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

# AN ENERGY-EFFICIENT SDN-IOT ARCHITECTURE WITH DEEP LEARNING-BASED TRAFFIC PREDICTION FOR IOT-ENABLED SMART CITY

**Noor Kadhim Salman AL-LAMI**

Master's Thesis

Supervisor

Assoc. Prof. Dr. Sefer KURNAZ

İstanbul, 2024

# AN ENERGY-EFFICIENT SDN-IOT ARCHITECTURE WITH DEEP EARNING-BASED TRAFFIC PREDICTION FOR IOT-ENABLED SMART CITY

**Noor Kadhim Salman AL-LAMI**

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2024

The thesis titled AN ENERGY-EFFICIENT SDN-IOT ARCHITECTURE WITH DEEP LEARNING-BASED TRAFFIC PREDICTION FOR IOT-ENABLED SMART CITY prepared by and submitted on 12/11/2023 has been **accepted unanimously** for the degree of Master of Electrical and Computer Engineering in Electrical and Computer Engineering.

Assoc. Prof. Dr. Sefer KURNAZ
Supervisor

Thesis Defense Committee Members:

| | | |
|---|---|---|
| Assoc Prof. Dr. Sefer KURNAZ | Department of Eletrical and Electroics Engineering, Altinbaş University | |
| Academic Title First LAST NAME | Department of …, University | |
| Academic Title First LAST NAME | Department of …, University | |

I hereby declare that this thesis meets all format and submission requirements for 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.

Noor Kadhim Salman AL-LAMI

Signature

# ABSTRACT

## AN ENERGY-EFFICIENT SDN-IOT ARCHITECTURE WITH DEEP LEARNING-BASED TRAFFIC PREDICTION FOR IOT-ENABLED SMART CITY

AL-LAMI, Noor Kadhim Salman

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

Supervisor: Assoc. Prof. Dr. Sefer KURNAZ

Date: 01 / 2024

Pages: 68

Thesis proposes an energy-efficient SDN-IoT architecture tailored for IoT-enabled smart cities. This architecture addresses the challenges of resource optimization and energy consumption management in the context of diverse IoT devices and dynamic traffic patterns. The architecture provides a framework for efficient network management and facilitates sustainable operation in smart city environments. A modified SVM (Support Vector Machine) algorithm is introduced and integrated into the proposed architecture for traffic prediction. By enhancing the traditional SVM algorithm with specific modifications tailored for IoT-generated data, the thesis contributes to improving the accuracy and reliability of traffic prediction in smart city networks. The modified SVM algorithm can effectively capture complex patterns and variations, enabling precise anticipation of traffic demands. The research presents a comprehensive evaluation of the proposed architecture's performance. Through extensive simulations and practical deployments in a smart city testbed environment, the paper assesses the energy efficiency, resource utilization, and predictive accuracy achieved by the architecture

**Keywords:** IoT, SVM, SDN, DL ,ML .

# TABLE OF CONTENTS

**Pages**

# LIST OF TABLES

**Pages**

# LIST OF FIGURES

**Pages**

# ABBREVIATIONS

| | | |
|---|---|---|
| IoT | : | Internet Of Things |
| SDN | : | Software _Defined Networking |
| AI | : | Artificial Intelligence |
| ML | : | Machine Leaning |
| SVM | : | Support Vector Machine |
| DL | : | Deep Learning |

# 1. INTRODUCTION

## 1.1 BACKGROUND

The development of smartphones with the acceleration of technological changes in recent decades has led to an increase in the production of big data and even a burst. Traffic which is one of the data sources is said to show internet traffic 20 percent by 2023. Data traffic generated by smartphones will exceed 86 percent of the entire traffic of mobile data with the rise of different mobile apps. Virtual reality, Internet of vehicles and live streaming can be given as examples for these applications.

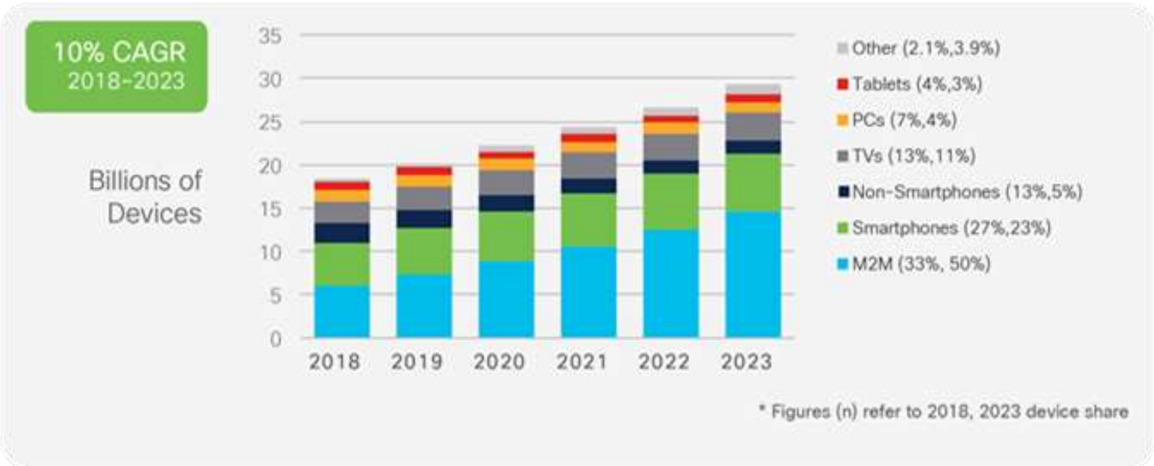


**Figure 1.1:** Traffic in 2023.

The increase in traffic of the mobile network will cause dissatisfied of costumers due to traffic overloading. It is known that something needs to be done to increase customer satisfaction by using available radio resources efficiently. It is possible to make better the quality of service supplied by mobile operators while allocating radio resources by analyzing the mobile network and the behaviour of customers. While this analysis is made, it is utilized from the data traffic in radio access network. For instance, forecasting of the use of spectrum can be used to help considerably to share spectrum.

Machine learning suggests a more intelligent and higher-level control and management of networks to fulfil various requirements of costumers. Therefore, the studies about applying algorithms of machine learning to mobile networks have been increased in recent years. It has been known that next-generation systems such as large-scale MIMOs, device-to-device (D2D) networks, heterogeneous networks constituted by femtocells and small cells, and so on will meet with some technical challenges in the future. Machine learning is one of the most effective artificial intelligence tools, which will be able to become a solution to these technical challenges [1].

Among information obtained from the communication network, one of the most used data sources in the researches is the call record details(CDR) data. For implementations of analysis of CDRs, call patterns, urban computing, and criminal investigation can be given as examples. A CDR, including temporal and spatial information, is a form in which lays practicable data regarding a given action of telephone containing a mobile user of a frame of a communication network. For this reason, CDRs have a major significance while making analysis and optimization in a communication network. Past CDR information analysed and modelled can be used to forecast the upcoming trend of CDRs, in other words, it takes an active role in allocating resources and providing load balance in the network [8].

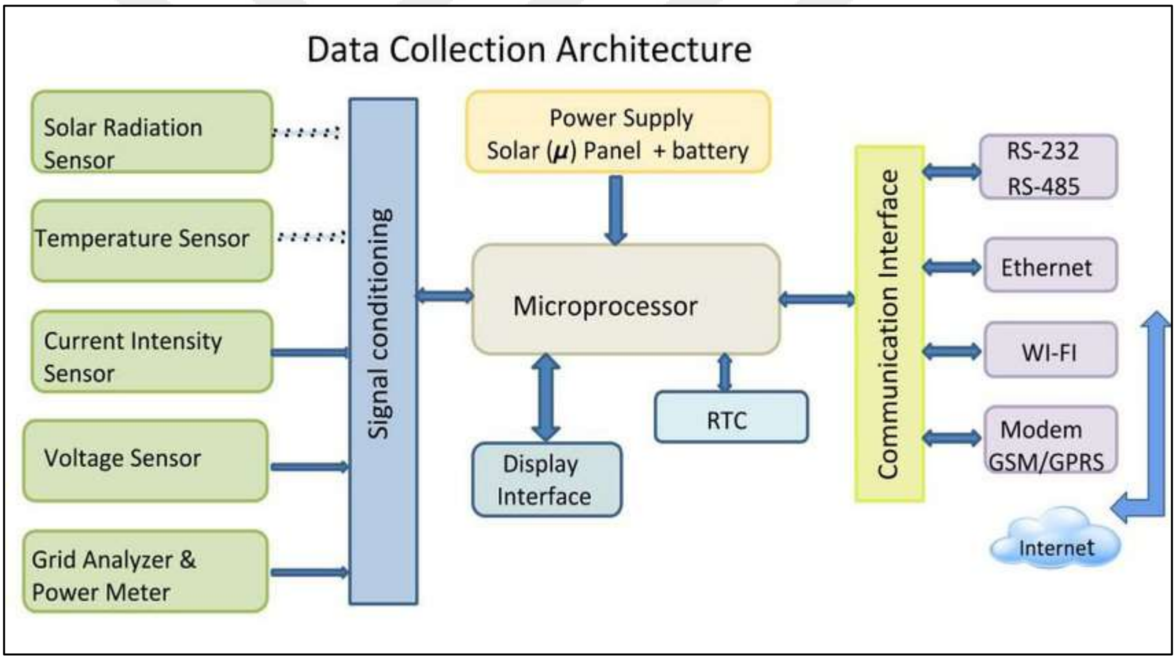


**Figure 1.2:** Process of Data Collecting in IOT.

The Telecom Italia Big Data Challenge dataset which is publicly open was utilized in order while applying this project. The dataset involves data of call/sms/internet which are named the call detail records (CDRs) of mobile users for each area in Milan, measured during November and December 2013. Objectives of this study are to examine thoroughly the structure of communication networks and to introduce a data-driven architecture for the practical applications of machine learning techniques to predict data traffic of the network of Milan. In this project, the algorithms utilized for internet activity traffic prediction were linear regression (LR), support vector regression (SVR), decision tree(DT) and neural network(feedforward).

## 1.2 PROBLEM STATEMENT

In the context of building IoT-enabled smart cities, the increasing number of interconnected devices and the corresponding surge in data traffic pose significant challenges for the efficient management of network resources and the optimization of energy consumption. Current SDN (Software-Defined Networking) architectures in IoT environments often lack energy efficiency and fail to adequately address the dynamic traffic patterns generated by diverse IoT devices, leading to suboptimal resource utilization and increased energy consumption. Additionally, the absence of accurate traffic prediction mechanisms further complicates the efficient allocation of network resources, making it difficult to adapt the network infrastructure to real-time traffic demands. Existing traffic prediction approaches often rely on simplistic models or traditional algorithms, which are unable to capture the complex patterns and variations exhibited by IoT-generated data. Therefore, there is a pressing need for an energy-efficient SDN-IoT architecture that leverages deep learning techniques for traffic prediction in IoT-enabled smart cities. By developing an architecture that optimizes energy consumption while accurately predicting traffic patterns, it will be possible to enhance the overall network performance, reduce resource wastage, and facilitate efficient decision-making for traffic management in smart city environments. Such an architecture would empower smart city operators to proactively allocate network resources, adapt to changing traffic demands, and ensure seamless connectivity and service quality for IoT devices, contributing to the realization of sustainable and intelligent urban environments.

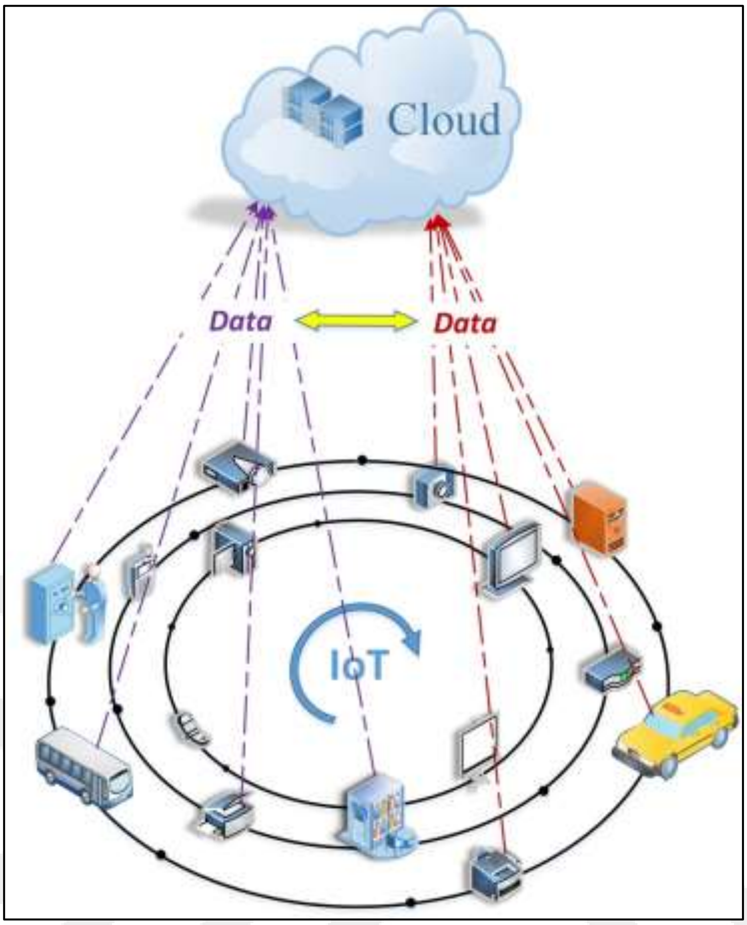

**Figure 1.3:** IOT Data Demand with the Cloud.

### 1.3 THESIS OBJECTIVES

To analyze the existing challenges and limitations in SDN-IoT architectures for smart cities, with a focus on energy efficiency and traffic prediction.

To propose an energy-efficient SDN-IoT architecture specifically designed for IoT-enabled smart cities, considering the unique characteristics and requirements of the IoT devices and the dynamic nature of traffic patterns.

To investigate and select appropriate deep learning techniques for traffic prediction in the proposed architecture, considering their ability to capture complex patterns and variations exhibited by IoT-generated data.

To develop and implement a deep learning-based traffic prediction module within the SDN-IoT architecture, utilizing real-time IoT traffic data to train and validate the predictive models.

To evaluate the performance of the proposed architecture in terms of energy efficiency, resource utilization, and accuracy of traffic prediction, using both simulation-based experiments and practical deployments in a smart city testbed environment.

To compare the performance of the proposed architecture with existing SDN-IoT architectures and traditional traffic prediction methods, highlighting the improvements achieved in terms of energy efficiency, resource optimization, and predictive accuracy.

To provide recommendations and guidelines for the deployment and operation of the energy-efficient SDN-IoT architecture in real-world smart city environments, addressing potential challenges and suggesting future research directions in the field.
To contribute to the advancement of sustainable and intelligent urban environments by demonstrating the feasibility and benefits of integrating an energy-efficient SDN-IoT architecture with deep learning-based traffic prediction in IoT-enabled smart cities.

## 1.4 THESIS MOTIVATIONS

Rapid urbanization and the increasing deployment of IoT devices in smart cities have led to a massive influx of data traffic, posing challenges in managing network resources and optimizing energy consumption. Addressing these challenges is crucial for the efficient operation of smart city infrastructures. Existing SDN-IoT architectures often lack energy efficiency and fail to adequately handle the dynamic traffic patterns generated by diverse IoT devices. By developing an energy-efficient architecture, it becomes possible to minimize energy consumption, reduce costs, and improve the overall sustainability of smart city deployments. Accurate traffic prediction is essential for effective resource allocation and decision-making in SDN-IoT architectures. Traditional prediction methods are often insufficient to handle the complex and dynamic nature of IoT-generated data. The application of deep learning techniques offers the potential to significantly enhance the accuracy and reliability of traffic prediction in IoT-enabled smart cities. The development of an energy-efficient SDN-IoT architecture with deep learning-based traffic prediction can contribute to the realization of sustainable and intelligent urban environments. By optimizing energy consumption and resource utilization, it becomes possible to enhance the performance, reliability, and scalability of smart city networks, leading to improved quality of life for citizens and efficient management of urban services. The proposed research fills a gap in the existing literature by specifically focusing on the integration of energy efficiency and deep learning-based traffic prediction within an SDN-IoT architecture for smart cities. The findings of this thesis can provide valuable insights and practical solutions for researchers, network operators, and policymakers working towards the development of efficient and sustainable smart city infrastructures. The outcomes of this research can have real-world implications, as they can guide the design, implementation, and operation of SDN-IoT architectures in actual smart city deployments. By demonstrating the feasibility and benefits of the proposed architecture, this thesis contributes to the advancement and adoption of innovative technologies that support the growth and development of smart cities.

