Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Telecommunication Engineering Graduate Program

Prediction of Radio Resource Allocation in Addis Ababa's LTE Network

By:

Solomon Teka

Advisor:

Dr. Eng. Yihenew Wondie

A Thesis Submitted to the School of Electrical and Computer Engineering Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Science in Telecom Network Engineering

Nov 2025

Addis Ababa, Ethiopia

Addis Ababa Institute of Technology
School of Electrical and Computer Engineering
Telecommunication Engineering Graduate Program

Prediction of Radio Resource Allocation in Addis Ababa's LTE Network

By:

Solomon Teka

Members of the examining committee's names and signatures:

Dr. Eng. Yihenew Wondie		
Advisor	Signature	Date
Internal Examiner	Signature	Date
External Examiner	Signature	Date
Chair or School Dean	Signature	Date

Declaration

In my opinion, this thesis is completely my own original work; no part of it has been submitted for a degree at any institution, and all sources have been properly referenced.

Solomon Teka

Name | Signature

Place: Addis Ababa Institute of Technology, Ethiopia.

Date of Submission: ____________

This thesis has been submitted for approval for examination by a university advisor.

Dr. Eng. Yihenew Wondie

Name | Signature

Abstract

The rapid growth of mobile data traffic in Addis Ababa places increasing pressure on Ethio Telecom's LTE network and highlights weaknesses in existing radio resource management practices. This thesis develops and evaluates a predictive framework to improve downlink Physical Resource Block (PRB) allocation by forecasting PRB demand and enabling proactive resource management. Using real operational data collected from Ethio Telecom for a multi-month period, the study formulates PRB utilization prediction as a univariate time-series problem and implements Long Short-Term Memory (LSTM) neural networks trained on preprocessed hourly and weekly PRB usage traces. The thesis describes the end-to-end pipeline: data collection and cleaning, feature preparation, model training and hyperparameter tuning, and evaluation using standard forecasting metrics (RMSE, MAE, MAPE, and R^2). Results indicate the proposed model produces accurate short-term PRB utilization forecasts and, when integrated with a reactive allocation policy, can reduce periods of congestion and improve throughput relative to static allocation strategies. The contributions include (1) a contextualized dataset and preprocessing approach for Addis Ababa's LTE environment, (2) an LSTM-based forecasting model adapted for PRB utilization prediction, and (3) a practical framework for integrating forecasts into operational PRB allocation.

Keywords: LTE, Physical Resource Block (PRB), predictive analytics, LSTM, radio resource allocation, time-series forecasting

Acknowledgment

First of all, I want to express my deepest gratitude to Almighty God and the Holy Virgin Mariam for their divine intervention and care, which I felt during the course of this research. Dr. Eng. Yihenew Wondie, my advisor, is certainly the one to be thanked most. His constant help, encouragement, and indispensable guidance were the main reasons that this research would not have been possible.

I also wish to thank highly the persons, who have been the judges, Dr. Ephrem Teshale and Dr. Byniyam Birhanu. The solutions that they gave, the criticisms that they made, and the changes that they provided throughout the thesis work were of great help to me. Their directions became a driving force for me to get over the challenges that I faced during the conduct of this research.

Besides that, I am deeply grateful that I have been given moral and financial support from the telecoms sector of Ethiopia and the Addis Ababa University (Addis Ababa Institute of Technology) in a form of a scholarship for my academic work.

In the end, I cannot be silent about my full gratitude to my beloved wife, Samerawit Demisse, my son, Boniyat, and my cute daughter, Yanet. The love and encouragement which they gave me during the process of this thesis cannot be compared to anything else.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
BE	Best Effort
CNN	Convolutional Neural Networks
CQI	Channel Quality Indicator
CSI	Channel State Information
DL	Downlink
eNodeB	Evolved Node B
eMBB	enhanced Mobile Broadband
EPC	Evolved Packet Core
EPF	Enhanced Proportional Fair
EXP-Rule	Exponential Rule
EXP/PF	Exponential/Proportional Fair
FLS	Frame Level Scheduler
GRU	Gated Recurrent Unit
IOT	Internet of Things
IQR	Interquartile Range
KPIs	Key Performance Indices
LTE	Long Term Evolution
LTE-A	LTE-Advance
LOG-Rule	Logarithmic Rule
LSTM	Long Short Term Memory
MAPE	Mean Absolute Percentage Error
MCS	Modulation and Coding Scheme
ML	Machine Learning
MSE	Mean Squared Error
MLWDF	Maximum Largest Weighted Delay First

PBCH	Physical Broadcast Channel
PDCCH	Physical Downlink Control Channel
PDSCH	Physical Downlink Shared Channel
PF	Proportional Fair
PRB	Physical Resource Block
QOS	Quality of Service
RMSE	Root Mean Squared Error
RRC	Radio Resource Connection
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
SNR	Signal to Noise Ratio
TTI	Transmission Time Interval
UE	User Equipment
XGBoost	eXtreme Gradient Boosting

Chapter 1

Introduction

1.1 Background and Motivation

The exponential expansion of mobile data traffic, facilitated by mobile phones, connected devices, and data-heavy applications, has occurred at such a rapid pace that it has placed significant stress on Long-Term Evolution (LTE) networks to deliver high-quality, reliable services. To address the increased demand for the use of available radio resources, it is important that they are allocated as efficiently as possible. The primary goal is to achieve the highest level of utilization possible from limited radio spectrum resources, while still providing acceptable levels of performance for Key Performance Indicators (KPIs) such as throughput, latency, and user satisfaction. The allocation of radio resources in LTE networks is accomplished through the use of PRBs (Physical Resource Blocks). Each PRB contains 12 sub-carriers, and is allocated to a user in a time-slot of 0.5 ms (180 kHz of frequency spectrum) [1].

Correct allocation of PRB's is key to meeting QoS criteria, reducing inter-cell interference, and maximizing spectral efficiency in urban areas which are busy and experience considerable variability in traffic. The way LTE-A networks support multimedia applications, provide higher data rates, and support large growth in capacity means that effective RRM (Radio Resource Management) processes will be required to address the highly dynamic characteristics of wireless channels [2]. Predictive modelling of PRB allocation has become increasingly popular as a strategy for predicting future demand for network resources and optimising PRB utilisation before they are actually needed, thus improving the entire network performance [3]. This thesis aims to provide such a model for the specific case of the LTE network in Addis Ababa, thus addressing an urgent need for context-specific approaches in urban wireless networks. While there has been much progress made in RRM, the available methods for the allocation of radio resources in LTE networks still have significant difficulties and constraints.

Conventional scheduling algorithms, such as Proportional Fair (PF), Maximum Throughput (MT), and Modified Largest Weighted Delay First (M-LWDF), are generally incapable of adequately responding a player's behaviour with regards to a dynamic traffic pattern or possible changes to the channel quality indicator (CQI) state [4]. An example of how this lack of responsiveness will affect a player is when thesealgorithms prioritize throughput at a cost to fairness or when they fail to consider the effects of real-time mobility of users and network load when operating over heterogeneous networks comprised of macro and small-cells [5]. Optimization-based approaches to scheduling have been shown to be effective in non-competitive/controlled environments, such as in Evolutionary Multi-Objective Optimization and game theory based solutions but tend to be computationally intensive and not practical for real-time applications found in large-scale networks [6]. Furthermore, when using static or semi-static allocation schemes, both primary and secondary users in cognitive LTE networks are unable to accommodate time-sensitive bursty traffic and maintain QoS guarantees [7]. There are significant limitations on the use of instantaneous channel state information (CSI) without the ability to predict will produce many problems that fall under the category of lack of proactive resource mazement, such as call blocking during handoffs and under-utilization of resources at the edges of cells [2]. These problems further support the need for sophisticated predictive models capable of accurately predicting PRB utilization as well as being responsive to the stochastic characteristics of a wireless environment.

