average accuracy rate for case 8 was 94.45% where the average response time is 2.12 ms.

ix. Case9: In this case the system was tested using Simple Home_XCS7_ 1003_WHT_ Security_ Camera data. The information was read and the accuracy of the system was measured.The average accuracy rate for case 9 was 94.15% where the average response time is 2.38 ms.

After reviewing the accuracy ratios in our proposed model in the above experiments, it was found that our model classified the data well with an average execution time of 2.5ms. Table 4.1 shows a summary of the experiments of our proposed system.

**Table 4.1:** Experiments summary

| # | IoT device | Accuracy | response time |
|---|---|---|---|
| Case 1 | Danmini_Doorbell | 92.76% | 2.19 ms |
| Case 2 | Ecobee_Thermostat | 94.01% | 3.55 ms |
| Case 3 | Ennio_Doorbell | 93.30% | 3.10 ms |
| Case 4 | Philips_B120N10_Baby_Monitor | 93.09% | 2.37 ms |
| Case 5 | Provision_PT_737E_Security_Camera | 94.22% | 3.01 ms |
| Case 6 | Provision_PT_838_Security_Camera | 93.87% | 1.88 ms |
| Case 7 | Samsung_SNH_1011_N_Webcam | 94.08% | 2.00 ms |
| Case 8 | SimpleHome_XCS7_Security_Camera | 94.45% | 2.12 ms |
| Case 9 | SimpleHome_XCS7_1_Security_Camera | 94.15% | 2.38 ms |

## 4.2. CONCLUSIONS

In this proposed method, we study the data of the Internet of Things and the devices that created the data. A framework and model for building a private artificial neural network training mechanism were developed. The model is used to classify data using a neural network and extract and study features. Our proposed method can be used to build high-resolution models in which data that cannot be used without classification are initialized. Real data collected and published is used for nine home-

use devices. Given the existence of previously used methods in this field, our results were compared with the results reported in the previous literature. Our method showed that it performs well, with high accuracy, and an average execution time of 2.5 milliseconds.

Our proposed method can be used to classify IoT data and associated devices. The importance of using classified data compared to unclassified data in building models of smart and efficient systems is not hidden. The main objective is to provide the possibility of classifying and extracting features with a relatively high accuracy compared to the previous methods.

### 4.3. FUTURE WORK

In the future we can apply the proposed model to data collected specifically for its application and extract features from it. Data collection tools such as cloud computing and distributed devices can be used to make the system work on our private data. The results can be compared with other literature that operates in different ways and according to the data used in it.

.

# REFERENCES


[1] Z. Bu, B. Zhou, P. Cheng, K. Zhang, and Z.-H. Ling, "Encrypted Network Traffic Classification Using Deep and Parallel Network-in-Network Models," IEEE Access, vol.8, pp.132950-132959, 2020.

[2] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, and A. Zanella, "IoT: Internet of Threats? A Survey of Practical Security Vulnerabilities in Real IoT Devices," IEEE Internet of Things Journal, vol.6, no.5, pp.8182-8201,2019.

[3] R. Song and T. Willink, "Machine Learning-Based Traffic Classification of Wireless Traffic," 2019 International Conference on Military Communications and Information Systems (ICMCIS), 2019.

[4] A. Bertello, A. Ferraris, S. Bresciani, and P. De Bernardi, "Big data analytics (BDA) and degree of internationalization: the interplay between governance of BDA infrastructure and BDA capabilities," Journal of Management and Governance, vol.25, no.4, pp.1035-1055, 2020.

[5] N. Rajak and R. Rajak, "Performance Metrics for Comparison of Heuristics Task Scheduling Algorithms in Cloud Computing Platform," Machine Learning Approach for Cloud Data Analytics in IoT, pp.195-226, 2021.

[6] A. J. Pinheiro, J. de M. Bezerra, C. A. P. Burgardt, and D. R. Campelo, "Identifying IoT devices and events based on packet length from encrypted traffic," Computer Communications, vol.144, pp.8-17, 2019.

[7] H. Tahaei, F. Afifi, A. Asemi, F. Zaki, and N. B. Anuar, "The rise of traffic classification in IoT networks: A survey," Journal of Network and Computer Applications, vol.154,pp. 102538, 2020.

[8] A. Azab, M. Khasawneh, S. Alrabaee, K.-K. Raymond Choo, and M. Sarsour, "Network traffic classification: Techniques, datasets, and challenges," Digital Communications and Networks, 2022.

[9] S. Izadi, M. Ahmadi, and R. Nikbazm, “Network traffic classification using convolutional neural network and ant-lion optimization,” Computers and Electrical Engineering, vol.101, p.108024, 2022.

[10] S. Roy, T. Shapira, and Y. Shavitt, “Fast and lean encrypted Internet traffic classification,” Computer Communications, vol.186, pp.166-173, 2022.

[11] Yin, J., Guo, L., Jiang, W., Hong, S., & Yang, J. “ShuffleNet-inspired lightweight neural network design for automatic modulation classification methods in ubiquitous IoT cyber–physical systems“. Computer Communications, vol.176, pp.249-257, 2021.