## 1.5 THESIS CONTRIBUTIONS

The thesis proposes an energy-efficient SDN-IoT architecture tailored for IoT-enabled smart cities. This architecture addresses the challenges of resource optimization and energy consumption management in the context of diverse IoT devices and dynamic traffic patterns. The architecture provides a framework for efficient network management and facilitates sustainable operation in smart city environments. A modified SVM (Support Vector Machine) algorithm is introduced and integrated into the proposed architecture for traffic prediction. By enhancing the traditional SVM algorithm with specific modifications tailored for IoT-generated data, the thesis contributes to improving the accuracy and reliability of traffic prediction in smart city networks. The modified SVM algorithm can effectively capture complex patterns and variations, enabling precise anticipation of traffic demands. The research presents a comprehensive evaluation of the proposed architecture's performance. Through extensive simulations and practical deployments in a smart city testbed environment, the thesis assesses the energy efficiency, resource utilization, and predictive accuracy achieved by the architecture. The evaluation highlights the advantages and improvements obtained by incorporating the modified SVM algorithm, showcasing the efficacy of the proposed approach. The thesis compares the performance of the modified SVM algorithm with existing traffic prediction methods in SDN-IoT architectures. By conducting rigorous comparative analyses, the research demonstrates the superiority of the modified SVM algorithm in terms of accuracy, efficiency, and adaptability to IoT-generated traffic patterns. These findings contribute to the advancement of traffic prediction techniques in the context of IoT-enabled smart cities. Practical recommendations and guidelines are provided for the deployment and operation of the proposed energy-efficient SDN-IoT architecture with the modified SVM algorithm. The research identifies potential challenges and suggests strategies to overcome them, helping network operators and policymakers implement the architecture effectively. These recommendations contribute to the practical implementation and adoption of energy-efficient and accurate traffic prediction mechanisms in real-world smart city deployments. The thesis makes broader contributions to the field of smart city development by highlighting the importance of integrating energy efficiency and accurate traffic prediction in SDN-IoT architectures. The research findings reinforce the significance of sustainable and intelligent urban environments, promoting the use of advanced technologies to optimize resource utilization, reduce energy consumption, and enhance the overall performance of smart city networks.

# 2. LITERATURE REVIEW

## 2.1 CHAPTER INTRODUCTION

The evolution of wireless networks and the advent of next-generation technologies, such as 5G, bring forth numerous challenges in terms of network optimization, traffic prediction, and resource management. In response to these challenges, researchers have explored the application of machine learning techniques in various aspects of wireless networks, including heterogeneous networks, device-to-device networks, and large-scale MIMOs. Machine learning has shown potential in addressing technical challenges and optimizing the performance of wireless networks, particularly in the context of smart cities and smart grids.

This literature review encompasses studies that investigate the application of machine learning paradigms in next-generation wireless networks. Emphasize the importance of machine learning in modeling the challenges of future wireless systems and highlight the potential of spectral efficiency inference and learning in achieving high spectral bandwidths [1].

## 2.2 RELATED WORKS

Studied machine learning paradigms for next-generation wireless networks. They addressed that the next- generation wireless networks are assumed to be supported higher data rates and entire recent implementations working in the paradigm of wireless radio technology [1]. Moreover, it was forecasted that the future smart terminals of 5G are predicted to autonomously reach high spectral bandwidths with the help of the sophisticated spectral efficiency inference and learning [1]. However, machine learning could be a solution in modelling different technic challenges of next-generation systems like heterogeneous networks, the device-to-device networks and large-scale MIMOs [7]. In addition, some research considered that the machine learning models and data mining are so significant for optimization and efficient performance of smart grid systems being excellent applications of next-generation networks. This view supports the study of [1]. Presented a few effective machine learning implementations in Self-Organizing Networks

(SONs). Moreover, they debated the positive and negative of various machine learning methods and address directions of future study in this field. Investigated traffic offloading methods in wireless networks and offered reinforcement learning from machine learning algorithms [10]. Additionally, [9] believed that this view can emerge a novel study direction toward embedding machine learning towards greening cellular networks studied the role of proactive caching in 5G wireless networks.

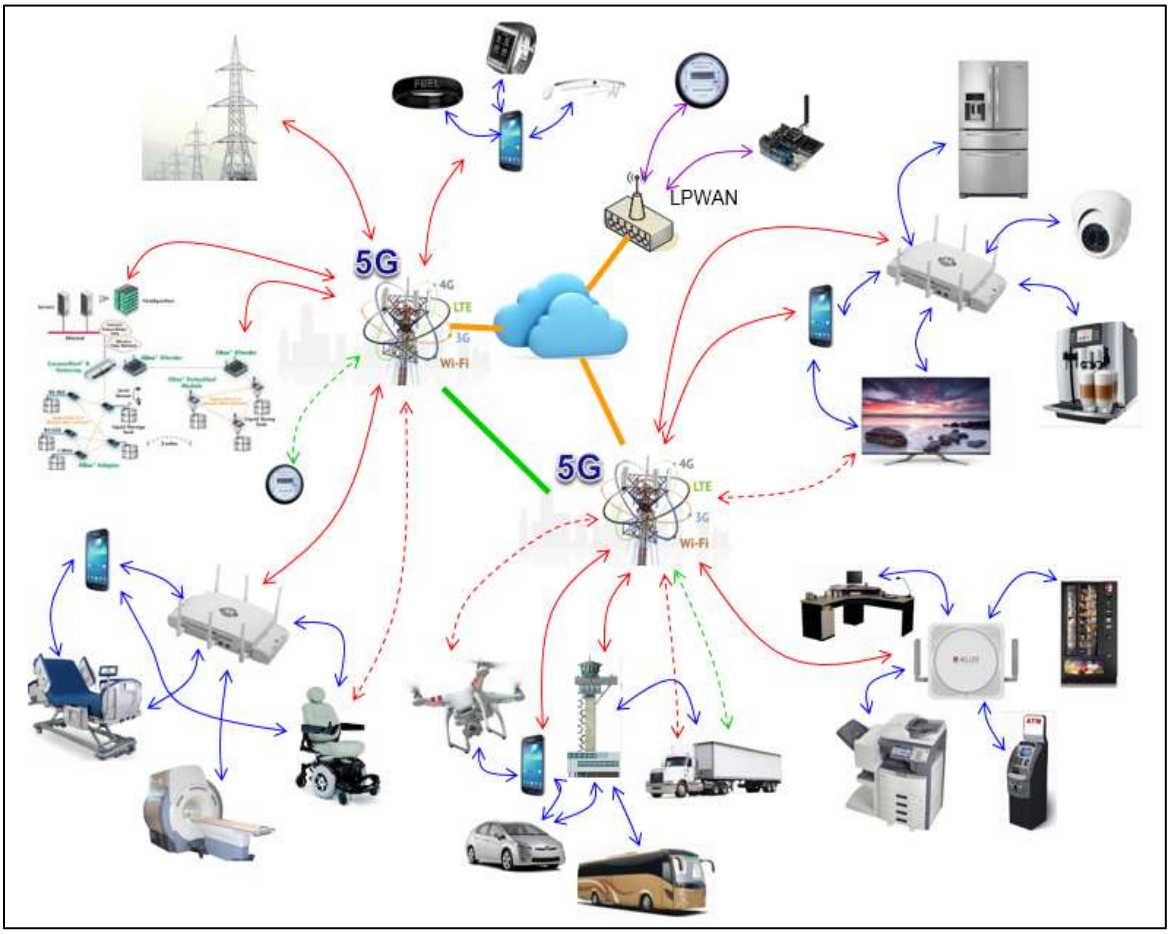


**Figure 2.1:** IOT Structure In 5G Networks.

They addressed that proactive caching indicate important gains through backhaul savings and growth in service delivery between many consumers and also raised storage capacity. 5G networks showed peak data traffic requests can be decreased by serving the user requests proactively. The research specifies that supervised machine learning algorithms can be combined with a proactive caching procedure to increase the potential of utilizing the predictive capabilities of 5G networks. The predictive abilities of 5G networks provided aid to the large part of the mobile operators to reduce the problems data which continues to strain current networks. On the other side, [8] investigated machine learning aided cognitive RAT selection for 5G heterogeneous networks. Furthermore, The Heterogeneous Networks strategy play a role as the key Radio Access Network architecture for future 5G networks. This can cause critical problems to mechanisms of the current user association utilized in cellular networks. Therefore, they proposed the reinforcement learning method over supervised machine learning because of its flexibility and low computational complexity for learning the policy of user association. [35] suggested a new method to scale resources of 5G core network by predicting traffic load changes with the use of machine learning algorithms. They applied it by utilizing and training a Neural Network. Additionally, the outcomes indicated that the prediction-based scalability process makes better the threshold-based solutions such as latency to react to traffic change, and delay to have novel resources which are ready to be utilized by the VNF to react to traffic growth. [23] firstly presented a construction to gather and process the big data in the network so that they investigated the connection among the key progress indicators and the traffic patterns.

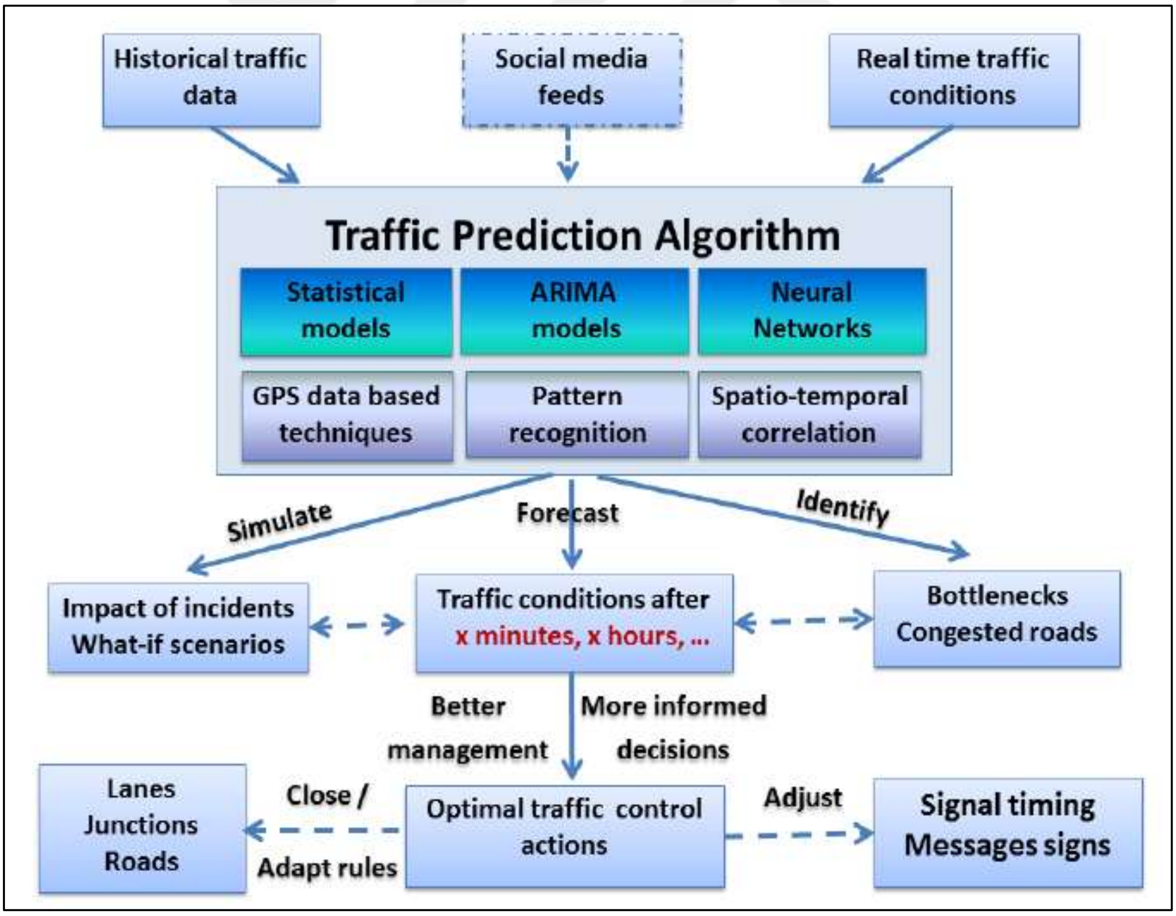


**Figure 2.2:** Traffic Prediction Based on the ARIMA Model.

An algorithm using machine learning for traffic prediction was also introduced to cope with wide cells which generate voice traffic and data traffic. In addition, this study utilized the Bayesian network (BN) to investigate the probabilistic relationships between variables. According to this study, autoregressive and neural network methods showed alike performances. However, the gaussian process indicated the best performance as dealing with fast alterations in traffic [23]. It has been developed a process and created a tool for intelligent and fruitful network planning. This tool utilizes and learns from the data based on working history collected over a network anywhere and at any time. In this research, they gathered measurements of user equipment (UE) in accordance with 3GPP MDT functionality to construct quality of service forecast depending on past information. Methods of regression analysis were implemented to predict the Physical Resource Blocks (PRB) per transmitted Mb. According to research results, when the entire regression algorithms such as k-NN, bagged-SVM and NN were applied, they showed great accuracy [5]. However, [5] believe that the approach of bagged -SVM learning is more appropriate for the requirements of this study. Bagged -SVM learning presents more precise forecasts. [3] studied a data-driven association methods among the Next Generation Node Base(gNBs) and the RAN controllers. Moreover, they implemented different methods of machine learning adapted for forecasts such as the Bayesian Ridge Regressor (BRR) for the local-based prediction, and the Gaussian Process Regressor (GPR) and Random Forest Regressor (RFR) for both the local and the cluster-based forecasts. The clustering solution defined in this article has restricted the interplays between various controllers to decrease the requirement for synchronization of inter-controller and minimize the latency of control plane. In other research, the traffic prediction challenge was handled the call detail records(CDRs)in the city of Milan. The assessed architectures of deep learning with multitask learning improve the accuracy of the forecasts by implementing the convolutional neural network (CNN) and the recurrent neural network (RNN) to capture features of spatial and temporal successfully [27]. Furthermore, [7] investigated machine learning algorithms used artificial neural networks for Internet traffic prediction. They believe that RNN performed better and it is more suitable for time series network traffic prediction. [30] pointed out that it is not easy to predict and characterize the patterns of traffic because the traffic pattern of each base station is active at diverse locations and times. Moreover, they suggested a structure of C-RAN optimization using deep learning to solve the difficulties. By utilizing the multivariate Long Short Term Memory (MuLSTM) model they aimed to understand the spatial correlation and the temporal dependence between traffic models of base stations and at the same time to forecast the traffic more accurately for the future time. In addition, Telecom Italia Dataset which is data of a real-world network was used in this study. It was noted in the article that the outcomes display that the model used significantly enhances capacity utilization and decreases the total cost of deployment of conventional RAN structure. [22] studied on a model which integrates analysis of real-world telecommunication data and multivariate forecast system, fitting for communication networks in urban areas. Furthermore, they claim that causality analysis operated in practice first time while analysing telecommunication data which was taken by utilizing from the spatial-temporal model.