The Addis Ababa LTE network provides an interesting new test-bed for the study of radio resource allocation in both technical and socioeconomic terms. As the capital of Ethiopia, Addis Ababa is a key region in the country which has an enlarged base of mobile users and government led digital initiatives like Digital Ethiopia 2025 push for high data service [8]. PRB assignment isvital to the efficient operation of the network given the challenges faced in one of the city's LTE networks predominantly by Ethio Telecom with respect to traffic load, non-uniform distribution of users, and interference in urban canyons, making efficient PRB allocation critical for network performance [9].

The efficiency of the Prophet model, which achieved an R-squared value of 0.95 compared to the Long Short-Term Memory (LSTM) and eXtreme Gradient Boosting (XGBoost) models, was proved in a recent thesis at Addis Ababa University, which focused on

PRB usage prediction for load balancing between frequency layers [9]. Nonetheless, only a few studies have focused on creating a predictive model of Addis Ababa's unusual characteristics and thus, very few predictive models have been developed in order to fit Addis Ababa's specific unique features of their network (topology, user behaviour and restrictions, etc.). This lack of predictive models emphasizes the need for customized solutions that will improve the overall functionality of this key area and its effect on other urban LTE networks around the world in developing countries [3].

The problem addressed by this thesis is that there are few strong, context-specific predictive models for predicting the use of physical resources blocks within the LTE network of Addis Ababa. Although machine learning methods such as Prophet have demonstrated some success in general LTE environments to date, the way these methods are used in the context of Addis Ababa's unique network will not be addressed in any detailed way in the body of the research [9]. Moreover, the vast majority of studies to date have either focused on generic LTE networks or other geographical locations without taking into account localized variables, such as high-density user populations, varying volumes of user traffic and minimal infrastructure upgrades, which have a considerable impact on resource allocation within Addis Ababa [7]. The dynamic nature of the way that physical resource blocks are allocated, coupled with the added complications presented by the movement of users resulting in many handovers and interference in densely populated urban environments makes it very difficult to accurately predict the volume of physical resource blocks needed for an LTE network in Addis Ababa [5].

The gap that exists is the lack of a predictive framework that incorporates local network data, accommodates the time-variant nature of user traffic, and employs advanced machine learning techniques that will allow for the effective prediction of physical resource block utilization. The existence of this gap materially restricts network operators' ability to proactively manage their resources, resulting in the inefficient management of resources, which exemplifies as a waste of resources, causes degradation in QoS and leads to increased CAPEX and OPEX costs for the operator [2]. In this research, the aim is to explore how predictive models that are based on machine learning can be constructed to provide accurate forecasts of the utilization of PRBs in Addis Ababa's LTE Network, taking into consideration the unique traffic values and constraints associated with the particular network under study. To accomplish this task, this research will pursue the following three objectives: (1) develop a viable machine learning model capable of predicting PRB utilization based on both historical and real time information from the Addis Ababa LTE Networks; (2) evaluate the operational performance of this model compared to other existing techniques with regards to accuracy, scalability and adaptability to dynamic conditions, and (3) present a framework that integrates predictive data into RRM strategies in order to improve network efficiency and become more efficient with respect to the quality of service. These objectives will bridge the gap between traditional generic predictive models and those developed to satisfy the Joint network requirements of the Addis Ababa LTE Network so that a customized solution for resource allocation can be developed.

The thesis contributes significantly to the literature in multiple areas. First, it expands the reach of machine learning for predicting PRB utilisation in the understudied area of the LTE network of Addis Ababa, as previous research has demonstrated that models such as Prophet are accurate and effective in predicting PRB utilisation but have been limited to research setting only [9].The second point of this thesis includes data from local networks including traffic flow rates, end-user mobility, and infrastructure characteristics in the predictive model, thus mitigating the challenges associated with generic models that do not capture the unique characteristics of any given location [3]. The third contribution of this thesis is to offer network operators in Addis Ababa concrete and tangible benefit by providing a scalable method for optimizing the allocation of physical radio bearers (PRB) leading to a decrease in call blocking and an increase in the quality of service (QoS) provided to cell-edge users [5] The fourth contribution of this thesis is the development of a data-driven predictive modelling framework for machine learning-based LTE radio resource management (RRM) through a comparison between LSTM and XGBoost and the newly developed model [7].Overall, these contributions advance the field of radio resource allocation by presenting a context-sensitive, data-driven method of optimising radio resources in urban LTE networks, while also introducing different concepts to be used whenever managing resources intelligently.

1.2 Statement of the Problem

In Addis Ababa's LTE network, the way radio resources are assigned is very reactive. The operator waits until congestion or demand spikes occur and then adjusts capacity, (broadband, power and channels) in real-time to try and relieve the congestion or respond to demand. There is no consideration of historical traffic data, user behaviour trends or predictive analytics, leading to inefficient use of resources; ongoing network overloads during peak times that can be easily predicted and create a poor customer experience with dropped calls, high latency and unreliable throughput (especially for high-demand applications, e.g., video streaming and real-time communications) due to a lack of proactive resource forecasting. The lack of proactive resource forecasting causes spatial imbalances in the network; for example, densely populated areas are continually under-provisioned while resources remain unused in less populated areas, straining infrastructure, driving up operating costs and limiting scalability for a market that is growing quickly with LTE subscribers.

The aim of this project is to create a prediction model based on machine learning that is designed specifically to fit both the population and the geography of Addis Ababa. This will assist in solving numerous different issues related to traffic congestion found within the area at the current time. To develop this model, we will take both historic data regarding traffic and also current hot spots of demand (areas experiencing higher levels of demand than others) into consideration as a basis for the prediction model. By doing so, we will be able to predict future needs for resources, proactively allocate those resources across the network before they are needed, and optimise the use of resources across multiple geographic areas. We hope that by moving from a reactive means of traffic management to a predictive management system, we will be able to reduce peak traffic congestion and eliminate resource waste, as well as improve quality of service metrics, including latency, throughput, and stability. Furthermore, our intention is to devise a validated predictive model of the LTE network using actual metrics from the LTE network in order to develop a scalable platform for managing and quantifying the reliability of services provided to millions of users located throughout Ethiopia and create a reference model that can be applied to other metropolitan areas within Ethiopia that are experiencing demographically similar populations or also experiencing similar traffic congestion challenges.

1.3 Objectives

1.3.1 General Objective

To predict DL PRB utilization in Addis Ababa's LTE network using machine learning-based algorithms.