[12] Yadav, R., Sreedevi, I., & Gupta, D. “Augmentation in performance and security of WSNs for IoT applications using feature selection and classification techniques“. Alexandria Engineering Journal, 2022.

[13] Capris, T., Takagi, Y., Figueiredo, D., Henriques, J., & Pires, I. M. “A Convolutional Neural Network-enabled IoT framework to verify COVID-19 hygiene conditions and authorize access to facilities“. Procedia Computer Science,vol.203, pp.727-732, 2022.

[14] Yamada, Y., Shinkuma, R., Iwai, T., Onishi, T., Nobukiyo, T., & Satoda, K. (2018). “Temporal traffic smoothing for IoT traffic in mobile networks“. Computer Networks,vol. 146, pp.115-124, 2018.

[15] Al-Turjman, F. “Information-centric framework for the Internet of Things (IoT): Traffic modeling & optimization“. Future Generation Computer Systems,vol.80, pp.63-75,2018.

[16] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, “Neural network architecture based on gradient boosting for IoT traffic prediction,” Future Generation Computer Systems, vol. 100, pp. 656–673, 2019.

[17] Atul, D. J., Kamalraj, R., Ramesh, G., Sankaran, K. S., Sharma, S., & Khasim, S. “ A machine learning based IoT for providing an intrusion detection system for security”. Microprocessors and Microsystems, vol.82, pp.103741, 2021.

[18] J. Manan, A. Ahmed, I. Ullah, L. Merghem-Boulahia, and D. Gaïti, “Distributed intrusion detection scheme for next generation networks,” Journal of Network and Computer Applications, vol.147, p.102422, 2019.

[19] E. M. Dovom, A. Azmoodeh, A. Dehghantanha, D. E. Newton, R. M. Parizi, and H. Karimipour, “Fuzzy pattern tree for edge malware detection and categorization in IoT,” Journal of Systems Architecture, vol. 97, pp.1-7, 2019.

[20] S. Rathore and J. H. Park, “Semi-supervised learning based distributed attack detection framework for IoT,” Applied Soft Computing, vol.72, pp.79-89, 2018.

[21] A. P. Namanya, I. U. Awan, J. P. Disso, and M. Younas, “Similarity hash based scoring of portable executable files for efficient malware detection in IoT,” Future Generation Computer Systems, vol.110, pp.824-832, 2020.

[22] J. Gracia, A. J. Mazon, and I. Zamora, “Best ANN Structures for Fault Location in Single- and Double-Circuit Transmission Lines,” IEEE Transactions on Power Delivery, vol.20, no. 4, pp.2389-2395, 2005.

[23] S. Agatonovic-Kustrin and R. Beresford, “Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,” Journal of Pharmaceutical and Biomedical Analysis, vol.22, no.5, pp.717-727, 2000.

[24] C.-C. J. Kuo, “Understanding convolutional neural networks with a mathematical model,” Journal of Visual Communication and Image Representation, vol.41, pp.406-413,2016.

[25] K. G. Lore, D. Stoecklein, M. Davies, B. Ganapathysubramanian, and S. Sarkar, “A deep learning framework for causal shape transformation,” Neural Networks, vol.98, pp.305-317, 2018.

[26] Z. Kons and O. Toledo-Ronen, “Audio event classification using deep neural networks,” Interspeech 2013, 2013.

[27] D. Byrne, “Introduction: Case-Based Methods: Why We Need Them; What They Are; How to Do Them,” The SAGE Handbook of Case-Based Methods, pp.1-10, 2009.

[28] M. T. Do, N. Park, and K. Shin, “Two-Stage Training of Graph Neural Networks for Graph Classification,” Neural Processing Letters, 2022.

[29] S. Qian, H. Liu, C. Liu, S. Wu, and H. S. Wong, “Adaptive activation functions in convolutional neural networks,” Neurocomputing, vol.272, pp. 204-212, 2018.

[30] E. Micheli-Tzanakou, “Artificial Neural Networks: An Overview,” Network: Computation in Neural Systems, vol.22, no.1-4, pp.208-230, 2011.

[31] M. H. SAZLI, “A brief review of feed-forward neural networks,” Communications, Faculty Of Science, University of Ankara, pp.11-17, 2006.

[32] L. Raff, R. Komanduri, M. Hagan, and S. Bukkapatnam, “Feed Forward Neural Networks,” Neural Networks in Chemical Reaction Dynamics, Feb.2012.

[33] J. Švec, A. Chýlek, and L. Šmídl, “Hierarchical discriminative model for spoken language understanding based on convolutional neural network,” Interspeech 2015, 2015.

[34] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” 2017 International Conference on Engineering and Technology (ICET),2017.

[35] M. Zubair and C. Yoon, “Cost-Sensitive Learning for Anomaly Detection in Imbalanced ECG Data Using Convolutional Neural Networks,” Sensors, vol.22, no.11, p.4075, 2022.