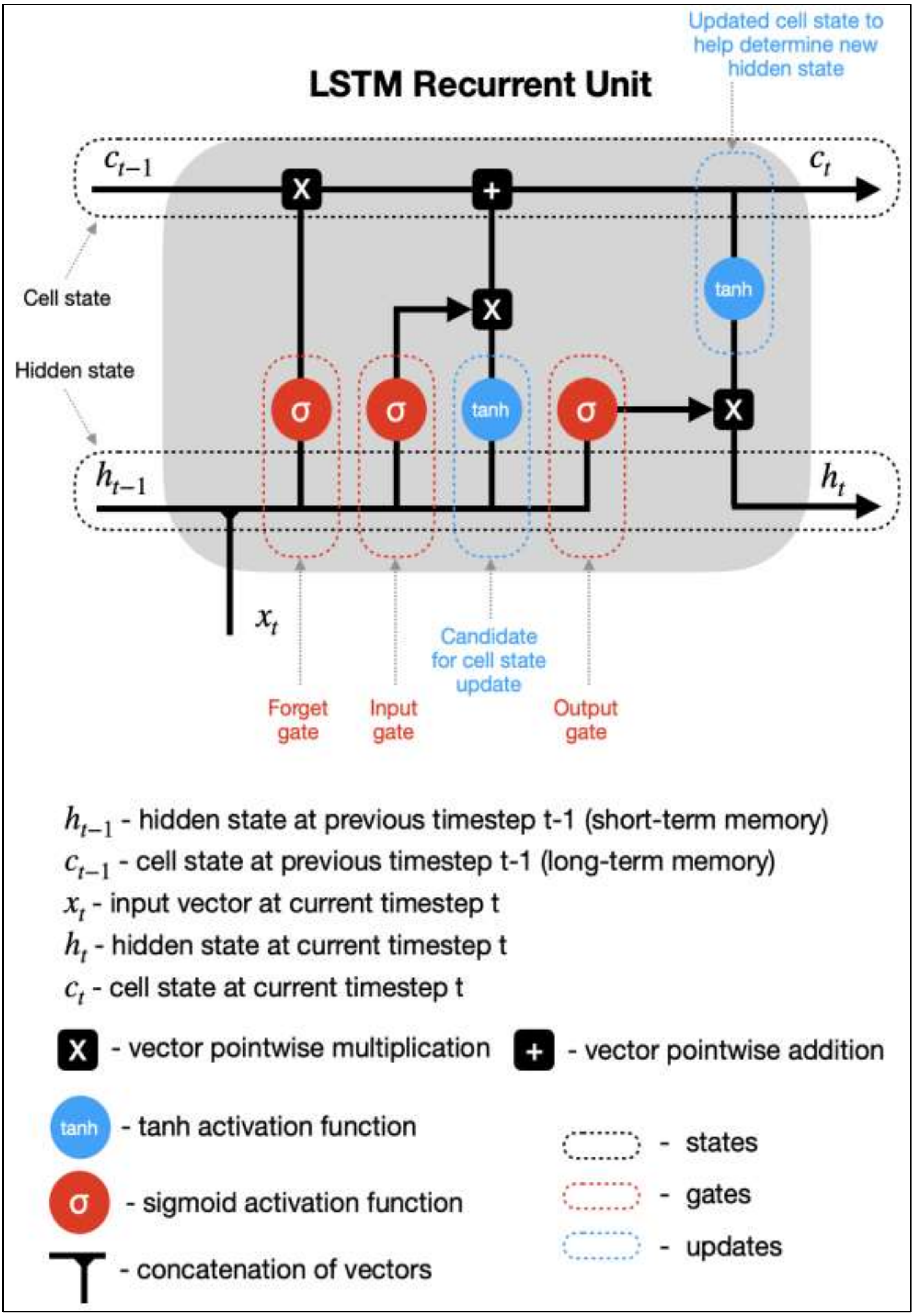


**Figure 2.3:** LSTM Algorithm Model.

Using the multivariate LSTM algorithm, future traffic data were also forecasted with the data obtained from the analysis. Based on the outcomes of the experiment, they emphasize that casualty analysis can increase the performance of forecasting of multivariate time series. [19] investigated Spatial-Temporal Cross-domain neural network(STCNET) which is one of the new structure of deep learning to get the complicated and confidential patterns in traffic data efficiently. They argue that STCNet, one of the components of convolutional long-short term memory, is good for overcoming problems associated with spatial-temporal dependence when forecasting traffic. Furthermore, three different domain datasets were used to capture the outer elements influencing the creation of traffic on this research. Since there are differences and similarities between traffic data in different regions of cities, the inter-cluster transfer learning strategy was designed to reuse information by using the clustering model to separate parts of the city into diverse categories. In addition to the STCNet model, performances of Linear Regression (LR), Support Vector Machine (SVM), LSTM networks, and Spatial-Temporaly Densely Connected CNN (DenseNet) algorithms were calculated and compared to prove STCNet's superiority in predicting traffic in mobile communication networks. This study emphasizes that transfer learning with STCNet improves performance even further.

## 2.3 SUMMARY OF RELATED WORKS

Discuss effective machine learning implementations in Self-Organizing Networks (SONs) and present directions for future research in the field [24].

Traffic offloading methods in wireless networks are investigated by [25] who propose reinforcement learning from machine learning algorithms to optimize traffic management. Additionally, [10] explore the role of proactive caching in 5G networks, emphasizing the benefits of utilizing supervised machine learning algorithms to enhance the predictive capabilities of these networks.

The thesis also delves into the application of machine learning in heterogeneous networks, particularly in cognitive Radio Access Technology (RAT) selection. [15] compare reinforcement learning methods with supervised machine learning for user association policies in heterogeneous networks. Furthermore, [33] propose a machine learning-based method for scaling resources in the 5G core network by predicting traffic load changes.

Traffic prediction is another crucial aspect covered in the literature review. Le et al. (2021) investigate the connection between key performance indicators and traffic patterns, introducing a machine learning algorithm for traffic prediction. Moysen et al. (2019) develop a tool for intelligent network planning using regression analysis and machine learning algorithms. [40] evaluate architectures combining deep learning and multitask learning for traffic prediction.

Other studies focus on the optimization of wireless networks through deep learning techniques. [19] propose a structure for Cloud Radio Access Network (C-RAN) optimization using multivariate Long Short-Term Memory (LSTM) models. [19] integrate causal analysis with multivariate forecasting systems for communication networks. [28] introduce the Spatial-Temporal Cross-domain neural network (STCNET) for efficient pattern recognition in traffic data.

**Table 2.1:** Summary of Related Works.

| Study | Key Focus |
|---|---|
| Jiang et al. (2019) | Machine learning paradigms for next-gen wireless networks |
| Klaine et al. (2019) | Machine learning implementations in SONs |
| Chen et al. (2020) | Traffic offloading with reinforcement learning |
| Baştuğ et al. (2014) | Proactive caching in 5G networks |
| Pérez et al. (2019) | Machine learning aided cognitive RAT selection |
| Alawe et al. (2021) | Scaling 5G core network with traffic prediction |
| Le et al. (2021) | Traffic prediction using machine learning algorithms |
| Moysen et al. (2019) | Intelligent network planning with regression analysis |
| Huang et al. (2019) | Deep learning architectures for traffic prediction |
| Chen et al. (2021) | C-RAN optimization with multivariate |

# 3. MATERIALS AND METHODS

## 3.1 INTRODUCTION

This project focused on which machine learning algorithms are able to be applied to traffic prediction for communication networks to ensure a requirement of a better quality of service, and how this process was followed. Therefore, this project accepted the method of qualitative research including the use of secondary data and resources like literature materials. This data was utilized to increase the deep comprehension of the developments to reinforce the suggestions of successful machine learning methods concentrated on increasing performance of communication networks and service delivery.

The first step of this study was to investigate the related work made on this topic in recent years. Secondly, after obtaining enough information about communication networks, radio access network(RAN), machine learning techniques and big data processing and analysing recent studies on this area, the dataset utilized was chosen and Matlab is a program which will be used during this project. Moreover, taking consideration of the existing machine learning implementations in communication networks, it was decided to use different supervised machine learning algorithms such as linear regression and support vector regression. Another step was to process the dataset to make them easy to use. Following this, the necessary features of the processed information were extracted in order to use for machine learning. Then forecasting was made based on the results of algorithms applied to data. Additionally, root mean square error (RMSE) and mean absolute error (MAE) values were calculated in order to evaluate accuracy and outcomes were compared. Finally, it was identified which is better among machine learning algorithms implemented in this project.

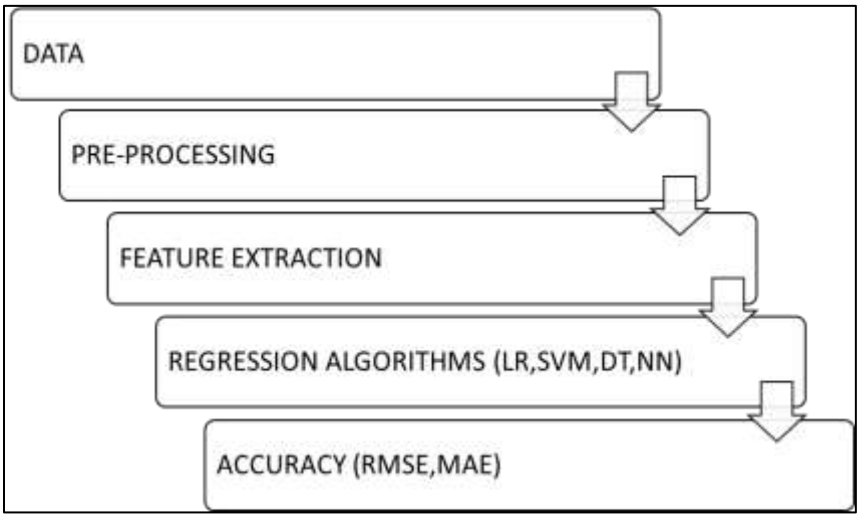


**Figure 3.1:** Flowchart for Methodology.

## 3.2 TRAFFIC PREDICTION

In this project, traffic prediction using machine learning in communication networks was studied. Unstructured data used in this study comes from Radio Access Network (RAN) where is responsible for resource allocation in a communication network. Machine Learning methods have been applied to unstructured big data in RAN to ensure patterns and to gather information about the status of network and make better resource allocation. In this part, the main topics mentioned are wireless cellular technologies, radio access network, some recent communication networks, machine learning and the algorithms used for this project, data processing, accuracy, and programming.

### 3.2.1 Wireless Cellular Technologies

The technologies of mobile wireless have begun to create, improve and make revolution since the beginning of the 1970s. Over the last few decades, 4-5 generations of technology revolution have taken place in technologies of mobile communication from 0G to 4G(Bhalla et al., 2010). In addition, it is expected to switch to 5G in early 2020. Some features of mobile communication generations mentioned are below:

a. 1G: Cellular concept was mentioned in 1G that had analog cellular technology for the first time. The data bandwidth of 1G was 2 Kbps and it provided basic voice service. In addition, it used Advanced mobile phone service(AMPS) as a standard and FDMA as a multiplexing technique. After, 2G was developed with more superior features than 1G, which incorporates digital communication technologies, taking into account the negative characteristics of 1G technology such as low capacity, bad voice connectivity and untrustworthy handoffs(Bhalla et al., 2010).

b. 2G: As mentioned above, when the second generation (2G) wireless mobile systems are compared with the first generation systems, it will be seen that contrary to 1G, the secondgeneration wireless systems used digital modulation methods such as time division multiple access (TDMA) and code division multiple access (CDMA). Moreover, Moreover, 2G provides SMS messaging in addition to the call service as primary services, and the most commonly utilized 2G standard is GSM. One of the weaknesses of 2G whose data bandwidth is 64 Kbps, is rates of limited data.

c. 3G: The technology of 3G intended to ensure 144kbps - 384kbps data rates for large coverage regions and 2Mbps for local coverage regions. Internet-based, E-mail, video teleconferencing and services of multimedia occurring of mixed audio and data streams are some conceivable implementations of 3G. 3G wireless communication standards come from the IMT2000 family and were advanced by ITU-R. The basic standards of IMT-2000 are below:

a. UMTS / WCDMA: The Universal Mobile Telecommunications System (UMTS), which utilizes Wideband Code Division Multiple Access (WCDMA), has been widely used in the rapidly expanding European Union system.
b. CDMA2000: A US system, the first standard distributed utilizing CDMA techniques.
c. TDS-CDMA: The system, which emerged in China and accepted many facts of GSM / UMTS, was optimized for Time Division Duplex (TDD).

d. 4G: 4G technology is a fully IP based network system. Integration is a word that defines the attributes of 4G. In other words, it has the ability to integrate present and technologies of the future wireless network (e.g OFDM) to provide freedom of mobility and uninterrupted roaming.
e. 5G: It, the next-generation communication system, will allow many new applications to emerge and be realized. 5G technology is expected to have many advantages such as high speed, tremendous capacity, IoT ability and low latency. In addition, 5G

networks have already begun to appear in some countries and are said to be spread around the world by 2020.

#### 3.2.1.1 Main factors determining demand for 5G

The main goal for 5G mobile networks is to provide distinct services according to customers' needs in an efficient way. In other words, the quality of service is the basic request.

Some of these requests are as follows:

a. The increment in the number of devices requiring connection
b. Need for higher data rates than currently offered data rates
c. Near-zero latency
d. Consumption of Energy

#### 3.2.1.2 Key attributes for 5G

a. Large Scale Antennas: Base stations(BSs) utilizing large scale antennas ensure more spatial multiplexing, which allows supporting a lot of customers(multi-access) per time-frequency signalling resource. Large scale antennas used can provide an increase in the gain of transmitter antenna(downlink) or the gain of the receiver antenna(uplink) in the formula of the Friis transmission, where causes higher SNR.
b. Milimeter Wave Communications**:** It is one of the favourite technologies of 5G, which is still under development. It enables the usage of signals having highfrequency in the unlicensed spectrum.
c. Flexibility: The capability of a network to comply quickly to an altering need.

Furthermore, other features contain cell size and density of base station,networks of a device to device communication with multiple radio access technology.

### 3.2.2 Radio Access Network

A radio access network (RAN) which links single devices to other network components by establishing a radio connection is an important element of the communication system. RAN consists of a base station and antennas with a varying coverage area relying on their capacities. Moreover, a RAN makes radio resource management possible. When a machine is wirelessly linked to the core network, the RAN conveys machine's signal to diverse endpoints, and this signal moves along with the traffic of other networks.

RAN, which has been used since the early days of mobile communication technology, has developed among generations of cellular communications. For example; GRAN, GERAN, UTRAN and E-UTRAN(Ries et al., 2007).

a. Generic Radio Access Network (GRAN): Its components are base transmission stations and controllers. It is in charge of radio connections of circuit-switched and packet-switched core networks.
b. GSM Edge Radio Access Network (GERAN): It is available for Real-time packet data.
c. UMTS Terrestrial Radio Access Network (UTRAN): It is used for both circuitswitched and packet-switched services
d. Evolved Universal Terrestrial Radio Access Network (E-UTRAN): E-UTRAN is only used on packet-switched services. It ensures low latency and high data rates.

RANs of the next generations of cellular communication requires to be more complicated to be able to support technologies of software-defined networking (SDN) and network functions virtualization (NFV) and millimetre wave (mmWave). Therefore, radio access networks as well will benefit from virtualization and cloud technologies. Virtualization (vRAN) and Cloud (C-RAN) are examples of RANs in the future.

#### 3.2.2.1 E-UTRAN in LTE network

The evolved Universal Terrestrial Radio Access Network (E-UTRAN), with evolved base stations called eNodeB or eNB, manages radio links between mobile and evolved packet cores.

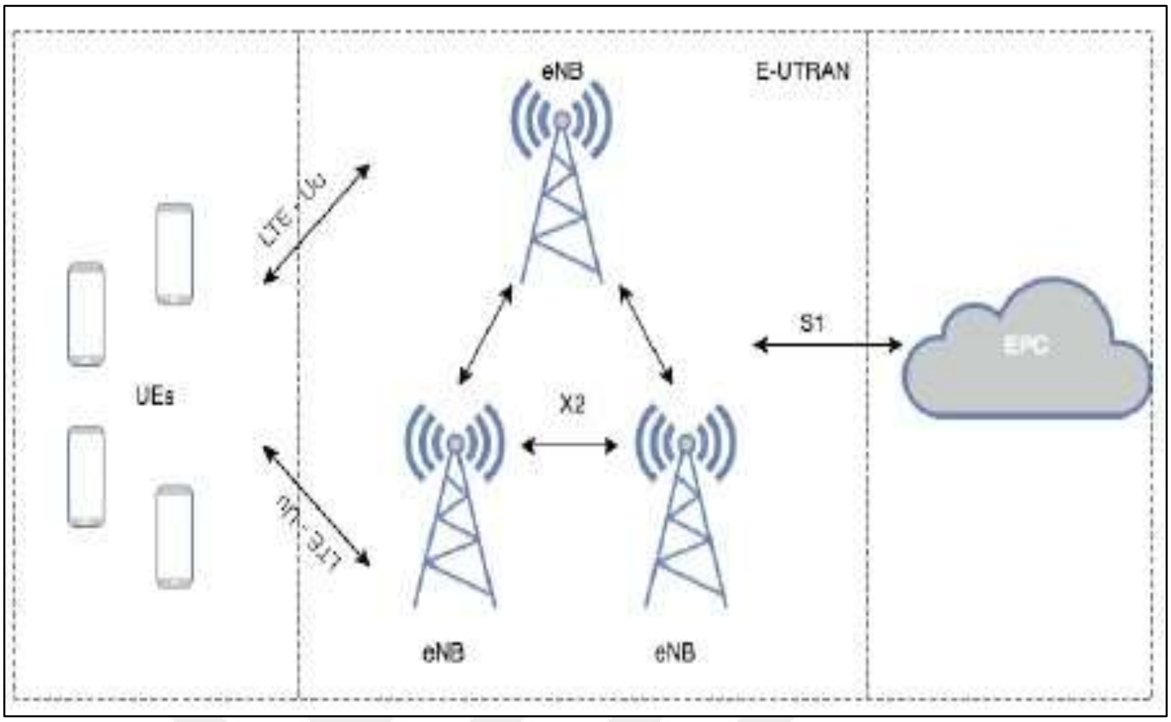


**Figure 3.2:** LTE Network Architecture.

E-UTRAN whose basic elements are defined below has known as RAN of LTE network.

a. UE: The user equipment (UE). For instance; a mobile phone, a computer or any remotely controlled device.
b. eNB: An evolved B (NB) is a base station in LTE. The eNB transmits and receives radio signals to entire mobile devices utilizing the analogue and digital functions of signal processing of the LTE air interface. The eNB is in charge of the low-level operation of the entire mobile devices, transmitting to them signalling messages like handover commands.
c. X2 interface: It is the interface among the eNBs ensuring links.
d. EPC: the evolved packet core (EPC) which is the core network of the LTE system is IP-based network.
e. S1 interface: It is the interface between the evolved Universal Terrestrial .