1.3.2 Specific Objectives

The following specific objectives have been specified in order to attain the overall goal:

- Review existing LTE performance and PRB allocation strategies.
- Identify and evaluate machine learning algorithms best suited for predicting DL PRB utilization based on localized data and network dynamics in Addis Ababa.
- Gather historical and real-time data specific to Addis Ababa's LTE network.
- A machine learning model forecasts DL PRB utilization, analyzing peak-hour congestion and off-peak underutilization to optimize proactive resource allocation strategies.
- Evaluate model performance, draw conclusions on its effectiveness, and recommend strategies to optimize DL PRB utilization for enhanced network efficiency and user experience.

1.4 Literature Review

As a result of the earlier thesis, this literature review includes a related thesis work analysis. The literature review will also make it evident how unique this thesis proposal work is the idea of resource allocation for LTE network issues and solutions has to be established by assured researchers. Several journal articles, conference papers, and theses were reviewed to understand those challenges and solutions.

In [10] investigates the complexities associated with resource allocation in Long-Term Evolution Advanced (LTE-A) networks, motivated by the increasing demand for efficient data transmission and improved Quality of Service (QoS) in heterogeneous network environments. The primary objective of the study is to compare the performance of various machine learning-based dynamic resource allocation methods utilizing actual data from the Ethiotelecom LTE-A network in Addis Ababa. To achieve this, the authors employ a comprehensive methodology that includes the collection and preprocessing of real-world data, followed by the implementation of advanced machine learning algorithms such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). The thesis has shown a number of major gaps in the literature available today, particularly related to the usage of traditional methods for allocating resources (such as pre-defined rules) and that these method do not tend to relate well to the changing patterns of user demand and traffic over time. Many studies have also used synthetic datasets to test their algorithms and are therefore unable to reflect the true behavior of a real world network. By the use of a machine learning approach, the resources allocated to a network can be done much more effectively than with traditional approaches can therefore be modified as the conditions of the network change and will therefore have a more positive impact on the overall performance of the network. In terms of the use of realistic datasets, the findings of the research provide significant insights into how machine learning-based approaches will be able to help enhance the effectiveness and efficiency of the resources allocated to a network and will therefore contribute to quality-of-service (QoS) improvements and higher user satisfaction in LTE-A networks.

The article [11] analyzes current LTE downlink scheduling algorithms and their performance for the ethiotechnic company, with LTE is facing two major issues: 1) Congestion due to limited resources and 2) Poor signal quality after deploying LTE. The aim of this research is to analyze downlink scheduling algorithms that are used in LTE and measure their performance against the QoS's defined resources and energy consumption before recommending one of them. The methodology consists of modeling, simulating, and evaluating five popular packet scheduling algorithms, MLWDF, EXP-Rule, EXP/PF, LOG-Rule and FLS, using the LTE-Sim tool. This study finds a void in the existing literature regarding the optimization of scheduling algorithms in the context of ethio telecom. It was found that all of the Algorithms considered were primarily biased towards real time services. The FLS algorithm was demonstrated to significantly outperform the other algorithms with regard to throughput, Packet loss ratio, fairness and energy efficiency for VoIP video traffic, while all other algorithms performed poorly on best effort (BE) traffic. The study also highlights that selecting the right Scheduling algorithm can significantly improve overall system performance, and that the research lays a foundation for subsequent research on this topic.

In [12], some authors focus on the problem of underutilized LTE Resources caused by static resource allocation at the eNodeB level. They develop an efficient machine learning model that considers KPIs like traffic volume, the number of concurrent users and RRC resources as input data for dynamically predicting RRC allocations. Three machine learning approaches are implemented: linear regression, convolutional neural networks (CNN), and long short-term memory (LSTM), with the LSTM prediction model showing 97% accuracy when predicting corresponding resource requirements. The researchers identified a major shortfall with most existing resource allocation methodologies which lead to either saturated or underutilized resources because of their incomplete capacity due to fixed resource allocations. The results indicate that by providing machine learning techniques to optimize resource allocation, mobile network operators could improve their overall efficiency by reducing the financial losses associated with incorrectly managed traffic. Ultimately, the authors conclude that this dynamic allocation model and its LSTM method can improve the operational efficiency of LTE networks through automated management of resources without any human involvement.

In [13] looks at how PRBs are allocated in the LTE downlink to improve user service quality by using efficient scheduling algorithms. The main goal is to evaluate various scheduling algorithms, including Modified Round Robin, Round Robin, Maximum Channel State Information (CSI), and Proportional Fair with respect to spectral efficiency and enode-B throughput. The methodology involves simulating user distribution within an enode-B's coverage area and assessing channel conditions via SNR caused by pathloss, shadow fading, and multipath fading. There is a large gap in existing literature concerning how different scheduling algorithms affect overall network performance under different user conditions.The results show that the Modified Round Robin algorithm has significantly higher performance than all other approaches, with a maximum enode-B payload of 158,512 Gbit as well as an enode-B throughput of 44.031 Mbps, demonstrating that this approach can optimize the allocation of resources and enhance spectral efficiency when compared with previous methods. Therefore, this research will provide further knowledge about improving the quality of service within LTE networks and emphasizes the importance of using adaptive scheduling techniques for optimal resource usage.

The paper reported on in [14] outlines the comprehensive work performed to forecast utilization of Physical Resource Blocks (PRBs) on LTE networks with a focus upon the Ethiopian telecommunications network (Ethiotelecom) in Addis Ababa. The research investigates important problems associated with allocating resources and balancing loads and highlights that traditional static methods used to model data traffic are insufficient due to dynamically changing characteristics of mobile data. The research utilizes machine learning techniques such as Prophet, Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) to establish intelligent algorithms capable of forecasting future PRB utilization. Resource distribution across frequency bands can also be optimally achieved based on this data.

The methodology outlined in this research paper is clearly defined and includes an exhaustive literature review, collection of actual data obtained from real-world resources (as well as performance evaluations of several models). In addition to providing needed insight into the LTE networks, the outcome of this study will help further advance LTE network optimization by providing a data-driven methodology to enhance the performance of the networks and improve the user experience, given the increasing demand for data, as well as by providing a path for continued development of intelligent load balancing methodologies utilizing theoretical and practical applications.