[36] J. Wang and R. Lee, “Chaotic Recurrent Neural Networks for Financial Forecast,” American Journal of Neural Networks and Applications, vol.7, no.1, p.7, 2021.

[37] S. V. Chakravarthy and J. Ghosh, “Function Emulation Using Radial Basis Function Networks,” Neural Networks, vol.10, no.3, pp.459-478,1997.

[38] D. I. Kumar and M. R. Kounte, “Comparative study of self-organizing map and deep self-organizing map using MATLAB,” 2016 International Conference on Communication and Signal Processing (ICCSP), 2016.

[39] A. M. Kalteh, P. Hjorth, and R. Berndtsson, “Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application,” Environmental Modelling & Software, vol.23, no.7, pp.835-845,2008.

[40] J. A. Michaels, S. Schaffelhofer, A. Agudelo-Toro, and H. Scherberger, “A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping,” Proceedings of the National Academy of Sciences, vol.117, no.50, pp. 32124-32135, 2020.

[41] B. L. M. Happel and J. M. J. Murre, “Design and evolution of modular neural network architectures,” Neural Networks, vol.7, no.6-7, pp.985-1004,1994

[42] O. I. Abiodun et al., “Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition,” IEEE Access, vol.7, pp.158820-158846, 2019.

[43] Z. Mosaffaei and A. Jahani, “Modeling of ash (Fraxinus excelsior) bark thickness in urban forests using artificial neural network (ANN) and regression models,” Modeling Earth Systems and Environment, vol.7, no.3, pp.1443-1452, 2020.

[44] M. A. A. M. Maki and S. Subramanian, “Using an Artificial Neural Network to Improve Email Security,” Research Anthology on Artificial Neural Network Applications, pp.1465-1477, 2022.

[45] S.-C. Wang, “Artificial Neural Network,” Interdisciplinary Computing in Java Programming, pp.81-100,2003.

[46] S. Hanif, T. Ilyas, and M. Zeeshan, “Intrusion Detection In IoT Using Artificial Neural Networks On UNSW-15 Dataset,” 2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT),2019.

[47] Z. Lv, B. Hu, and H. Lv, “Infrastructure Monitoring and Operation for Smart Cities Based on IoT System,” IEEE Transactions on Industrial Informatics, vol.16, no.3, pp.1957-1962, 2020.

[48] F. Alshehri and G. Muhammad, “A Comprehensive Survey of the Internet of Things (IoT) and AI-Based Smart Healthcare,” IEEE Access, vol.9, pp.3660-3678, 2021.

[49] Xia, F., Yang, L. T., Wang, L., & Vinel, A. “IEEE Internet of Things Journal,” IEEE Internet of Things Journal, vol.3, no.2, pp. C3-C3,2016.

[50] M. Zhang, F. Sun, and X. Cheng, “Architecture of Internet of Things and Its Key Technology Integration Based-On RFID,” 2012 Fifth International Symposium on Computational Intelligence and Design, 2012.

[51] Y. Jie, J. Y. Pei, L. Jun, G. Yun, and X. Wei, “Smart Home System Based on IOT Technologies,” 2013 International Conference on Computational and Information Sciences, 2013.

[52] M. Kocakulak and I. Butun, “An overview of Wireless Sensor Networks towards internet of things,” 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), 2017.

[53] M. A. Matin and M. M. Islam, “Overview of Wireless Sensor Network,” Wireless Sensor Networks - Technology and Protocols,2012.

[54] M. Mohammed Sadeeq, N. M. Abdulkareem, S. R. M. Zeebaree, D. Mikaeel Ahmed, A. Saifullah Sami, and R. R. Zebari, “IoT and Cloud Computing Issues, Challenges and Opportunities: A Review,” Qubahan Academic Journal, vol.1, no.2, pp.1-7,2021.

[55] S. Huh, S. Cho, and S. Kim, “Managing IoT devices using blockchain platform,” 2017 19th International Conference on Advanced Communication Technology (ICACT), 2017.

[56] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. H. D. N. Hindia, “An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges,” IEEE Internet of Things Journal, vol.5, no.5, pp.3758-3773, 2018.

[57] https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

[58] I. Lee and K. Lee, “The Internet of Things (IoT): Applications, investments, and challenges for enterprises,” Business Horizons, vol.58, no.4, pp.431- 440, 2015

[59] https://www.kaggle.com/datasets/mkashifn/nbaiot-dataset

[60] How to Shuffle Pandas Dataframe Rows in Python - November 29, 2021.

[61] Giorgos Myrianthous (2021) “ Shuffling Rows in Pandas DataFrames - Discussing how to shuffle the rows of pandas DataFrames” ,2021.

[62] A. Daly, T. Dekker, and S. Hess, “Dummy coding vs effects coding for categorical variables: Clarifications and extensions,” Journal of Choice Modelling, vol.21, pp.36-41, 2016.

[63] Minitab Blog Editor, 10 February, 2016 “ When Is It Crucial to Standardize the Variables in a Regression Model?” , 2016.