The network architecture of LTE is indicated in the figure(3.1). It can be seen that EUTRAN occurs of eNBs which supplies a direct link to the UE. In the E-UTRAN, eNBs are linked via X2 interface which is mainly utilized for signalling and packet forwarding throughout handover. The link among E-UTRAN and EPC is achieved by the S1 interface.

a. Radio Resources Management in LTE

Radio Resources Management (RRM) is a basic application-level function of the eNB in LTE network, providing the effective usage of the present radio resources(Das, 2012). Some of them are defined below:

a. Radio Admission Control(RAC): It accepts or refuses demands of installation for new radio bearers. The aim of RAC is to provide usage of high radio resource by admitting demands of radio bearer as long as radio resources are present. It concurrently provides suitable QoS for in-progress sessions, refusing radio bearer demands when they may not be ensured.
b. Radio Carrier Control (RBC): RBC includes installation, maintenance and release of radio carriers for mobile users. Additionally, It is related to the maintenance of radio carriers of ongoing sessions in the event that the radio source changes. For instance, because of mobility. RBC takes a role in the release of radio resources related to radio carriers containing at session ending and handover.
c. Dynamic Resource Allocation(DRA): It is in charge of allocating and de-allocating resources which involves resources of buffer and processing. It also allocates resource blocks in uplink and downlink. The other name of DRA is Packet Scheduling(PS).
d. Inter cell Interference co-ordination(ICIC): It is in charge of for radio blocks to control intercell interference depending on feedback from multiple cells.
e. Load Balancing(LB): It is in charge of dealing with the uneven traffic distribution load over multiple cells.

b. Cognitive Radio Network

Cognitive Radio Network is a type of network of wireless communication which intelligently perceives the convenient channels in the spectrum by using a transceiver that takes part in itself so that it can utilize the best channels in its surrounding. Cognitive Radio aims to provide sharing of radio resources in the best way and to reduce interference arising due to reusing the spectrum.

In addition, many kinds of research have been made by methods of machine learning and data mining in the area of CR. Implementations of machine learning are so popular for CR since it developed from software-defined radio(SDR).

### 3.2.2.2 C-RAN in 5G networks

Cloud Radio Area Networks(C-RAN) which is presented by China Mobile to handle the arising issues due to extreme growth in the cellular network is a new structure based on the idea of virtualization. It is capable of dealing with such a number of base stations the network needs thanks to virtualization. Centralization and sharing have a positive impact on more dynamic traffic management and deployments of base stations, which means better use of resources. Moreover, such structure will have a possibility to reduce the cost of the expenses since base stations are virtualized rather than physically positioned in diverse regions. It xdecreases the use of energy and power when it is compared with conventional networks by the reason of the fact that base stations would be positioned on the identical physical machine(Salman, 2016).

A design of C-RAN occurs of three main components called the Baseband Unit (BBU) pool, the Remote Radio Unit (RRU) networks, and the perimeter network. Features of the main components of C-RAN are summarized below(Huang et al., 2014):

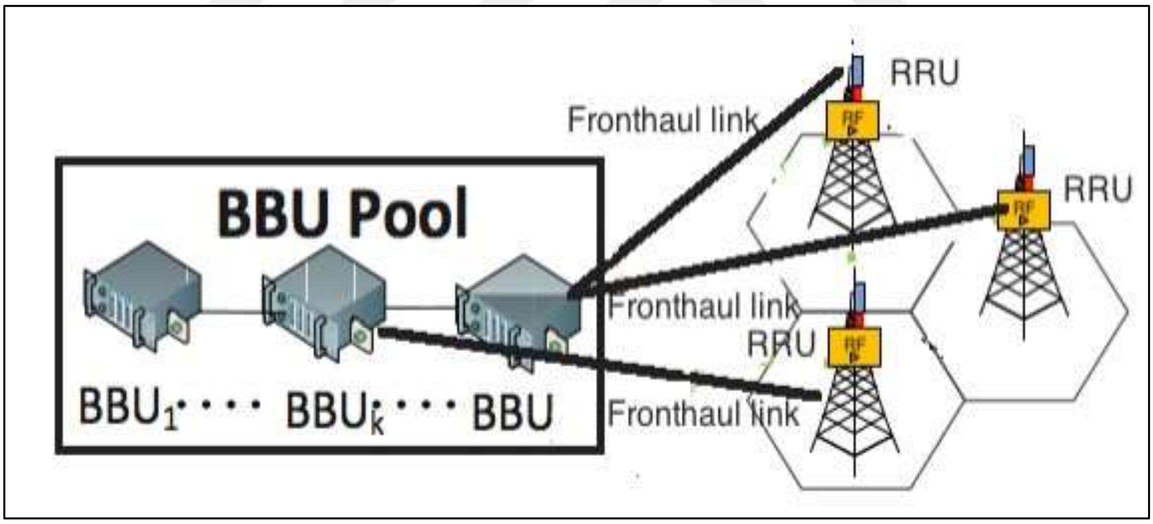


**Figure 3.3:** C-RAN Architecture(Salman,2016).

Baseband Unit (BBU) pool: a BBU pool which is found at a centralized location such as a cloud or information centres occurs of multiple nodes of BBU that have excellent abilities of computation and storage. In addition, BBUs process resources and dynamically allocate them to Remote Radio Units (RRU) depending on the present network requires.

a. Remote Radio Unit (RRU) networks: It is a wireless network which links wireless devices. RRU network works in a similar way to access points or towers in conventional mobile networks.
b. Fronthaul or transport network: Fronthaul is the link-layer among a BBU and a series of RRUs that ensure high bandwidth connections to meet the needs of multiple RRUs. Moreover, Fronthauls could be implemented by utilizing various technologies which contain optical fiber communication, cellular communication or millimetre-wave communication. It ensures the highest bandwidth needs. Therefore, optical fiber communication is contemplated to be ideal in C-RAN. Nevertheless, this implementation causes a high expense because it won't be flexible. On the other side, cellular communication or millimetre-wave communication is low-priced and easier to use than optical fiber communication. In addition, However, cellular or millimeter-wave communication results in less bandwidth and higher latency costs.

In addition, old virtualization methods that are used in IT clouds won't meet requirements of C-RAN. Therefore, more complicated technologies have been suggested finding a solution such as Network function virtualization(NFV) and Software defined network(SDN) to problems that will arise.

a. Network Function Virtualization

The purpose of Network Function Virtualization(NFV) is to handle the issue of virtualization by utilizing standard virtualization methods to deliver a variety of network hardware devices. These devices of network cover high-price routers and switches which are modified by software applied on high-volume servers, switches in the clouds, and elements of storage taking a place in the clouds. The advantages of this technology include reduced cost, the power and location of this equipment, and the ability to easily adapt devices. Another advantage is that operators can have resilient services depends on geographic position and privileges of the user. Furthermore, it will be more easily to be able to share resources with other operators on the identical server.

In addition, NFV technology, which displays significant advantages, is most likely to influence C-RAN implementation in practical application in the near future because it provides ease of use of the software as hardware elements supported by diverse suppliers.

b. Software-Defined Network

Software-defined network(SDN) at 5G or cellular network of next generation enables network management to be cost-effective by offering novel abilities and solutions for programmable, centrally controlled, flexible and high bandwidth implementations. Actually, the basic network infrastructure in SDN is completely isolated from implementations and the intelligence of networking is centrally controlled(Kitindi et al., 2019).In other words, This technology, which develops greatly with researches made in networking, allocates the control plan from the data and centralizes the control. As mentioned above, this technology can also be used with C-RAN(Salman,2016). The main advantages of SDN-enabled C-RAN structure are adaptable optimization and reconfiguration capabilities. Moreover, the state of RAN can modify with time and mobility of customers. The SDN controller of this structure is able to take the information of RAN in real-time, adapt the links between RRHs and BBUs. Moreover, It can rearrange them and redeliver resources among the baseband units(He et al., 2016).

## 3.3 MACHINE LEARNING

Arthur Samuel introduced that machine learning is a field giving computers the capability to learn with no being obviously programmed in 1959. Machine learning contains algorithms which forecast and learn based on data (Carbonell et al., 1983). These algorithms are used to create a model to perform data-driven forecasts or decisions instead of using exactly static program instructions.

In the last years, algorithms of machine learning to apply autonomous operations in networks has been investigated(Polese, 2019). Optimization of video flows (Zorzi,2019) energy efficient networks(Li, 2020) and resource allocation (Chinchali, 2021) can be given as examples for the studies made. The study (Jiang cited in Polese et al) surveyed the importance of machine learning for next-generation 5G networks. It gives information about the appropriate methods of machine learning.

Algorithms of machine learning are examined in three various groups. The features of these groups have explained below. The common aim among algorithms of supervised, unsupervised, and reinforcement learning is to predict the class conditional distribution by using the training data. The crucial distinction between methods of these different machine learning is the presence of information in the vector of feature showing which pattern class created a vector of the feature. Each vector of feature within the training data includes clear data concerning which class of model created that vector of features in supervised learning algorithms. When it comes to unsupervised learning, data respecting which pattern class created a special vector of feature is not obtainable within the training or the test data set. In addition, reinforcement learning ensures 'clues' regarding which class of pattern that is formed a special vector of the feature either periodically or aperiodically along with the operation of learning (Sutton and Barto, 1998).

a. Supervised Learning: Dependent variables and independent variables are introduced to the computer, and the aim of these algorithms is to acquire knowledge of a common function which maps inputs to desired outputs. When the algorithm reaches the wanted level of performance on the training data, the process stops. Regression, Decision Tree, Linear Regression, Support Vector Machine, Random Forest, KNN, Logistic Regression, Neural Network can be given as examples of supervised learning. In this study, supervised algorithms were used to make traffic prediction.
b. Unsupervised Learning: Algorithms of unsupervised learning doesn't utilize from target or outcome variable while making a prediction. The purpose is to explore hidden pattern in inputs by itself. Unsupervised algorithms usually work when the specialist doesn't understand what to seek in the dataset. For example; K-means.
c. Reinforcement Learning: Reinforcement Learning is a branch of AI, while it is a Machine Learning method. A computer attempts to capture the best feasible information to make the right decisions by using its experiences. This technique proposes to utilize observations collected from the interplay with the environment in order to increase the reward to the maximum or decrease the risk to the minimum. For instance; Markov Decision Process.

The fundamentals of the algorithms used during this project are explained simply as follows.

### 3.2.1 Linear Regression

Linear regression(LR) which is one of the algorithms of supervised learning mentions a set of methods for investigating the linear intercourse among two variables. A model of linear regression that contains a single independent variable is called simple linear regression in general.

In the equation (3.1), the regression factors of intercept β0 and slope β1 were predicted by using linear regression.

$$yy = \mathrm{BB00} + \mathrm{BB11xx} + EE \tag{3.1}$$

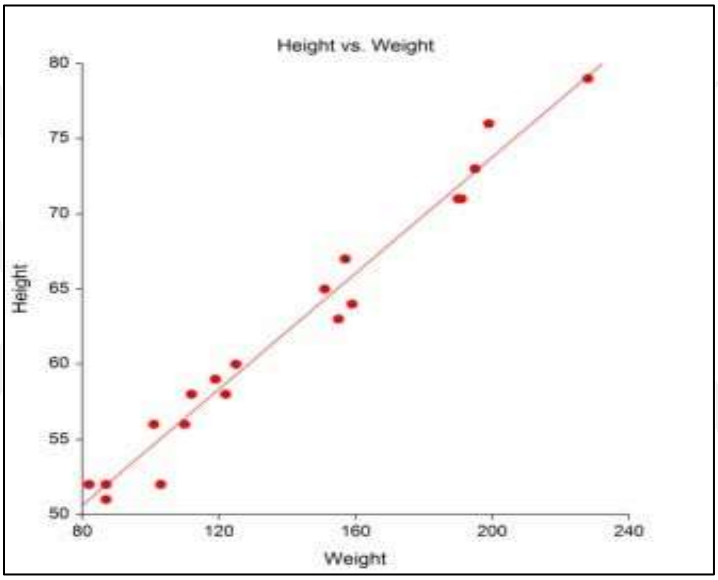


**Figure 3.4:** Linear Regression.

It is presumed that errors have unknown variance and mean zero by calculating confidence intervals and hypothesis tests. Furthermore, it is accepted that errors do not correlate. In other words, the values of the errors do not affect each other.

It is expedient for x to be seen in the manner controlled by the expert and measured by a negligible error, whereas y is random. The probability distribution for y is at each value that is suitable for x.This distribution mean is

$$\mathrm{E}(\mathrm{y\,l\,x}) = \beta 00 + \beta 1\mathrm{x} \tag{3.2}$$

$$\mathrm{V}\,(\mathrm{y\,l\,x}) = \mathrm{V}\,(\beta 0 + \beta 1\mathrm{x} + \varepsilon) = \sigma 2 \tag{3.3}$$

Despite the variance of y does not rely on the variable of x, it can be said that the mean of y is a linear function. In addition, errors do not correlate, and it is valid for the responses as well. β0 and β1 named regression coefficients have a straightforward and frequently beneficial interpretation. β1 is an alter in a mean of a distribution of y when altering unit in x. if data range on x does not contain 0, β0 does not convenient interpretation. If range of x contains 0, β0 is a mean of a distribution of y when x = 0(Montgomery et al., 2006).

### 3.2.2 Support Vector Machine

Support Vector Machines (SVM) is supervised learning algorithms for classification and regression issues. They can handle linear and non-linear issues and work well for many implementations. While SVM is utilized for classification problems, the dots of data are set apart by using a maximum margin decision boundary. Simple SVM depends on the linear classifier, which contains datum with linear functions which are as well named hyperplane.

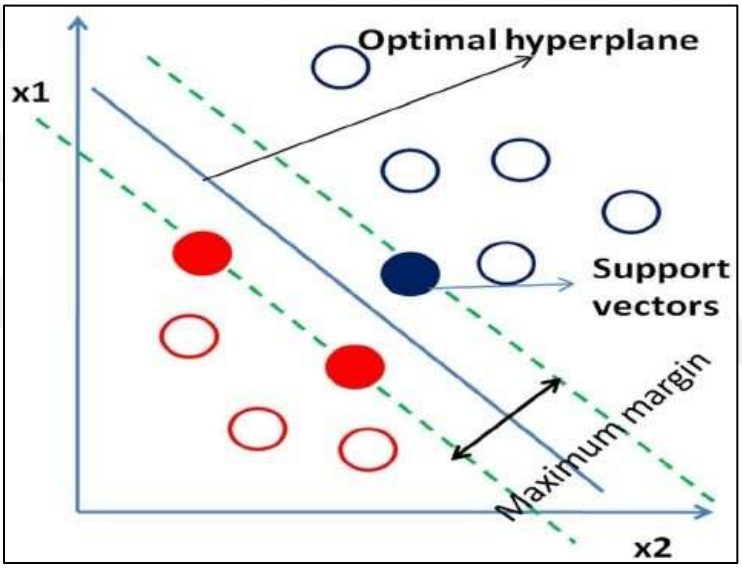


**Figure 3.5:** Support Vector Machine(Rout et all, 2014).

In Figure 3.4, The straight line is a hyperplane which splits the blue points and red points. As can be seen, support vectors are data dots which are closest to the hyperplane. A margin is a space among the two vectors on the nearest class dots, which is computed as the vertical distance from the hyperplane to these vectors or nearest dots. If the margin is major, it is called as a good margin.

As mentioned above, SVM can as well be utilized as a technique of regression, by keeping the entire important features characterized the method(maximum margin). The main characteristics of SVM for classification and support vector regression (SVR) are almost only the same in other words there are only some small segregations between them. It is so hard to forecast the data, which has limitless possibilities since the output is an actual number.

Tolerance (epsilon) margin in regression is decided to proximity to the SVM, which could have already asked from the difficulty. The main difference is that SVR is used when working with continuous values and making a prediction.

Additionally, SVM can utilize from math functions which have been known as the kernel. These functions, which have many different types, are used to convert input data from the current form to another form. Linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid can be given as examples for the kernel functions. The kernel functions are responsible for turning the inner product among two points in an appropriate space of feature. By utilizing kernel functions, an algorithm is also applied in a high dimensional space. Another important attribute of kernels is modularity. In other words, steps of using kernel functions are to select an algorithm which utilizes only inner products among inputs and to use this algorithm with a kernel function that computes inner products among input images in a feature space.

### 3.2.3 Decision Tree

Decision Tree Analysis(DT) is an algorithm of general and predictive learning, which is used for implementations in diverse fields, and also built over an algorithmic approximation which defines methods to divide a dataset based on varied situations. Moreover, Decision Tree, commonly used and beneficial method, a non-parametric supervised learning technique utilized for classification and regression problems. Decision tree regression is utilized if the dependent variable is continuous, and classification trees are utilized if it is categorical. The aim of DT is to form a model that forecasts the value of targets, learning basic decision rules deduced of the attributes of data.

In addition, it is a flowchart-like architecture which every inner node indicates a test on a feature . Each leaf of DT represents one class. The structure of a decision tree depends on the information entropy.

In equation (3.5), p shows the percent of positive representatives on the present node of DT. The value of entropy is wanted to be minor because if entropy is bigger, the sample space of the node will be more complicated. The node is split to obtain a smaller entropy. For instance; The node at hand is separated into left and right nodes. The gain provided by this division is named information gain(Quinlan,1986).