In [15]'Hotspot Prediction Using Deep Learning: In the Case of Addis Ababa LTE Network," presents a pioneering approach to predict network hotspots in Addis Ababa's Long-Term Evolution (LTE) network using a Long Short-Term Memory (LSTM) deep learning model, driven by the urgent need to proactively manage cell congestion to enhance Quality of Service (QoS) and reduce operational costs in the face of escalating mobile data demands; motivated by the inefficiencies of traditional manual hotspot identification methods, such as drive tests and KPI analysis, and the shortcomings of prior models that relied on data traffic volume or user density—approaches that fail to account for varying cell capacities and QoS factors like throughput, which is influenced by radio conditions, user mobility, and cell load—the study sets out to develop a model centered on the number of active users and their throughput, with specific objectives to preprocess LTE cell data, build an LSTM-based prediction model, and evaluate its performance using metrics like accuracy, precision, recall, and F1 score; the methodology entails collecting hourly cell counters and KPIs from 200 cells over 54 days (July 1 to August 23, 2021) via Ethiotelecom's Performance Reporting System, followed by comprehensive preprocessing steps including data cleaning to remove duplicates and irrelevant features, feature selection to retain 31 correlated features, imputation of missing values using per-cell per-hour historical means leveraging 24-hour seasonality, labeling cells as hotspots (.12 users, !4 Mbps throughput) or low traffic based on Ethiotelecom thresholds, under-sampling to address class imbalance (achieving a 32.32% minority class ratio), data normalization for neural network stability, and splitting into 80% training and 20% testing sets, culminating in training a sequence-to-sequence LSTM model to predict hotspot status for four future hours, optimized with the Adam algorithm and hyperparameters like 32 neurons and a 0.0001 learning rate, yielding an impressive 89.13% accuracy and 85.5% F1 score; however, gaps include the model's sensitivity to class imbalance, with performance deteriorating (F1 score drops to 74.81% at 23.11% minority class ratio), its exclusive focus on downlink data, omitting uplink KPIs critical for holistic network optimization, reliance on Ethiotelecom-specific thresholds that may not generalize, and a relatively limited dataset size that could be expanded for greater robustness; key lessons learned highlight the superiority of throughput and user-based metrics over volume or density for accurate hotspot identification, the effectiveness of LSTM in capturing long-term dependencies in time-series data, the critical role of meticulous preprocessing (e.g., handling missing values and resampling) in enhancing model performance, the significant impact of class imbalance on classification outcomes, and the model's practical potential to enable realtime network optimization, reduce capital and operational expenditures, and improve

QoS, suggesting future research directions such as incorporating uplink measurements, exploring balanced datasets, and testing the model across diverse network environments to enhance its applicability and robustness.

In [16], the growth in mobile data traffic has been dramatic since the introduction of the 3G system. In fact, mobile data consumption has increased by over 30 times in the last three years. The development of technology and data-intensive applications has contributed to this pace of growth. Establishing reliable predictive models of mobile data traffic is crucial for optimizing the performance of LTE networks. The purpose of this study is to create a deep learning multi-step prediction model based on multiple variables that affect the variance of data traffic so that LTE cell sites can be optimized and upgraded before there is any congestion or downgrades. The data processing method used in this study consisted of collecting monthly data from 690 eNodeBs over a four month period (one hour per eNodeB), which included a range of variables including the volume of data traffic in both downlink and uplink directions, throughput, and user count. The processes used to pre-process the collected data included checking for missing, outlier and "noisy" data before applying feature selection and scaling. One major aspect of this study's methodology is clustering of base stations with similar traffic patterns based upon time using K-means clustering to improve the accuracy of predictions.

Additionally, the article provides an evaluation of other popular modeling methods for traffic forecasting, which would include SARIMA along with Extreme Learning Machines and LSTM/GRU networks. The article emphasizes the significance of adequately managing both linear and nonlinear traffic characteristics. The research article discusses several shortcomings of current models inability to take into account multivariate characteristics and their inability to capture long-term dependencies; as such, the research article argues that there is a need for advanced deep learning methodologies such as CNN-LSTM architectures. The article presents many key takeaways including the necessity of a complete dataset for data preprocessing purposes, the potential for clustering and deep learning to create more reliable predictions through better predictions, checking the value of variables being predicted with the additional usage of multimodal datasets to build superior traffic forecasting models. Enhancing the accuracy of traffic forecasting models provides an opportunity for maximizing the efficiency of network operations and assisting in managing networks more effectively.

Researchers in [17] study how load imbalance in LTE (Long Term Evolution) networks has become increasingly problematic as a result of growing data traffic and the need for optimal resource management in order to deliver a good quality of service (QoS) to end-users. To achieve this goal, the researchers developed and validated high accuracy machine-learning models for predicting the use of physical resource blocks (PRBs) at different frequency bands of the spectrum so that load balancing could occur between frequency bands. The methodology included gathering a detailed dataset from Ethio Telecom, creating engineering features from the datasets, and measuring/assessing the model's performance using such metrics as R-squared value and mean absolute error (MAE). It was noted in the research, as a limitation, that the models were created so that their characteristics were like those of the Ethio Telecom network. This means that these models may not be used or generalized across different telecommunications networks. Furthermore, Prophet had a higher performance than all the other models when comparing the R-squared value, which was 0.95, and the MAE value, which was 4.98. Prophet is then effective as a predictor of load and an appropriate tool for load balancing purposes.

This research highlights the cruciality of identifying the right predictive models for optimizing a network and discusses how machine learning can improve the allocation of capacity to provide better utilization of resources. The key takeaway is that accurate, data-based models for predicting network behavior can greatly increase the performance of the network, allowing for improved load management strategies and development of more intelligent and dynamic load management in LTE networks.

The literature review provides approximately a summarised version of everything presented (this includes the various types of findings, methods used in the studies, and conclusions/implications drawn) in the accompanying Literature Review Table and provides a clearer perspective on the findings of the reviewed publications. The reviewed publications discuss the advancements made in using predictive techniques to optimize radio resource distribution in LTE (and specifically in the Addis Ababa context). All of the studies [10] to [17] apply different machine-learning models (e.g. LSTM, CNN, XG-

Boost) to address various issues such as inefficiencies in resource allocations, downlink scheduling performance, and load balancing issues.

The reviewed studies show that machine learning techniques outperform traditional static techniques significantly based on the overall improvement in the Quality of Service (QoS) and overall network efficiency. Some of the key limitations and/or gaps in the current body of research include: reliance on synthetic datasets; limited exploration and/or analysis of uplink metrics; and the sensitivity of predictive models to class imbalances. Lessons learned emphasize the importance of comprehensive data preprocessing, the effectiveness of multivariate feature incorporation, and the potential of advanced deep learning architectures to capture longterm dependencies. These insights pave the way for future investigations into more robust, generalizable models that can further improve resource management and user experience in LTE networks.

Table 1.1: Literature Review on LTE Resource Allocation Studies

Ref.	**Motivation**	**Objectives**	**Methodology**	**Limitations**	**Lesson Learned**
[10] 2023	Optimize LTE-A resource allocation dynamically.	Compare machine learning resource allocation methods.	Evaluate strategies using realistic LTE-A data.	Static methods insufficient for dynamic demands.	Feature selection improves model performance.
[11] 2021	Optimize RRC resource allocation in LTE.	Develop dynamic allocation using machine learning.	Compare LSTM, CNN, and linear regression.	Static data may not capture variability.	LSTM outperforms other models significantly.
[12] 2023	Improve LTE resource management efficiency.	Analyze LTE scheduling algorithms performance.	Simulate and compare scheduling algorithms.	No novel algorithm proposed.	Traditional methods lack adaptability and flexibility.
[13] 2022	Improve LTE resource allocation efficiency.	Enhance QoS for dynamic traffic.	Propose EURA with adaptive mechanisms.	Dynamic conditions may complicate implementation.	Adaptability is crucial for performance.