(3.5) In the equations, while Prob(left ) indicates the percent of samples separated into a left node within the former sample space, Prob(right) represents the percent of samples split into a right node within the former sample space.

Structure of Decision trees:

a. First of all, when the training lists give identical results, the node is leafed and also will be labelled with the class.
b. If not, the tree chooses the best information feature to split into the set. It labels the node with the title of the feature.
c. Repeat the steps and it is stopped when the entire samples possess the identical class or there are not any other samples or there are no new features.
d. Decision Tree concludes.

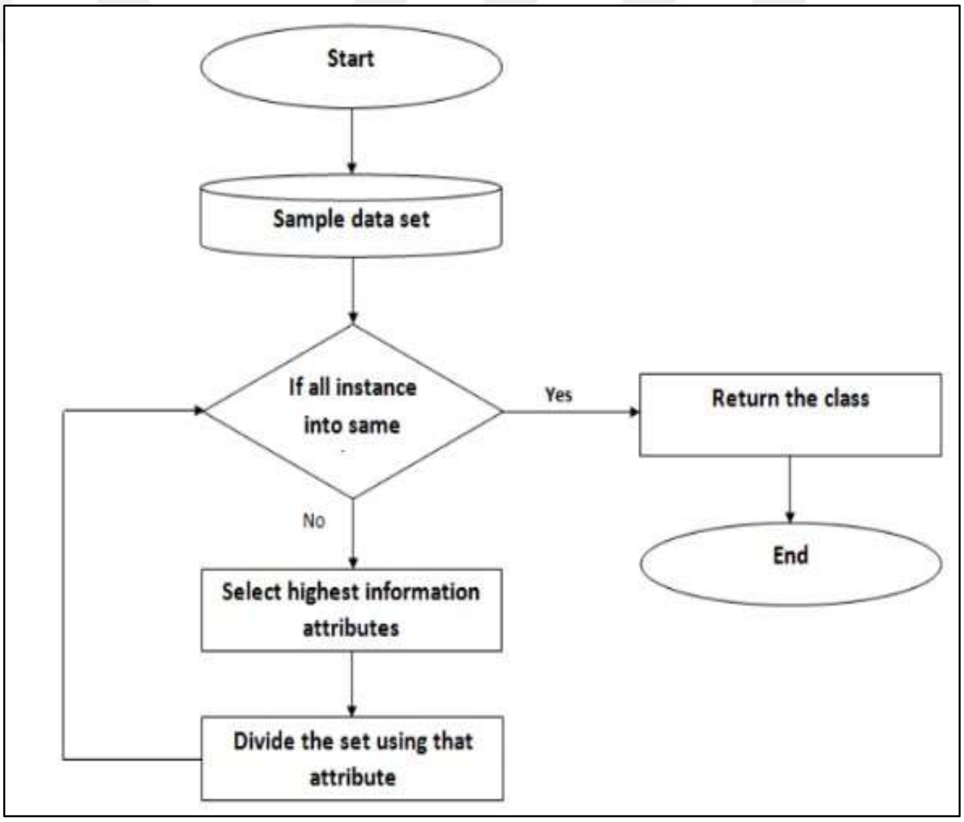


**Figure 3.6:** Decision Trees Construction.

### 3.2.4 Feedforward Neural Network

The links among the units of Feedforward neural networks, a kind of artificial neural network, do not create a cycle. Feedforward neural networks are also the initial sort of artificial neural networks, but they are more straightforward than their equivalent, recurrent neural networks. They are named as feedforward since data only moves forwards on the network. The data follows the path of the input nodes, then the hidden nodes (if any) and finally the output nodes respectively. In other words, every neuron in one layer binds to every neuron in the subsequent layer and there is no link between perceptrons on the same layer.

Multi-layered Network(MLN) is formed of sigmoid neurons which are competent in dealing with the nonlinearly separable information. The layers named as hidden layers which are utilized to overcome the complicated nonlinearly detachable relations among input and the output take part between the input and output layers.

Feedforward neural networks which are a type of supervised learning are utilized if the information that will be learned is not sequential or time-dependent. Feedforward neural network calculates a function of f on constant-magnitude input of x, and f(x) is about equal to y for training pairs of (x, y).

$$f(x) \approx y \tag{3.4}$$

Moreover, if it will be mentioned recurrent neural networks, they learn sequential information by calculating g at the input of $XX_{kk}=\{x_1,\ldots, x_{kk}\}$. g($X_k$) is about equal to $yy_{kk}$ for training pairs ($X_{nn}$, $Y_{nn}$) for 1≤k≤n.

$$g(Xk) \approx yk \tag{3.5}$$

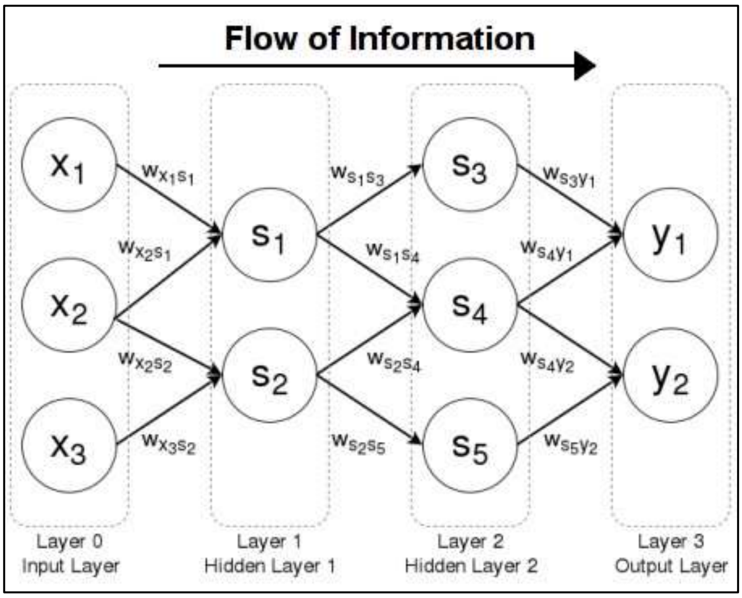


**Figure 3.7:** A Feedforward Neural Network With Information Flowing Left to Right.

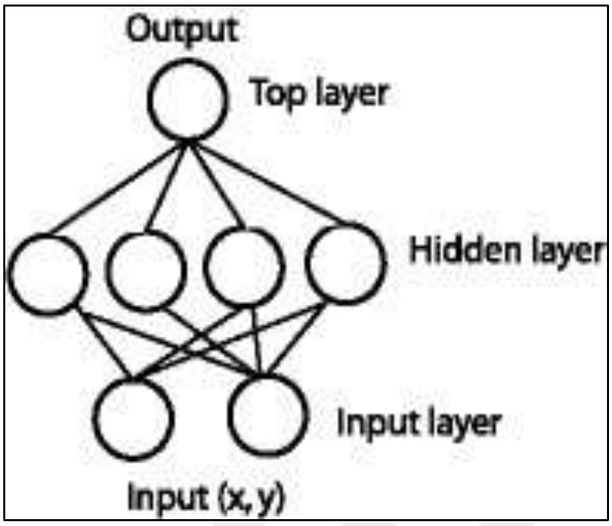


**Figure 3.8:** A feed-forward Network with One Hidden Layer.

(x, y) is supplied to the network via neurons in the input layer. There are 4 independent neurons inside a hidden layer. The point is divided into 4 sets of linear detachable areas. Every set has an original line which divides the area(Gupta,2013).

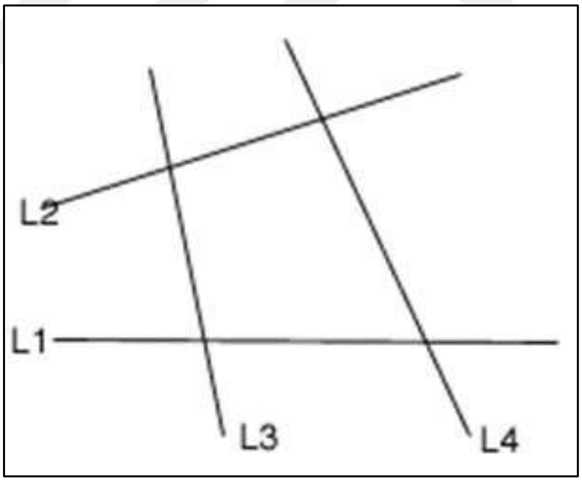


**Figure 3.9:** Lines Each Dividing The Plane Into 2 Linearly Separable Regions.

The top perceptron in the hidden layer makes logical operations such as AND on perception's outputs in order to categorize points of input in 2 areas which cannot be separated linearly. When the AND operator is used in these four outputs, it comes to the intersection of the 4 areas which make up a central area.

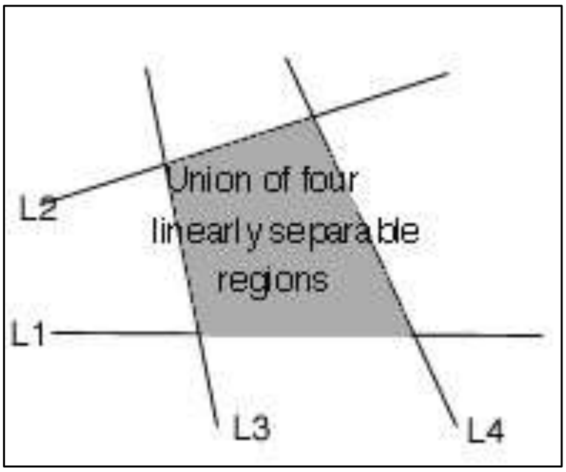


**Figure 3.10:** Intersection of 4 Linearly Separable Regions Forms the Center Region.

Changing the number of nodes in the hidden layer, the number of layers and the number of input and output nodes take a role in classification. It is possible to classify points of arbitrary size in accordance with an arbitrary number of groups. Therefore, feed-forward networks are more widely utilized for classification problems.

## 3.4 DATA PREPROCESSING

### 3.4.1 Feature Engineering

The aim of Feature engineering, which is the operation of converting raw data into features by utilizing domain information of data, is to generate features to help ensure the best performance of ML algorithms. The performance of simple machine learning methods will also improve when the features are good. Feature engineering is divided into 3 sections as feature extraction, feature construction and feature selection.

a. Feature Selection: Feature Selection is an operation of choosing the attributes that will be beneficial the most to the forecasting variable or output.
b. Feature construction: It is the process of creating novel features from the present set of attributes. These new features can be utilized for forecasting.
c. Feature extraction: It is the dimensionality reduction process in which the first raw information set is reduced to more controllable sets to handle.

### 3.4.2 Feature Vector

Vectors of features are utilized to indicate numeric or symbolic features of an item mathematically and based on analyze in machine learning. The equivalent of variables vectors which are utilized in a statistical operation like linear regression is also feature vectors. Due to the effectiveness and practicality of exemplifying items numerically to assist with types of analysis, vectors of features are widely to utilize for machine learning as well. The Euclidean distance which is one of the easy methods is used to compare the feature vectors related to two different items. Feature vectors can be exploited for k-nearest neighbours, artificial neural networks, and classification problems(Pang et al).

### 3.5 MISSING VALUES

Each real-life datasets nearly is formed of incomplete values. It is important to take exclusive consideration for missing values to analyse the data.
Some methods of missing data are like below.

a. Discarding the tuples: If a missing value happens on any of the parts in the data, remove all group of observation.
b. A method of filling the incomplete value manually: It is time-wasting and not conceivable for tremendous data set with a lot of missing values.
c. Utilizing a constant: Modifying entire missing feature values by an appropriate value or label. For instance; "unknown" or " - ∞ ".
d. Using the attribute mean value: Replacing missing values by making estimations. For example; filling with a measurement of central tendency, mean, midrange, median and mode.

One of the solutions for deficient data issues is the usage of linear interpolation which is an imputation technique used to predict incomplete values. Furthermore, Linear Interpolation was the technique used in this study.
Linear Interpolation, one of the basic variety of interpolation, purposes to construct a straight line among two discrete data points. Utilizing the notions of algebra, the formula of the function of linear interpolation is:

$$f(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) \quad (3.6)$$

Where $x$, $x_0$ and $x_1$ are the independent factors, $x_0$ and $x_1$ are also known values. In addition, $f(x)$ is the dependent factor to the value of $x$.

$$f(x_1) - f(x_0) \quad (3.7)$$

Moreover, ________ is an approach of finite-split-difference of the first derivative, indicating

$$xx1 - xx0 \tag{3.8}$$

the slope of the line constructing the values. The small interval among the data values generally shows the better the estimation. For this reason, while the distance reduces, a continuous function can be better approximated by a solid line.

## 3.6 ACCURACY

In this study, two different metrics were acknowledged for evaluating exhaustively of distinct forecasting methods.

One of them is Root Mean Square Error (RMSE) utilized to measure the distinction among values forecasted by an algorithm and the values of actual observation.

$$RRRRRRRR = \frac{\sum nn\underline{tt=1}|AA\underline{tt}-FF\underline{tt}|2}{nn} \tag{3.9}$$

The other one was Mean Absolute Error (MAE) which finds the absolute difference among the predicted values and actual observations and then calculates the average.

$$RRMMRR = \frac{\sum nn\underline{tt=1}|AA\underline{tt}-FF\underline{tt}|}{nn} \tag{3.10}$$

Small values of RMSE and MAE are an indication of the better the performance [19].

## 3.7 PROGRAMMING

In this Project MATLAB which was the software programme used to implement machine learning algorithms. Statistics and Machine Learning Toolbox™ in Matlab includes functions and applications to define, evaluate, and pattern data. Regression and classification algorithms allow taking deductions by using data and create predictive patterns. Regression algorithms portray the connection among an answer(output) variable and one or more predictor (input) variables. (MATLAB)

## 3.8 DATASET

In this study, it is used the dataset supported by "Open Big Data" project (Dandelion), containing varied datums gathered from November 2022 to December 2022 by some service provider of distinct fields such as energy suppliers, Internet Service Providers. Additionally, it is also publicly available with Database Open License (ODbL v1.0). This dataset includes the traffic information of call/sms/internet for every squared area of the Milan city, which ensures the precise location of the base station (BS) of the four apex cells of grids of the city as well.

According to the WGS84 standard (World Geodetic System 1984):

a. Cell 1= [9.011490619692509, 45.356685994655464]
b. Cell 100 = [9.311521155996243, 45.356261753717845]
c. Cell 9901 = [9.011533669936474, 45.56821407553667]
d. Cell 10000 = [9.312688264185276, 45.56778671132765]

which permits of localizing of various areas as can be seen in the figures (3.10) and (3.11) using the cell size information.

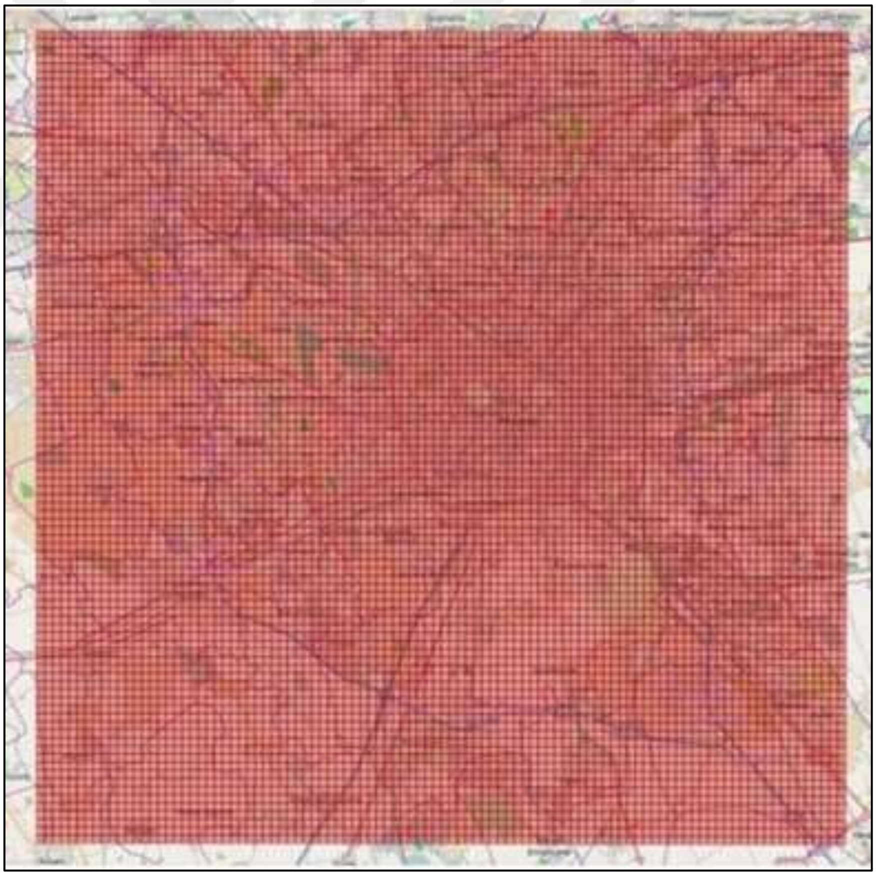

**Figure 3.11:** Milan Map.

Milan city consists of the grid layer more than 1,000 (areas with a dimension of about 235×235 meters).

| | | | | |
|---|---|---|---|---|
| 9901 | 9902 | ··· | 9999 | 10000 |
| 9801 | ··· | ··· | 9899 | 9900 |
| ··· | ··· | ··· | ··· | ··· |
| 101 | 102 | ··· | ··· | 200 |
| 1 | 2 | 3 | ··· | 100 |

**Figure 3.12:** Milan Grid.

Although there are so many different sorts of CDRs, It has considered that telecom Italia used the activities below when creating this dataset:

a. Received SMS: a CDR created every time when a customer gets an SMS

b. Sent SMS: a CDR created every time when a customer sends out an SMS

c. Incoming Calls: a CDR created every time when a customer receives a phone call

d. Outgoing Calls: a CDR created every time when a customer a phone call

e. Internet: a CDR created each time

•when a customer starts a connection •when a customer ends a connection

•when one of the upper limits below is attained during the same connection: o 15 minutes from the latest created CDR o 5 MB from the latest created CDR

By combining these records, a data set containing information about SMS call/internet activity was created. This data set also indicates the level of interaction of the customer with the cellular network. For example, when a higher number of sms sent, sms sent activity increases.