1.5 Methodology

The purpose of this thesis is to evaluate the improvement of efficiency of the LTE Network in ethiotelecom through resource allocation by the use of different methodologies, the overall goal of this thesis is to determine whether it is feasible to implement with specific resource allocation of service assurances on LTE networks in ethiotelecom. The impact of LTE technology using resource allocation strategy was focused on the developed analytical models for investigating the issue in ethiotelecom. Therefore, the thesis involves adopting a proof-by-construction methodology for creating new mechanisms and procedures and creating and putting into practice solutions in machine learning algorithm prediction frameworks. To gain a thorough understanding and knowledge of the appropriate areas and where the potential problem and solution may emerge, the thesis methodology of this thesis comprises the study of the difficulties and solution areas that should be directly associated with the research title. The thesis will follow the following methodology steps:

1. **Literature Review**
    - Conduct a comprehensive review of existing literature on radio resource allocation, congestion management, and predictive modeling techniques in telecommunications.
2. **Data Collection**
    - Historical Performance Data of Network Performance - Gather performance characteristics of the LTE network from Ethiotelecom, including various performance metrics such as the amount of physical resource blocks (PRBS) in use, patterns of user demand, and Quality of Service (QoS).
3. **Data Analysis**
    - Descriptive data analysis - Examine collected data for congestion and underutilization patterns throughout greater Addis Ababa.
    - Predictive Modeling - Use machine learning to create predictive models that predict the distribution of traffic based on previously collected historical data

and demographic information. Such as Regression Analysis, Time Series Forecasting, and/or Clustering.

4. **Implementation of Resource Allocation Strategy**

    - Dynamic Resource Allocation Framework - Create and employ a dynamic resource allocation framework which allows for real-time reallocation of PRBs from under-populated geographic locations to overloaded geographic locations based on data produced by the above predictive models.
    - Simulation Testing: Conduct simulations to test the effectiveness of the dynamic allocation strategy under various traffic scenarios and peak usage conditions.

5. **Evaluation of Impact**

    - Performance Measurements - Measurement of the effectiveness of the allocation methods that have been put into place based on the major performance measurements such as data rates, latency and user satisfaction by comparing the measurement data prior to implementing the allocated resources with the data collected after the implementation of the allocation strategies.

6. **Recommendations and Reporting**

    - Data analysis and impact measurements will create a set of actionable recommendations for Ethiotelecom based on the impact measurement results.

This methodology will develop a comprehensive framework to improve radio frequency resource allocation in Ethiotelecom's LTE network in Addis Ababa so they can have better service quality and higher end-user satisfaction. The methodology for the network study is illustrated in Figure 1.1.

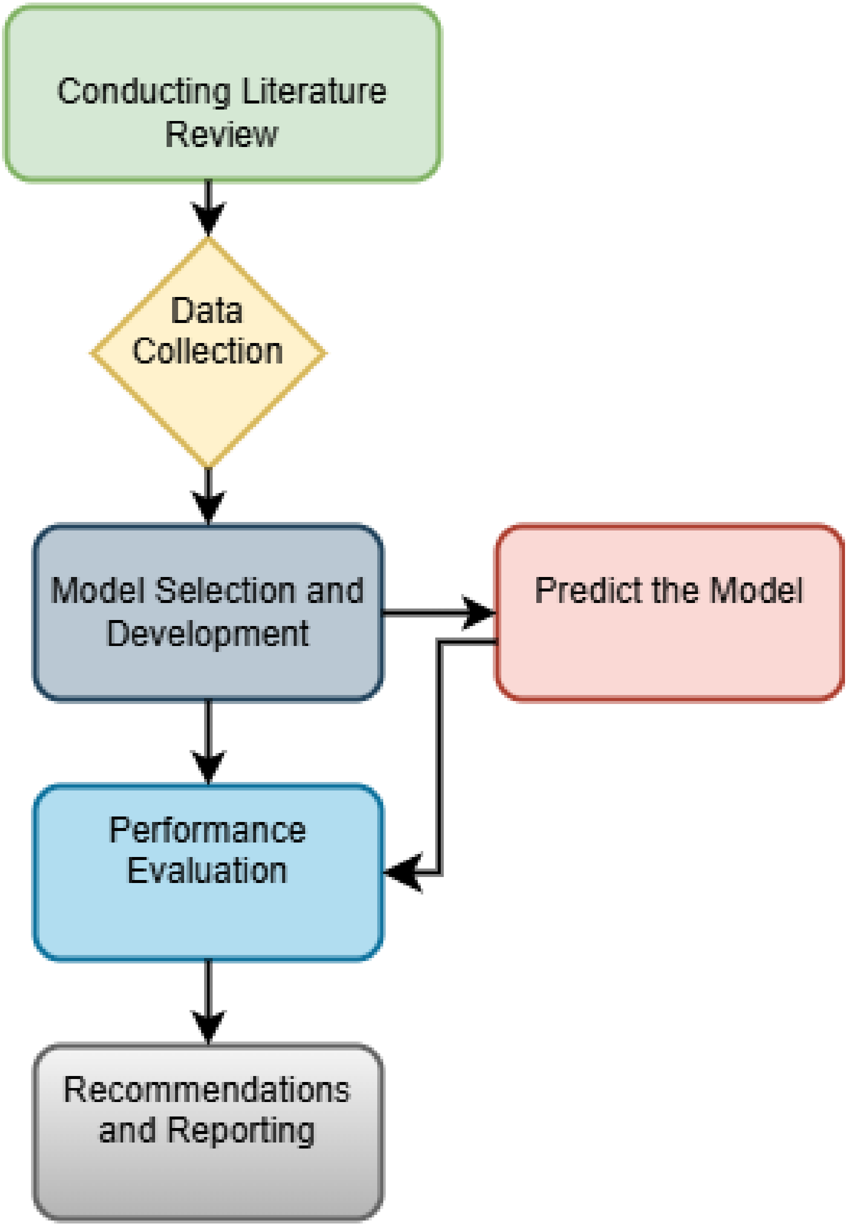

Figure 1.1: Methodology of the Thesis

1.6 Scope and Limitation

1.6.1 Scope of the Study

The research will be conducted in the Akaki Kality sub-city of Addis Ababa. It is expected that findings from the research will apply to other urban cities with the same problems regarding infrastructure and traffic congestion. The researcher will use real data from Ethiotelecom to have practical significance in performing a predictive modeling approach that may be applied to the overall city in future studies.

1.6.2 Limitations of the Study

The main restrictions which limit the applicability and deployment of these technologies are:

- Data coverage. The dataset is limited to the provided cells and period; traffic patterns may change with infrastructure upgrades or adoption rates.
- Univariate formulation. Modeling only downlink PRB utilization ignores potentially useful multivariate signals (e.g., CQI, user counts, uplink activity).
- Integration of operations: The integration of practical application with the operations of the eNodeB vendor and the management of real-time constraints is critical; thus, the research will establish forecasting as advantageous but will not provide a production level of availability.
- Assumed scheduling model. The allocation simulation uses an eNodeB scheduler abstraction; real schedulers may have additional constraints.

1.7 Contribution

This research proposes a practical and efficient predictive model for downlink PRB utilization in Addis Ababa's LTE network. By leveraging real traffic data and applying LSTM neural networks, the study aims to improve resource utilization before network saturation occurs. This proactive approach reduces data congestion, enhances user experience, and provides Ethiotelecom with actionable insights for dynamic resource planning. The proposed framework is scalable and adaptable to similar urban LTE environments across Ethiopia.

1.8 Thesis Outline

The following is how the rest of the paper is organized: Chapter 2 covers the fundamentals of LTE technology, as well as KPIs categories, their applications, and data traffic characteristics. System Model and Data Preparation and its methods are discussed in Chapter 3. The DNNs employed in the proposed framework are also described in this chapter. The experimental settings for DL PRB Utilization prediction are presented in Chapter 4 together with the results and discussion section, and the results of the prediction models and findings are then examined. Finally, Chapter 5 presents the main conclusion and future work.

Chapter 2

Radio Resource Allocation in LTE Networks

2.1 Overview of Radio Resource Allocation in LTE Networks

The efficient management of radio resources is a critical aspect of Long-Term Evolution (LTE) networks, especially in dynamic urban environments such as Addis Ababa. The LTE standard, a fourth-generation (4G) mobile telecommunications standard, was created by the 3rd Generation Mobile Telecommunications (3GPP) project. LTE aims to deliver broadband wireless service that enhances mobile user connectivity by delivering high-speed data transfers with low latency and guaranteed quality of service (QoS) [18]. In LTE, RRM consists of allocating time-frequency resources (called Resource Blocks or RBs) to User Equipment (UE) to optimize resource usage while meeting various QoS needs. The high density of people and the different traffic patterns of users in urban contexts such as Addis Ababa create significant challenges for RRM. Therefore, efficient resource allocation in urban LTE networks is critical to maintaining the reliability of the network and providing an acceptable level of satisfaction to end users [17]. This chapter discusses the LTE network architecture and resource allocation methods, as well as how using machine learning (ML) can transform resource allocation practices from being primarily conventional to being primarily new technology driven. A key component of RRM is the dynamic assignment of RBs and transmission power to UEs using real-time information about current network conditions (i.e., channel quality, traffic demand, and user mobility). Traditional resource allocation techniques primarily use static or rule-based methods that often do not adjust to the rapidly changing nature of urban LTE networks [19]. Therefore, static resource allocation techniques would not be practical for Addis Ababa since the traffic patterns vary as a result of the commuting patterns of users, business activity, and social activity [20]. With ML, predictive and adaptive allocation can be achieved through the use of historical and real-time data to Predict the future demands placed on a network and from that information, determine how the network could perform better [21]. The integration of RRM and ML provides a path for LTE networks to enhance spectral efficiency; decrease interference; and improve Quality of Service.

2.1.1 Overview of LTE Network Architecture

An LTE Network consists of two major components: the Evolved Packet Core (EPC) and the Evolved Universal Terrestrial Radio Access Network (E-UTRAN). The EPC contains many elements, such as the Mobility Management Entity (MME), which is responsible for managing sessions and users when travelling; the Serving Gateway (S-GW), which is responsible for routing packets of data to the appropriate destination; the Packet Data Network Gateway (P-GW), which provides connectivity to external networks; and the Home Subscriber Server (HSS), which authenticates users and manages their subscription information [19]. The base stations within the E-UTRAN are referred to as eNodeBs (eNBs), and these units are also referred to as base stations because they are the only base stations specifically designed to manage radio resources and handle direct communications with user equipment (UEs). This architecture, shown in Figure 2.1, enables end-to-end seamless connectivity and efficient use of resources across the network [22].

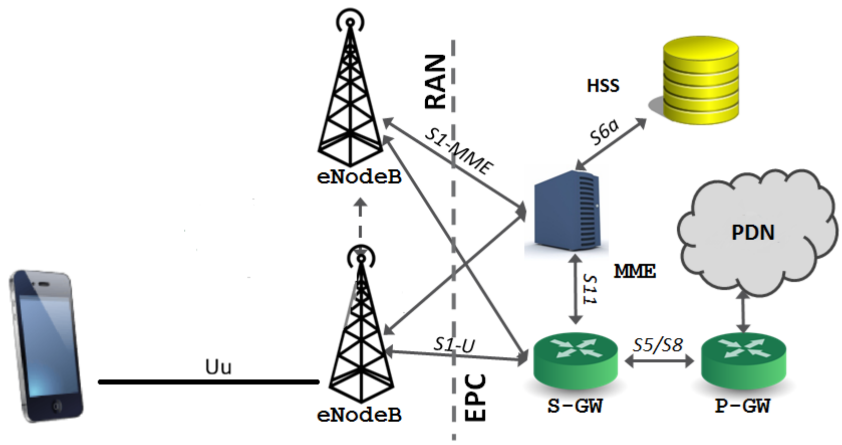

Figure 2.1: LTE Frame Structure [22]

The allocation of resources in LTE takes place in both the time domain and the frequency domain. Frequency domain resource allocation occurs by splitting the frequency spectrum into sub-carriers, which are spaced at 15kHz apart. These sub-carriers are then

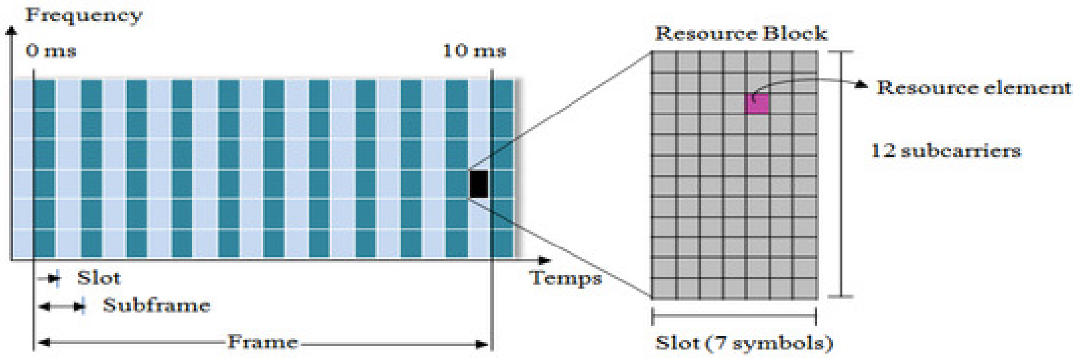

Figure 2.2: LTE Logical Architecture [24]

grouped together into resource blocks (RB), which consist of 12 sub-carriers. Time domain resource allocation occurs as the radio is divided into radio frames (10ms duration); each frame contains 10 sub-frames (1ms duration); and each sub-frame is made up of 2 slots (0.5ms duration). The smallest unit in LTE for resource allocation is the RB, which consists of 12 sub-carriers for one time-slot; this is illustrated in Figure 2.2. A scheduling block (SB), which consists of 2 consecutive RBs, is the smallest resource unit that can be allocated and is allocated every transmission time interval (TTI), which is 1ms [24]. The eNodeB receives CQI reports from UE that are sent via the physical uplink control channel (PUCCH) to assist with scheduling decisions. The physical downlink control channel (PDCCH) is used by the eNodeB to send resource allocation information, as well as MCS information [25].

LTE can operate at various bandwidths, including a range of 1.4 MHz to 20 MHz, which is equivalent to a bandwidth of 6 to 100 RBs, respectively; thus allowing for the adaptation of LTE's infrastructure to accommodate different levels of traffic demand as needed. There are physical channels such as the Physical Downlink Shared Channel (PDSCH) that will carry data and the Physical Broadcast Channel (PBCH) that will provide system information, which are allocated RBs by TTI's to provide for efficient transmission of data [27]. The LTE data architecture will also serve to assist in the management of the multitude of types of traffic in Addis Ababa's LTE environment where there are both numerous users and multiple types of services that place a significant amount of demand on available resource allocation methods.