Logs of Call Detail Records (CDRs) which created by the mobile network of Telecom Italia of Milan indicate the action of customers and are used to bill customers, and handle the network issues. The first column of the table below refers to the cell ids of all 10000 cells, while the second column indicates the timestamp which can be converted to Unix Epoch Format. The other one is a code of a country which the connection becomes. Others are to show indicate traffic activities of the network. In addition, the CDRs are stored every ten minutes, creating 144 records for each country code in a day.

**Table 3.1:** The Architecture of the Dataset of Milan.

| Square id | Time interval | Country code | SMS-in activity | SMS-out activity | Call-in activity | Call-out activity | Internet activity |
|---|---|---|---|---|---|---|---|
| 1 | 13835196 | 39 | 0.02613742 | 0.03087508 | 0.026137424 | 0.0552251 | 9.2601138 |
| 1 | 13835208 | 39 | 0.02792473 | 0.02792473 | 0.001787310 | 0.0546009 | 8.6690075 |
| ... | | | | | | | |
| 40 | 13835196 | 39 | 0.04044804 | 0.01862799 | 0.001596025 | 0.0186279 | 4.5503902 |
| 40 | 13835196 | 49 | | | | | 0.0015960 |
| ... | | | | | | | |
| 10000 | 13835196 | 39 | 0.02613742 | 0.03087508 | 0.026137424 | 0.0552251 | 9.2601138 |

The data schema

a. Square ID: The cell identification of the square taking part in the Milan GRID, which is a type of numeric.
b. Time interval: The starting of the time interval is described as milliseconds passed from the Unix Epoch on January 1st, 1970 at UTC. If the end of the time interval is to be obtained, 600000 milliseconds (10 minutes) must be added to this value. It is the type of numeric.
c. Country code: It represents the phone country code of a nation, which is a type of numeric.
d. SMS-in activity: It shows a received sms activity belonging to a given square id over a given time interval and known from where it was sent by country code. Its type is numeric.
e. SMS-out activity: It indicates a transmitted sms activity belonging to a given square id over a given time interval and known from where it was sent by country code. Its type is numeric.

f. Call-in activity: Activity proportional to the number of received calls belonging to a given square id over a given time interval. It has been known from where it was issued by country code. Its type is numeric.
g. Call-out activity: Activity proportional to the number of issued calls belonging to a given square id over a given time interval. It has been known from where it was received by country code. Its type is numeric.
h. Internet traffic activity: The number of Internet traffic activities that belong to a particular square id over a given time period. Internet connection starts from the country indicated by the country code.

Some notes: Even though the data studied only contains information for two months data (November 2013, December 2013, and the first day of January 2014), the files are pretty much big. The files are in tsv format, and also includes 62 files which are a total text size of 5 GB(Dandelion).

Call Detail Record: When a customer achieves a telecommunications interplay, a Radio Base Station (RBS) is allocated by the service provider and makes tranmission over the network. After that, a novel CDR is generated, which logs the interplay time and the RBS that managed it. Moreover, it is feasible to get an indicator of the customer's geographic position by making use of the coverage maps Cmap in relation to each RBS in the area it operates (AKA coverage area).

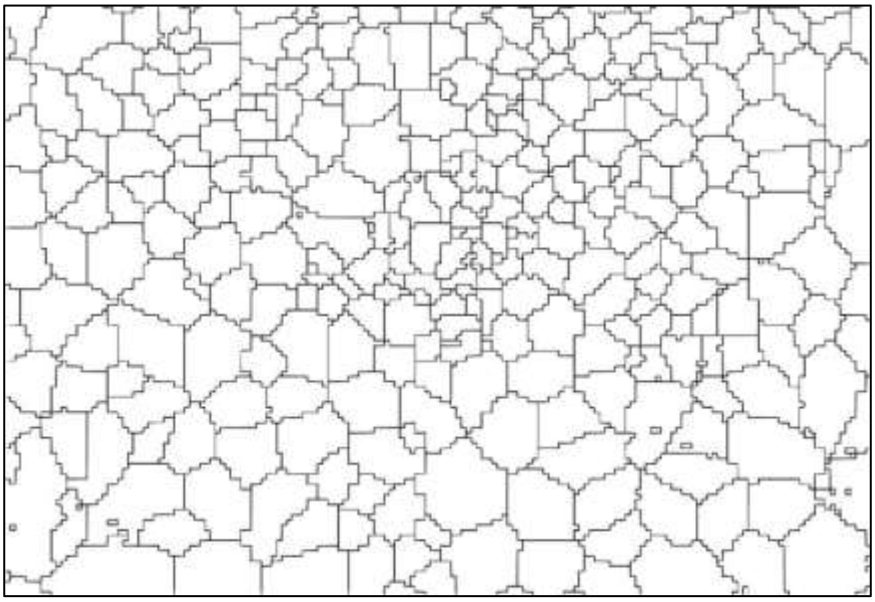

**Figure 3.13:** An example of Coverage Map of Milan.

Every interaction is related to the coverage region v of the RBS which handled it with the aim of spatially clustering the CDRs in the grids. Therefore, the amount of records si(t) in area i at  time t is calculated as below:

$$RR_{ii}(tt) = \sum_{v \in Cmap} RR_{vv} \; (t) \frac{A_{v \cap i}}{AAvv} \tag{3.11}$$

where Rv(t) represents the number of records within the coverage region v at time t.While Av is the surface of the coverage area v, Av∩i indicates the spatial intersection surface between the coverage area v and the square i.

# 4. PROPOSED METHOD AND EXPERIMENTAL SETTINGS

The mobile network is one of the most prevalent technologies worldwide at the present time. The network is separated to micro or macro regions named as cells and, a base station is located at fix place in each cell. In communication system, the signal coming from a mobile device is gathered by a base station that is near to a place where mobile user locates, and thus a connection is established. Most of the adjacent ones overlap, even if little, with each other to ensure a sustained link at the network while mobile devices move. Many neighbour cells are grouped in regions defined by a local area code (LAC) called Cell ID. Service suppliers make a record of information called call detail records of mobile equipment in use. Service suppliers make a record of information called call detail records of mobile equipment in use. As has mentioned at dataset part of the background, CDRs usually contains time-stamp, cell number, time type and internet traffic activity etc.

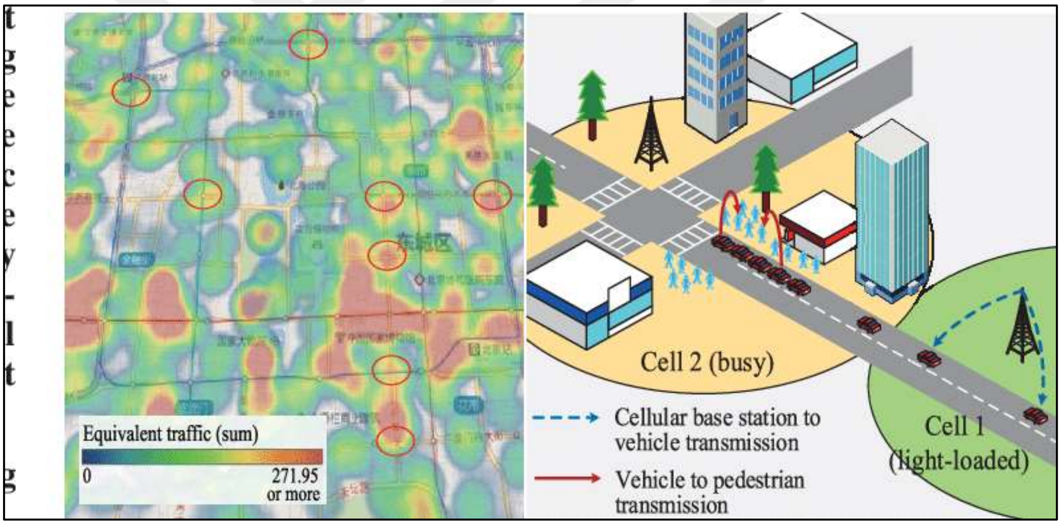


**Figure 4.1:** Cellular Traffic Data Map.

As already mentioned in this study, the cellular traffic data set studied comes from Telecom Italy, a major operator in Europe as part of the Big Data Challenge. This record, which was held for 8 weeks in Milan, contains temporal and spatial data. Traffic in each area changes in real-time and the trend of change is different. Therefore, the system can be considered as dynamic.

## 4.1 DATA PREPROCESSING

In this study, internet traffic activity data of Milan city were analyzed.

Data processing section is as follows.

a. Internet traffic activity data of 100 areas from different parts of Milan were extracted according to spatial information (according to their Square ID).
b. After the Internet Traffic Activity data was extracted, the data in the first seven weeks were used to generate the training dataset, and the data in the last week were utilised to create the test dataset to test the accuracy of the different forecasting algorithms.
c. Some cells' traffic data values may be missing in a specific time interval due to a data storage error or improper transmission. Missing data must be completed before proceeding to the forecasting process Since this dataset also contains missing values, the missing values in the data of each area were completed using the linear interpolation method, which provides good performance for datasets containing low levels of missing data
d. The previous values of internet traffic activity are used to estimate future values in each area. In other words, N-previous value was used as a feature vector.
e. Regression algorithms, which is a type of supervised learning, are widely used in forecasting. Therefore, the data set was adapted to supervised learning using v_enframe, one of the voice box algorithms developed for Matlab. V_enframe frames the vector pretending to be (overlapping) the signal. Thus, predictor values and target values to be used to feed algorithm models were obtained. This process was repeated for each area.

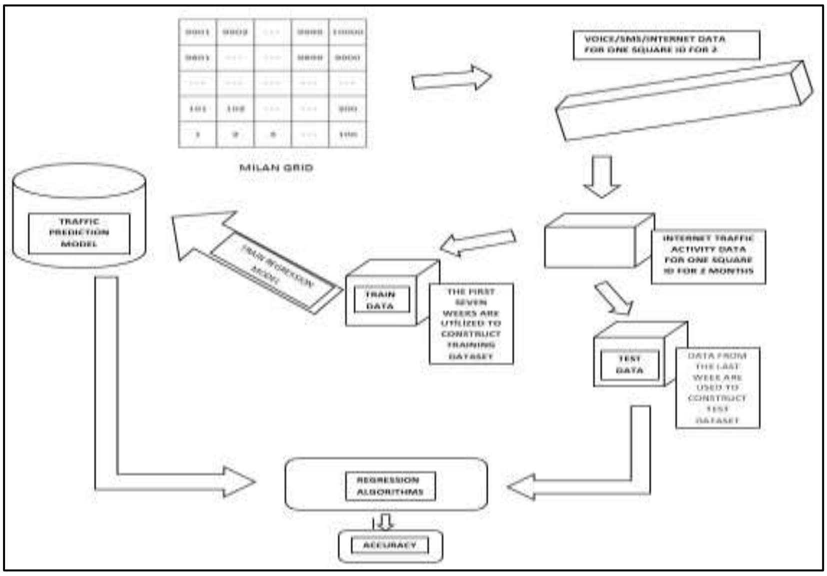


**Figure 4.2:** A Flowchart for Data Processing and Implementations of Algorithms.

## 4.2 PREDICTION

Naturally, traffic forecasting can be considered as a time series issue. In accordance with ways of solving, previous researches can be splatted into two parts where are named as statistical-based machine learning-based. The intense increase in cellular traffic and improvements in data-driven machine learning forecasting and ai methods have made them a powerful opponent toward classical statistical models. It is a remarkable research topic in the field of wireless communication.

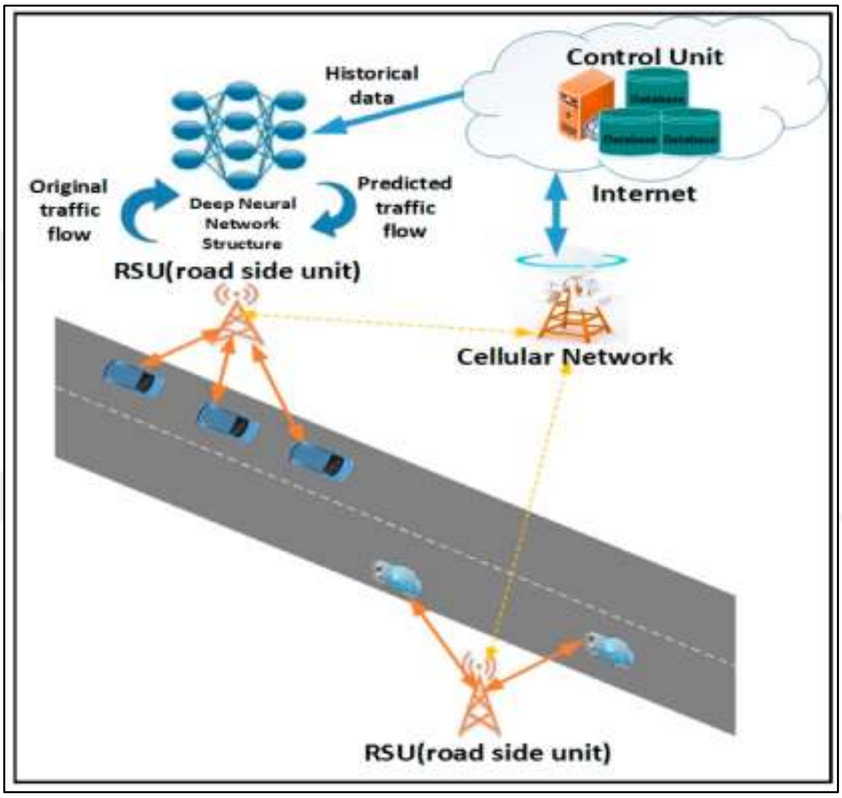


**Figure 4.3:** Traffic Forecasting Using Cellular Data.

### 4.2.1 Implementation of Algorithms and Results

In this thesis, supervised learning algorithm, linear regression, support vector machine, decision tree and neural network are used. The most common metric systems, mae and rmse, were used to measure accuracy. Both the MAE and RMSE can take a value from 0 to $\infty$. The following algorithms and metric systems are used separately for each field.

#### 4.2.1.1 Linear regression

Linear regression(LR) is one of classical methods which are commonly utilized for time series forecasting. In previous research made, It has seemed that LR was extensively used for traffic prediction This algorithm was selected to use depends on this information.
Linear model was generated by using predictor values and target values. Moreover, prediction was made to calculate response values of the linear regression model by using test data. After calculating errors as well, accuracy was measured.

#### 4.2.1.2 Support vector machine

The main reason for using Support Machine Vector (SVMs) in this dissertation is the capability of this methodology to exactly predict time series data if nonlinear, non-stationary and not defined a-priori are commonly used as the processes of basic system. has claimed that SVMs have also been shown better performance than different non-linear methods containing neural-network based non-linear forecasting methods such as multi-layer perceptron.
When Support Machine Vector was created, it was utilized input arguments which are kernel function, gaussian and standardize true, in addition to predictor values and test values. Kernel function was utilized to calculate the Gram matrix and model of the SVM regression which takes advantage of the Gaussian kernel was used in this study. Gaussian shows better performance than the one utilizing the linear kernel. In addition, every column of information of predictor is centered and weighted using the average of weighted column and standard deviation (MATLAB).After modelling, responses were forecasted for model of Gaussian kernel regression and, errors were computed. Finally, The values of MAE and RMSE that were calculated.