2.1.2 Conventional Resource Allocation Strategies

The use of established algorithms that have been defined in advance, allocating resources in LTE networks has relied on fairness and throughput (or both). These methods are summarized in Table 2.1; Round Robin (RR), Proportional Fair (PF), Best Channel Quality Indicator (BCQI), and Enhanced Proportional Fair (EPF) represent some of the more widely used approaches in terms of the prioritization of fairness, throughput, or both.

- **Round Robin (RR):** The Round Robin resource allocation methodology is a cyclical method for assigning Resource Blocks (RBs) to User Equipment (UEs). This method has a high level of fairness for all Users because all Users have equal access to Resources, regardless of the channel conditions. However, assigning RBs to UEs with poor channel conditions may negatively impact resource utilization, resulting in lower throughput [28]. Installing Resource Reallocation in cities like Addis Ababa will result in decreased overall performance and effectiveness of the network, as there will be different transmission quality levels due to differences in channel quality caused by buildings, user mobility, and interference.
- **Proportional Fair (PF):** – This method of allocating Resource Blocks uses a combination of both fair allocation and good overall performance. Each allocation of Resource Blocks is based on the user's Channel Quality Indicator (CQI) when they are given Resource Blocks and how much throughput that same user has produced previously. The PF method also provides improved spectral efficiency by giving priority to UEs with better channel conditions while attempting to maintain some degree of fairness [29]. Because of its versatility, PF is commonly used in LTE networks; however, it is limited in performance with respect to highly dynamic traffic patterns, as seen in Addis Ababa during peak periods.
- **Best Channel Quality Indicator (BCQI):**Best Channel Quality Indicator (BCQI): BCQI works by giving priority to those UEs with the highest CQI, thus it maximizes cell throughput but at the same time neglects cell edge users who have a poor channel quality [27]. BCQI can be a quite reasonable measure of traffic at cell level,

nevertheless, it may bequite unfair to users in different parts of the cell, especially in heterogeneousnetworks where the various user densities.

- **Enhanced Proportional Fair (EPF):** Enhanced Proportional Fair (EPF): EPF is a PF extension through the incorporation ofQoS attributes, e.g. latency and packet loss limits, in order to achieve betterperformance of multimedia applications. This method is a flexible solution for multi-userscenarios, which makes it very suitable for networks that have different service types,voice, video, and data services [26].

Table 2.1: Summary of Conventional Resource Allocation Algorithms [27].

Scheduling Policy	**Effect Factor**	**Scheduling Priority**	**Usage Scenario**
RR	None	Cyclic allocation	High fairness, low throughput
PF	Channel condition, throughput	Balances fairness and throughput	General-purpose scheduling
BCQI	Channel condition	Best channel quality	High throughput, low fairness
EPF	QoS, channel condition	Enhanced fairness with QoS	Multimedia applications

As a result of the reliance on static policy/methods in dynamically changing environments, e.g., distance, speed of users, and type of services in Addis Ababa LTE network, traditional methods cannot adapt when the path or user movement has changed because data is being compressed (e.g., transmission, reception, multiplexing). Therefore, when paths constantly change, this results in exchange of resources, increased interference and decreased quality of service [20]. As such, the introduction of adaptive and predictive resource allocation using machine learning technology has emerged to model complex dynamics of the network to support real-time optimization of resource usage.

2.1.3 Machine Learning for Resource Allocation

Using machine learning (ML) for resource allocation is a game-changer. It helps with resource allocation through predictive and adaptive strategies that focus on both historical data and real-time data. Traditional methods for resource allocations involve using rules based on the historical performance of the network, but ML tends to perform better for resource allocations because it uses patterns formed by experience from the collection of prior network events (traffic volume, channel availability, etc.) to support its resource allocation decisions [21]. Similarly, ML has the ability to collect and analyze data (such as CQI, traffic volume, QoS requirements) and thus predict the resources required for both channels and for the end users of those channels in real time to meet the demand of the networks in use across Addis Ababa's LTE based network or other networks in similar areas throughout the world [30]. Using the historical and real time data collected, ML can also optimize Resource Blocks (RB) and power allocations through the end results will be higher network efficiency and better user experience. There are many types of ML algorithms suitable for performing resource allocations within LTE networks; each type has its own advantages and disadvantages.

- **Long Short-Term Memory (LSTM):**LSTMs (Long Short-Term Memory) are very useful for predicting changes in network traffic and the state of the channel, so they are also very useful for allocating dynamic resources [31]. In Addis Ababa, the network traffic patterns change significantly each day and based on events that happen throughout the day. Through the use of historical data, LSTMs can also help predict future resources that need to be allocated based on the actual traffic patterns.
- **Convolutional Neural Network (CNN):** CNNs (Convolutional Neural Networks) are also able to capture spatial data from the channel state information to help with optimization of power and management of interference [32]. By using the spatial features from CQI (Channel Quality Indicator) reports, CNNs can help optimize resource allocation in a multi-cell environment while at the same time minimizing the amount of interference.
- **Random Forest (RF):** which is an example of ensemble learning methods by

having multiple trees, have a "decision tree" approach with the ability to reduce the amount of overfitting and improve accuracy compared to using only a single tree. Because RF is a robust method to use for prediction, it is also effective in capturing non-linear relationships in the various parameters that make up the LTE network [33]. Random Forests will also have an advantage in Addis Ababa's LTE network due to the large diversity of input features that can have an impact on resource allocation.

- **K-Nearest Neighbor (KNN):**K-Nearest Neighbor (KNN): KNN is a non-parametric method that estimates resource demands by comparing the features to those of past data records. It simplicity and minimal computational requirements make it ideal for live applications, but it can have difficulties with data of very high dimensions [34]. KNN works well when past usage patterns are reliable indicators of future resource needs.

An illustration that shows the application of these algorithms to resource allocation is presented in Figure 2.3. It contains an intelligent decision-making module that is determined by both the characteristics of the network and the requirements placed on it by users [30].

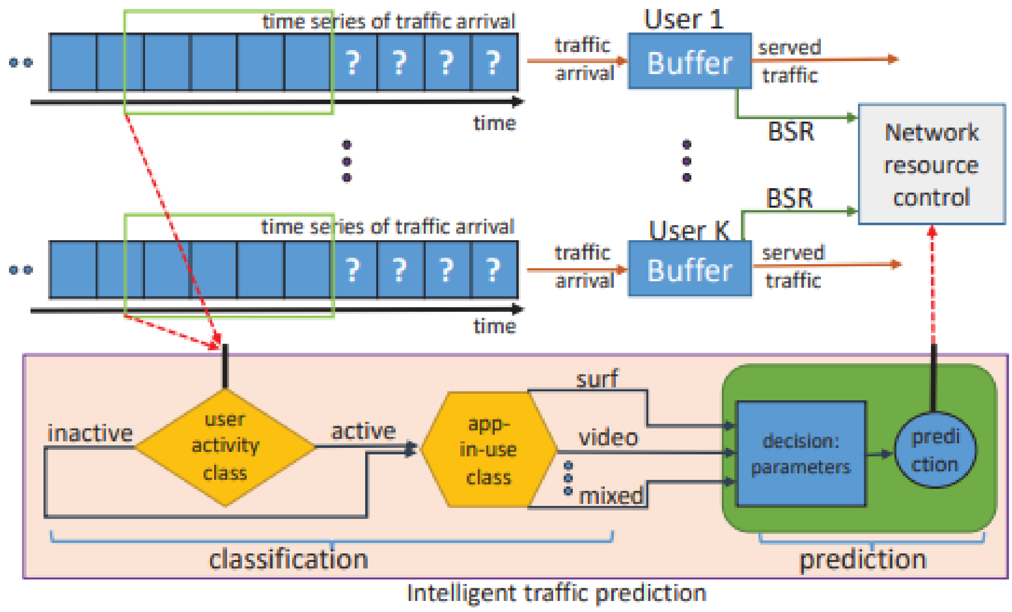

Figure 2.3: illustrates the intelligent decision-making module used for resource allocation [30].