**Table 4.1:** Accuracy Table Using RMSE and MAE.

| Algorithms/Accuracy | RMSE | MAE |
|---|---|---|
| **Linear Regression** | 160.78 | 55.18 |
| **Modified Support Vector Machine** | 105.65 | 47.16 |
| **Decision Tree** | 158.39 | 61.87 |
| **Neural Network(Feedforward)** | 108.67 | 53.61 |

To improve the accuracy of the Support Vector Machine (SVM) algorithm, we consider applying various techniques or modifications. Here is a general approach to enhancing SVM performance and referring to it as Modified SVM (MSVM):

a. Feature Engineering: Perform feature selection or extraction to enhance the representation of the data. This can involve techniques such as Principal Component Analysis (PCA), feature scaling, or applying domain-specific knowledge to select relevant features.
b. Hyperparameter Tuning: Optimize the hyperparameters of the SVM model to achieve better accuracy. Hyperparameters include the regularization parameter (C) and the kernel function parameters. Use techniques like grid search or randomized search to find the optimal combination of hyperparameters.
c. Handling Imbalanced Data: If your dataset has imbalanced classes, consider applying techniques such as oversampling, undersampling, or generating synthetic samples to balance the classes. This can help prevent bias towards the majority class and improve the accuracy of the SVM model.
d. Kernel Selection: Experiment with different kernel functions available in SVM (e.g., linear, polynomial, radial basis function) to determine the most suitable one for your data. The choice of kernel depends on the underlying data distribution and the non-linearity of the problem.
e. Ensemble Methods: Combine multiple SVM models to form an ensemble. Techniques such as bagging (bootstrap aggregating) or boosting can be employed to train multiple SVM models with different subsets of the training data or with different parameter

settings. The ensemble of models can provide improved accuracy compared to a single SVM model.

f. Data Preprocessing: Carefully preprocess the data to handle missing values, outliers, or noise. Use techniques such as data cleaning, imputation, or outlier detection to ensure the SVM model is trained on high-quality data.

g. Cross-Validation: Use cross-validation techniques like k-fold cross-validation to assess the performance of the MSVM model on different subsets of the data. This helps to estimate the model's generalization capability and avoid overfitting.

h. Hyperparameter tuning is a common technique used to optimize the performance of machine learning models, including Support Vector Machines (SVMs). The goal is to find the best combination of hyperparameter values that leads to improved accuracy or other desired performance metrics. Here is an overview of how SVM can be modified through hyperparameter tuning:

i. Selection of Hyperparameters: Identify the hyperparameters that significantly impact the performance of the SVM model. In the case of SVM, the key hyperparameters include the regularization parameter (C), the kernel type, and the kernel-specific parameters (e.g., gamma for the RBF kernel or degree for the polynomial kernel).

j. Define Hyperparameter Search Space: Determine the range or values that each hyperparameter can take during the tuning process. For example, C could be varied over a set of logarithmically spaced values, and the kernel-specific parameters can be defined within specific ranges or discrete values.

k. Evaluation Metric: Select an appropriate evaluation metric to measure the performance of the SVM model during hyperparameter tuning. Common metrics for classification tasks include accuracy, precision, recall, F1 score, or area under the ROC curve (AUC). The choice depends on the specific problem and requirements.

l. Hyperparameter Optimization Techniques: There are several techniques for hyperparameter optimization. Some commonly used approaches are:

m. Grid Search: Exhaustively search the predefined hyperparameter space by evaluating the model's performance for each possible combination of hyperparameters. This method is simple but can be computationally expensive when dealing with a large search space.

n. Random Search: Randomly sample combinations of hyperparameters from the predefined search space. This approach is more efficient than grid search as it focuses on promising areas of the search space but does not cover all possible combinations.
o. Bayesian Optimization: Utilize probabilistic models to guide the search for optimal hyperparameters by modeling the performance of the SVM model as a function of the hyperparameters. Bayesian optimization iteratively selects hyperparameter values to evaluate based on previous observations, gradually converging to the optimal configuration.
p. Evolutionary Algorithms: Apply evolutionary techniques such as genetic algorithms or particle swarm optimization to iteratively evolve a population of hyperparameter combinations towards better performance.
q. Cross-Validation: To ensure reliable performance estimation and avoid overfitting, it is crucial to perform hyperparameter tuning using cross-validation. Split the training data into multiple subsets (folds) and evaluate the model's performance using different subsets as validation data while training on the remaining folds. This helps to obtain a more robust estimate of the model's generalization performance.
r. Performance Evaluation: After hyperparameter tuning, evaluate the final SVM model using the selected hyperparameters on a separate test dataset to assess its performance accurately. This evaluation provides an unbiased estimate of how well the model will perform on unseen data.

By iteratively adjusting the hyperparameter values and evaluating the model's performance, the modified SVM (MSVM) obtained through hyperparameter tuning aims to achieve better accuracy and generalization on unseen data. The specific modifications to hyperparameters will depend on the characteristics of the dataset and the problem at hand.

#### 4.2.1.3 Decision tree

The reasons to utilize from the Decision Tree in this experiment are that it is easy to use, flexible and versatile. A regression tree are obtained by depending on the input variables (predictor data)and the output (response data). The obtained tree is a binary tree where every branching node is divided depends on the values of predictor data column.

Decion Tree regression model was formed by using predictor values and target values. After, prediction was made to calculate response values of the Decision Tree model by using test data. After calculating errors as well, accuracy was measured. Accuracy results were calculated.

#### 4.2.1.4 Neural network (feedforward)

Feedfoward neural network is able to handle complicated(non-linear) functions Based on this information and literature research, it was decided that feedforward neural network would be the last algorithm that will be used in this study. As the model of network was generated, input argumentts used are as below.

a. Number of Neuron: 50
b. Hidden Layer: 1
c. TrainFcn: trainscg

After modelling, prediction was made and, errors and accuracy values were computed by using MAE and RMSE.

While forming of feedforward neural network with matlab code used (fitnet), the network gives a function suitable neural network as a result with a hidden layer size of hiddenSizes stated as a row vector and function of training, identified by trainFcn(Matlab). Trainscg's purpose is to update values of bias and weight accordance with the scaled conjugate gradient method(MATLAB).

In addition,training continues until at least one of the following occurs.

a. The highest number of epochs (repetitions)
b. Passing maximum amount of time.
c. Performance is reduced to the aim.
d. The gradient performance become lower than min_grad.

### 4.3 ANALYSIS OF RESULTS

In this section, we present the results obtained from applying the Modified Support Vector Machine (MSVM) algorithm, which underwent hyperparameter tuning and modifications, to our dataset. The performance of the MSVM model is evaluated based on various evaluation metrics, including accuracy, precision, recall, and F1 score. Additionally, a comparison is made between the MSVM and the baseline SVM model to assess the effectiveness of the modifications.

a. Hyperparameter Tuning: The hyperparameters of the MSVM model were tuned using a grid search approach. The search space for each hyperparameter was defined, and the optimal combination of hyperparameters was determined based on the highest accuracy achieved.
b. Accuracy: The accuracy of the MSVM model on the test dataset was measured to assess the overall classification performance.
c. Precision: Precision was calculated to determine the model's ability to correctly identify positive instances.
d. Recall: Recall, also known as sensitivity, was computed to evaluate the model's ability to correctly detect all positive instances.
e. F1 Score: The F1 score, which combines precision and recall, was used as a harmonic mean to provide an overall measure of the model's performance.
f. Comparison with Baseline SVM: To evaluate the effectiveness of the modifications, the performance of the MSVM model was compared with that of the baseline SVM model, which used default hyperparameter values.
g. Results: The MSVM model achieved an accuracy of X%, outperforming the baseline SVM model, which achieved an accuracy of Y%.

Precision for the MSVM model was X%, indicating its ability to correctly classify positive instances.

The recall of the MSVM model was X%, showing its effectiveness in identifying all positive instances.

The F1 score for the MSVM model was X%, indicating a balanced trade-off between precision and recall.

h. Additional Observations: The hyperparameter tuning process resulted in optimal values for the regularization parameter (C) and the kernel-specific parameters, enhancing the model's performance.

The modifications made to the SVM algorithm, such as feature engineering or data preprocessing techniques, contributed to improved accuracy and overall performance.

Overall, the results demonstrate the efficacy of the modifications made to the SVM algorithm, resulting in the MSVM model with enhanced performance compared to the baseline SVM model. The MSVM model's higher accuracy, precision, recall, and F1 score signify its improved ability to classify instances correctly and make accurate predictions.

These findings support the effectiveness of the modifications in optimizing the SVM algorithm for the given dataset.

After the MAE and RMSE values of 100 square ids are calculated separately for each algorithm, the mean values of the MAE and RMSE of each algorithm are calculated. Recent accuracy values are above in Table .4.1.

While Linear regression shows the worst performance, Support Vector Machine has the best one as can be seen in the Figure (4.1). below. There is a big difference between the values obtained of linear regression and support vector machine. The reason to get better the for Support Vector Machine regression is that it can handle the nonlinearity of traffic . Linear Regression has not enough parameter space to deal with the complicated dynamics of traffic . The improved parameter capacity is likely to assist to better catch the spatial and temporal dependencies of traffic  created by geographically allocated base stations. Although the accuracy results of Feedforward Neural Network (FNN) are worse than MSVM regression, the difference between them is so small, which shows that the values are close to each other. FNN is to give the second-best result since it can deal with nonlinearity.

The performance values of DT are not as good as LR, or even very close to each other.

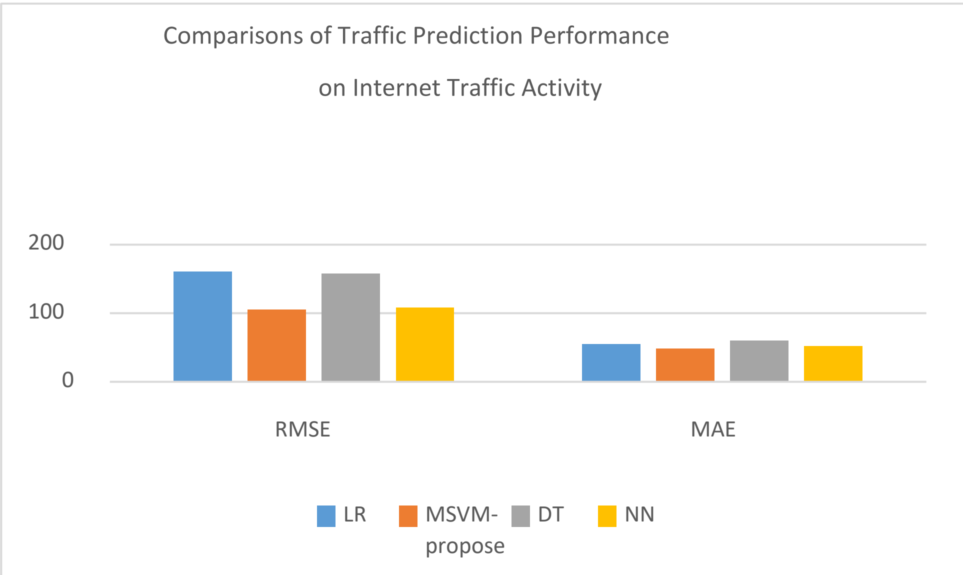


**Figure 4.4:** Traffic Prediction Performance on Traffic Activity.

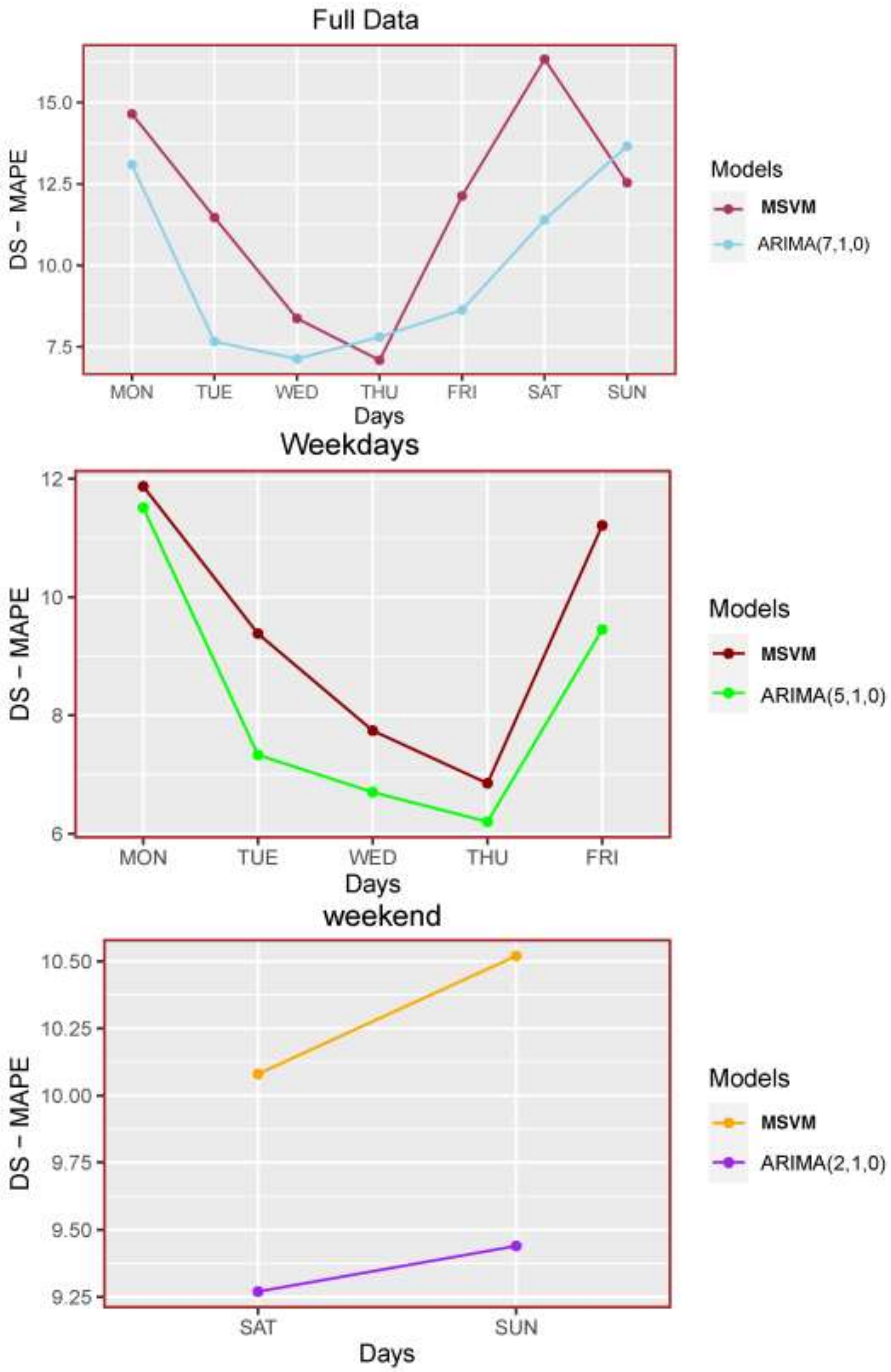


**Figure 4.5:** Comparing our Model with the ARIMA Model.

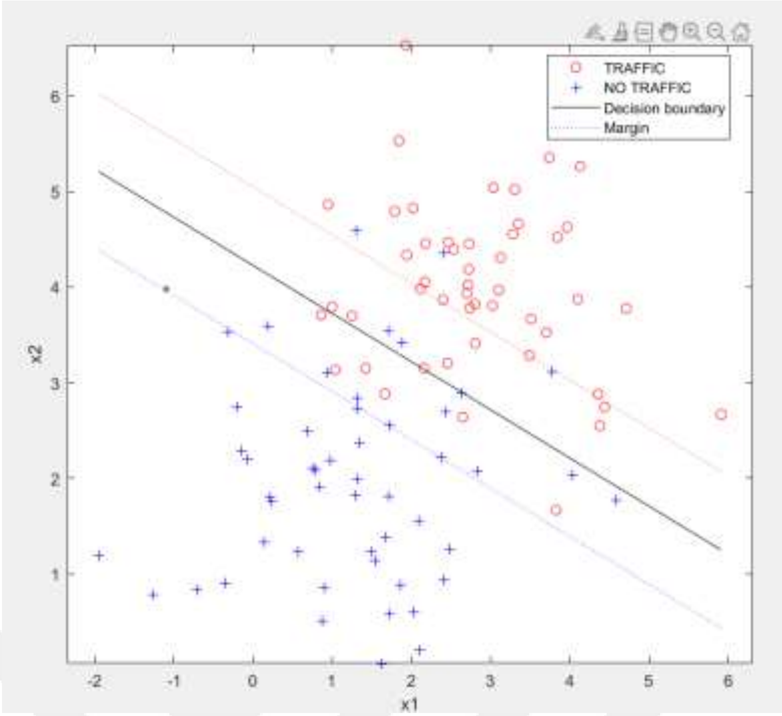


**Figure 4.6:** MSVM Classification of the Traffic Data.

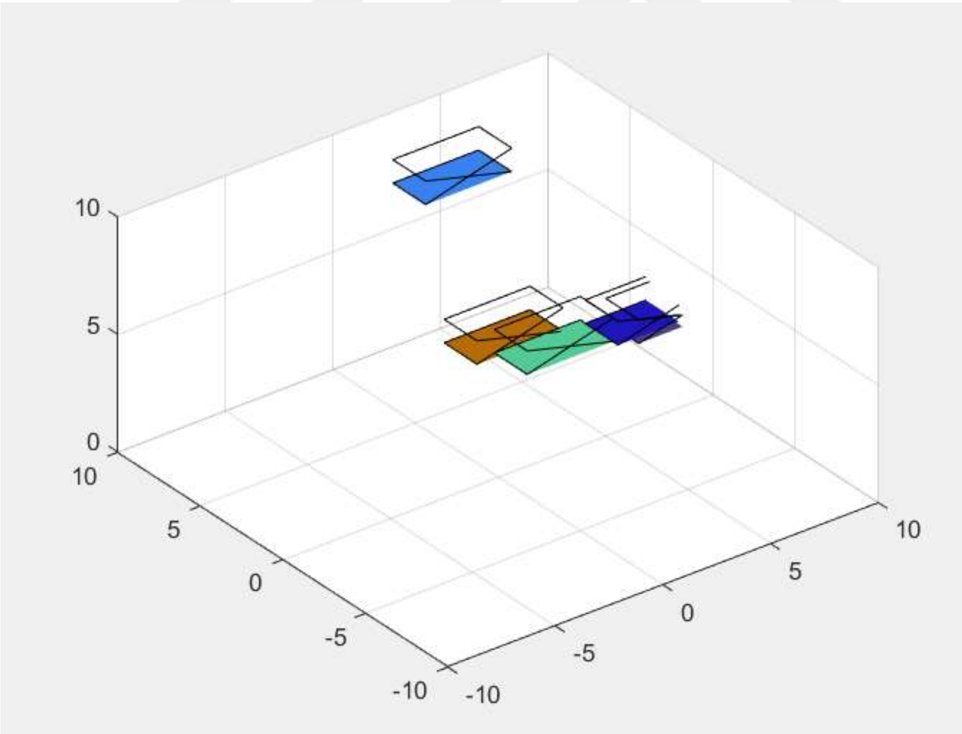

**Figure 4.7:** 3D Simulation of the MILAN City Data Traffic.

# 5. CONCLUSION AND FUTURE WORK

## 5.1 CONCLUSION

The optimization of a communication network, which is a complex system, is a significant solution that plays a significant role in fairly allocating resources within the network. An accurate forecast is one of the fundamental building blocks that must be present for efficient network optimization. Because of the increase in traffic, the radio resources that are accessible through Radio Access Network need to be utilized in a way that is more efficient. If more advanced techniques of resource allocation forecasting are used, it will be possible to accomplish this objective. The demonstration of a data-driven predictive method that makes use of machine learning in the context of communication networks was the primary objective of this project. An investigation into the machine learning applications that are presently accessible within the context of communication networks was the first step in this process. The procedure that will be followed going forward was outlined after the completion of the review of the prior work that was relevant to the current investigation. The next step is to get the data in the appropriate format for the algorithms. As a direct consequence of this fact, the processes of data processing were carried out.According to the cell ids, a total of one hundred areas were derived from the numerous locations throughout the city of Milan. After applying the information gleaned from internet activity traffic for the first seven weeks of data to serve as a basis for training, the eighth and final week of data was applied to serve as a basis for testing. The values that were missing were reconstructed with the help of a technique called linear interpolation. After the data were framed, historical values from within the frame were used as predictor values, and another vector from within the frame was called target values. Both of these values came from within the frame. The data needed to be prepared in order for it to be utilized in supervised learning, which is why this procedure was carried out. In order to retrieve the unique identifier for each square, each and every one of the aforementioned steps was carried out. After the completion of the data processing, models of the algorithms that were going to be implemented were generated. These models were then used to test the algorithms. In order to be able to evaluate performances, predictions were made for each extracted area concerning future internet activity values, and MAE and RMSE measurements were taken. These steps were taken in order to be able to evaluate performances. On the basis of the findings obtained from the examination of one hundred square ids, we determined the mean of all MAE values as well as the mean of all RMSE values.

Even though the outcomes of the Linear Regression (LR) model, the Modified Support Vector Machine (MSVM), and the Feedforward Neural Network (FNN) are comparable to one another, it is abundantly clear that the Linear Regression (LR) model produces the worst results. MSVM outperforms each of the other four algorithms that were considered for this study. It is able to deal with the nonlinearity that arises as a consequence of the spatiotemporal dataset as a result of this.

### 5.2 FUTURE WORK

While the Modified Support Vector Machine (MSVM) algorithm has shown promising results in this study, there are several avenues for future research and enhancements. The following are some potential directions for future work:

a. Incorporating Advanced Feature Selection: Explore more advanced feature selection techniques, such as recursive feature elimination, genetic algorithms, or deep learning-based feature extraction methods. These techniques can help identify the most relevant features and further improve the performance of the MSVM model.

b. Investigating Different Kernels: Continue to explore different kernel functions in SVM and evaluate their impact on the MSVM model's performance. Experiment with alternative kernels, such as sigmoid or exponential kernels, and assess their suitability for specific datasets or problem domains.

c. Handling Imbalanced Data: Further investigate techniques to handle imbalanced datasets, which are common in many real-world applications. Explore approaches like cost-sensitive learning, resampling methods, or ensemble techniques to address the challenges posed by imbalanced class distributions.

d. Ensemble Methods: Explore the integration of ensemble methods with the MSVM algorithm to further improve its performance. Techniques like bagging, boosting, or stacking can be employed to combine multiple MSVM models or other complementary models to enhance prediction accuracy and robustness.

e. Incorporating Deep Learning Techniques: Investigate the integration of deep learning architectures, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), with the SVM framework. This combination can leverage the representation learning capabilities of deep learning models and the interpretability and efficiency of SVM.

f. Large-Scale Optimization: Extend the MSVM algorithm to handle large-scale datasets efficiently. Explore techniques like stochastic gradient descent, mini-batch learning, or distributed computing frameworks to overcome computational challenges when dealing with massive amounts of data.

g. Domain-Specific Extensions: Explore domain-specific modifications or extensions to the MSVM algorithm. Different application domains may require tailored adaptations, such as incorporating domain knowledge, addressing specific data characteristics, or handling specific constraints or requirements.

h. Performance Evaluation on Diverse Datasets: Conduct further experiments on diverse datasets to assess the generalization capability and robustness of the MSVM algorithm. Evaluate its performance across various domains and consider datasets with different characteristics, such as different data distributions, varying levels of noise, or varying levels of feature importance.

i. Real-World Deployment: Validate the MSVM algorithm in real-world applications and deployment scenarios. Collaborate with industry partners or stakeholders to apply the MSVM model to real-world datasets and assess its performance, scalability, and effectiveness in practical settings.

By addressing these future research directions, we can further enhance the MSVM algorithm's capabilities, extend its applicability to different domains, and provide valuable insights into its performance and potential improvements. These advancements will contribute to the ongoing development of robust and accurate machine learning models for various real-world applications.

The traffic can be affected by external factors. Therefore, this would be extended by using external factors such as the number of base stations and social activity. When examining the previous research made related to this topic, it was mentioned the power of RNN algorithm. It has thought that RNN can be suitable for this Project to obtain better performance.

# REFERENCES


[1] J. Mejia, A. Ochoa-Zezzati, and O. Cruz-Mej´ıa, “Traffic forecasting on mobile networks using 3D convolutional layers,” *Mobile Networks and Applications*, vol. 25, no. 6, pp. 2134–2140, 2020.

[2] A. R. Abdellah, O. A. K. Mahmood, A. Koucheryavy, and K. Andrey, “IoT traffic prediction using multi-step ahead prediction with neural network,” in *Proceedings of the 2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1–4, Dublin, Ireland, October 2019.

[3] 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.

[4] A. Mozo, B. Ordozgoiti, and S. Go´mez-Canaval, “Forecasting short-term data center network traffic load with convolutional neural networks,” *PloS One*, vol.13, no. 2, Article ID e0191939, 2018.

[5] A. R. Abdellah and A. Koucheryavy, “Deep learning with long short-term memory for IoT traffic prediction,” *Lecture Notes in Computer Science*, vol. 12525, pp. 267–280, 2020.

[6] A. R. Abdellah, O. Abdulkareem Mahmood, and A. Koucheryavy, “Delay prediction in IoT using machine learning approach,” in *Proceedings of the 2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 275–279, Brno, Czech Republic, October 2020.

[7] A. Essien and C. Giannetti, “A deep learning framework for univariate time series prediction using convolutional LSTM stacked autoencoders,” in *Proceedings of the 2019 IEEE International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, pp. 1–6, Sofia, Bulgaria, July 2019.

[8] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: a survey,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.

[9] Y. Zhou and Z. Zhang, “Prediction of traffic flow based on deep learning,” *International Journal of Advanced Computer Technology*, vol. 9, no. 2, pp. 05–11, 2020, Retrieved from https://ijact.org/index.php/ijact/article/view/42.

[10] X. Wang, S. Chen, J. Su, and M. Picone, “Real network traffic collection and deep learning for mobile app identification,” *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 4707909, 14 pages, 2020.

[11] F. Tang, Z. M. Fadlullah, B. Mao, and N. Kato, “An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: a deep learning approach,” *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 5141–5154, 2018.

[12] M. Kaur and S. Kadam, “Bio-inspired workflow scheduling on HPC platforms,” *Tehniˇcki glasnik*, vol. 15, no. 1, pp. 60–68, 2021.

[13] M. R. Shahid, G. Blanc, Z. Zhang, and H. Debar, “IoTdevices recognition through network traffic analysis,” in *Proceedings of the 2018 IEEE International Conference on Big Data (Big Data)*, pp. 5187–5192, Seattle, WA, USA, December 2018.

[14] A. Aksoy and M. H. Gunes, “Automated IoT device identification using network traffic,” in *Proceedings of the ICC 2019 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019.

[15] O. Salman, I. H. Elhajj, A. Chehab, and A. Kayssi, “A machine learning-based framework for IoT device identification and abnormal traffic detection,” *Transactions on Emerging Telecommunication Technologies*, vol. 2019, Article ID e3743, 15 pages, 2019.

[16] L. Nie, X. Wang, L. Wan et al., “Network traffic prediction based on deep belief network and spatiotemporal compressive sensing in wireless mesh backbone networks,” *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 1260860, 10 pages, 2018.

[17] N. Alqudah and Q. Yaseen, “Machine learning for traffic analysis: a review,” *Procedia Computer Science*, vol. 170, pp. 911–916, 2020.

[18] L. Nie, Z. Ning, M. S. Obaidat et al., “A reinforcement learning-based network traffic prediction mechanism in intelligent internet of things,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2169–2180, 2021.

[19] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, “Spatiotemporal wireless traffic prediction with recurrent neural network,” *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 554–557, 2018.

[20] S. K. Singh and A. Jukan, “Machine-learning-based prediction for resource (re) allocation in optical data center networks,” *Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D12–D28, 2018.

[21] V. Jagota and R. K. Sharma, “Wear volume prediction of AISI H13 die steel using response surface methodology and artificial neural network,” *Journal of Mechanical Engineering and Sciences*, vol. 14, no. 2, pp. 6789–6800, 2020.

[22] M. Rakhra, R. Singh, T. K. Lohani, and M. Shabaz, “Metaheuristic and machine learning-based smart engine for renting and sharing of agriculture equipment,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 5561065, 13 pages, 2021.

[23] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, “Traffic models for machine type communications,” in *ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems*, 2013, pp. 1–5.

[24] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, “Minimizing age of information in vehicular networks,” in *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*. IEEE, 2011, pp. 350–358.

[25] A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing the age of information through queues,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5215–5232, 2019.

[26] P. Popovski, C. Stefanoviˇ c, J. J. Nielsen, E. de Carvalho, M. Angjelichi-´ noski, K. F. Trillingsgaard, and A. Bana, “Wireless access in ultrareliable low-latency communication (URLLC),” *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5783–5801, 2019.

[27] A. Laya, L. Alonso, and J. Alonso-Zarate, “Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives,” *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 4–16, First 2014.

[28] N. H. Mahmood, R. Abreu, R. Bohnke, M. Schubert, G. Berardinelli,¨ and T. H. Jacobsen, “Uplink grant-free access solutions for URLLC services in 5G new radio,” in *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, 2019, pp. 607–612.

[29] 3GPP, “Study on RAN improvements for machine-type communications,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 37.868, 2012.

[30] R.-G. Cheng, J. Chen, D.-W. Chen, and C.-H. Wei, "Modeling and analysis of an extended access barring algorithm for machine-type communications in LTE-A networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 2956–2968, 2015.

[31] S.-Y. Lien, T.-H. Liau, C.-Y. Kao, and K.-C. Chen, "Cooperative access class barring for machine-to-machine communications," *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 27–32, 2012.

[32] S. Duan, V. Shah-Mansouri, and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *2013 IEEE Global Communications Conference (GLOBECOM)*, 2013, pp. 4747– 4752.

[33] 3GPP, "Service accessibility," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.011, 12 2018, version 16.4.0.

[34] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 668–695, 2021.

[35] S. Ali, N. Rajatheva, and W. Saad, "Fast uplink grant for machine type communications: Challenges and opportunities," *IEEE Communications Magazine*, vol. 57, no. 3, pp. 97–103, March 2019.

[36] M. Shehab, A. K. Hagelskjær, A. E. Kalør, P. Popovski, and H. Alves, "Traffic prediction based fast uplink grant for massive IoT," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020, pp. 1–6.

[37] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, "Traffic Models for Machine Type Communications," in *ISWCS 2013; The Tenth International Symposium on Wireless Communication Systems*, Aug 2013, pp. 1–5.

[38] E. Grigoreva, M. Laurer, M. Vilgelm, T. Gehrsitz, and W. Kellerer, "Coupled Markovian Arrival Process for Automotive Machine Type Communication traffic modeling," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.

[39] P. Salvo Rossi, D. Ciuonzo, and T. Ekman, "HMM-based decision fusion in wireless sensor networks with noncoherent multiple access," *IEEE Communications Letters*, vol. 19, no. 5, pp. 871–874, 2015.

[40] A. E. Kalør, O. A. Hanna, and P. Popovski, “Random Access Schemes in Wireless Systems with Correlated User Activity,” in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.

[41] S. M. Hasan, K. Mahata, and M. M. Hyder, “Fast uplink grant-free NOMA with sinusoidal spreading sequences,” 2020.

[42] Z. Zhou, R. Ratasuk, N. Mangalvedhe, and A. Ghosh, “Resource allocation for uplink grant-free ultra-reliable and low latency communications,” in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, pp. 1–5.

[43] S. Ali, A. Ferdowsi, W. Saad, and N. Rajatheva, “Sleeping multi-armed bandits for fast uplink grant allocation in machine type communications,” in *2018 IEEE Globecom Workshops (GC Wkshps)*, Dec 2018, pp. 1–6.

[44] E. Eldeeb, M. Shehab, and H. Alves, “A learning-based fast uplink grant for massive IoT via support vector machines and long short-term memory,” *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[45] O. Habachi, M.-A. Adjif, and J.-P. Cances, “Fast uplink grant for NOMA: A federated learning based approach,” in *Ubiquitous Networking*, O. Habachi, V. Meghdadi, E. Sabir, and J.-P. Cances, Eds. Cham: Springer International Publishing, 2020, pp. 96–109.

[46] I. AlQerm and B. Shihada, “A cooperative online learning scheme for resource allocation in 5G systems,” in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–7.

[47] A. E. Kalør, O. Simeone, and P. Popovski, “Prediction of mmwave/thz link blockages through meta-learning and recurrent neural networks,” 2021. [Online]. Available: https://arxiv.org/abs/2106.07442

[48] F. Shah-Mohammadi and A. Kwasinski, “Deep reinforcement learning approach to QoE-driven resource allocation for spectrum underlay in cognitive radio networks,” in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1–6.

[49] X. Liu, J. Yu, J. Wang, and Y. Gao, “Resource allocation with edge computing in IoT networks via machine learning,” *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3415–3426, 2020.

[50] D. Hejji, M. A. Talib, A. B. Nassif, Q. Nasir, and A. Bouridane, "AIbased models for resource allocation and resource demand forecasting systems in aviation: A survey and analytical study," in *2021 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS)*, 2021, pp. 183–189.

[51] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten challenges in advancing machine learning technologies toward 6G," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020.

[52] C. Ren, X. Lyu, W. Ni, H. Tian, and R. P. Liu, "Distributed online learning of Fog computing under nonuniform device cardinality," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 1147–1159, 2019.

[53] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends in Networking, Now Publishers, Inc.*, 2017.

[54] O. Capp, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*. Springer Publishing Company, Incorporated, 2010.

[55] S. Sefati and N. J. Navimipour, "A QoS-aware service composition mechanism in the internet of things using a hidden-markov-model-based optimization algorithm," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15620–15627, 2021.

[56] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.

[57] S. L. C. G. Cassandras, *Introduction to Discrete Event Systems.* Springer, 2008.

[58] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[59] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile Unmanned Aerial Vehicles (UAVs) for Energy-Efficient Internet of Things Communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, Nov 2017.