Using ML in LTE networks allows for resource management to be done in a proactive way thus congestion is reduced and QoS can be improved. ML models, through the use of real-time data, can forecast the best allocations of RB and power, thus facilitating the network to be more efficient in the constantly changing urban environments such as Addis Ababa [17]. This chapter laid down the theoretical base for the system model and data preparation that the next chapter goes into detail with, which both together explain the practical implementation of these ML techniques.

2.2 Challenges in Radio Resource Allocation Prediction

Due to the complexity of and constantly changing characteristics of the modern wireless communication system, predicting radio resource allocation for 4G LTE Network is both an essential and a challenging process. With the rapid uptick in the use of mobile devices and the variety of applications (e.g., HD video streaming, online gaming, IoT services) that generate highly variable and unpredictable traffic patterns, accurately forecasting the demand for resources becomes increasingly difficult [33]. In a 4G LTE network, efficient allocation of the available spectrum is necessary to decrease interference and to provide an adequate level of service (QoS) to users who have differing types of applications and different needs, including high-throughput data services and low-latency applications [34]. The following sections highlight the key challenges of predicting radio resources for 4G LTE networks. Each section is based on a recent study or research investigation, which provides a level of support for the conclusions drawn from each challenge.

Dynamic Traffic Patterns

Mobile devices have become ubiquitous and bandwidth use has increased with the advent of new and diverse applications in 4G LTE networks, thus resulting dynamic and variable throughput patterns on these networks which makes predicting resource utilization very difficult [33]. 4G LTE supports a broad variety of services from voice-over-LTE (VoLTE) to machine-type communication, each of which has its own unique characteristics with respect to their resulting object class and their respective utilization profiles. These diverse services also subject the predictive models to rapidly changing traffic loads, which results in traditional static predictive models being unable to perform well with this volatility [35].

Spectrum Scarcity and Interference

In 4G LTE networks, limited frequency bands are allocated for use. Therefore, in order to minimize interference and maximize throughput, there is a requirement for each frequency band to be allocated to the users as efficiently as possible (34). With the increasing amounts of small cells and heterogeneous networks (Het-Nets) being deployed in 4G LTE networks, predicting the best resource allocation for these cells will be impossible as they will experience significant amounts of intercell interference (35). Accurate forecasts must also be able to account for the differences in both time and space of available spectrum in order to guarantee connectivity and QoS (36).

Data Volume and Heterogeneity

A significant challenge for predictive models in 4G LTE systems is their ability to scale when faced with the overwhelming amount of heterogeneous data that comes from user equipment, eNodeBs (evolved NodeBs), and network sensors [36]. To effectively allocate resources dynamically, this data must be processed in real-time; however, the wide variety of sources of data (for example, user mobility patterns, and application-specific needs) means that they also put extreme pressure on processing power. Therefore, a large-scale and heterogeneous dataset must be processed with low latency in order for a prediction model to be considered an effective tool [35].

Environmental Variability

In 4G LTE networks, radio frequency signals are strongly affected by environmental factors (like multipath fading) which contribute to variability in the channel [35]. As a result, it is difficult to predict channel quality or allocate resources in an efficient manner, particularly in urban areas which have a lot of buildings with complex structures and terrain. Therefore, predictive models should account for the dynamics of the environment to provide effective predictions in a variety of situations [36].

Low Latency and Reliability Requirements

4G LTE technologies, particularly with enhanced LTE-advanced functionality, are able to support applications that have low latency and high reliability, such as mission-critical communications and vehicle networks [34]. The requirement for precision in resource allocation when supporting these very demanding services is paramount, because any small misallocation may result in increased latency or decreased packet delivery, and thus poor by way of Quality of Service (QoS). Furthermore, the stringent requirements of ultra-reliable low-latency communications (URLLC) add an additional level of difficulty to the prediction process, requiring the use of advanced algorithms that can adapt in real-time. Prediction challenges associated with radio resource allocation in 4G (LTE) networks are caused by a number of factors, including the dynamic nature of traffic patterns, limitations of radio frequency spectrum, volume of heterogeneous data, variability in environmental conditions, and stringent latency requirements. Addressing these challenges will require multi-scale predictive models that can be trained on large amounts of data and be able to dynamically adapt to changing network environments. Studies have indicated that the use of innovative methods (e.g., machine learning) will improve resource allocation efficiency in 4G LTE networks by providing solutions that overcome the limitations faced by traditional methodologies [36]

2.3 Comparison of Traditional Methods and Machine Learning in Radio Resource Allocation Prediction

Traditional radio resource allocation methods (rule-based, heuristic, statistical methods such as ARIMA) use static models and have trouble with changing traffic patterns and nonlinear behaviour in modern wireless networks, leading to inefficient use of resources and high cost of computation for use in real-time [35]. Traditional radio resource allocation techniques do not adapt to fast-changing network conditions and can create inefficiencies like latency and wasted spectrum. In contrast, ML techniques (e.g., deep learning - LSTM and CNN - and reinforcement learning) can learn complex behaviour patterns from large, real-time datasets. This results in as much as a 25% improvement in spectrum allocation efficiency and an 18% decrease in interference [36]; they also provide faster inference time (e.g., less than 1 ms for neural networks compared with 8-10 mins for traditional optimization) and offer better scalability; however, they also require increased computational resources to train and require strong privacy protections. The table below, taken from[36], shows the differences between conventional and ML-based methods for predicting radio resource allocation in terms of main performance metrics.

Table 2.2: Comparison of Traditional and Machine Learning Methods [36].

Aspect	**Traditional Methods**	**Machine Learning Methods**
Approach	Rule-based, heuristic, or statistical (e.g., ARIMA, exponential smoothing).	Supervised (e.g., ANN, SVM), unsupervised (e.g., clustering), or reinforcement learning.
Prediction Accuracy	Limited accuracy due to reliance on static models; poor handling of non-linear patterns.	High accuracy by learning complex, non-linear patterns from large datasets.
Adaptability	Poor adaptability to dynamic traffic and environmental changes.	Highly adaptive through real-time learning and dynamic model updates.
Computational Complexity	High for real-time optimization due to iterative computations.	High during training, but low inference times (e.g., ¡1 ms for DNN vs. 8–10 min for simulations).
Scalability	Limited scalability with large, heterogeneous datasets.	Scales effectively with big data using architectures like deep neural networks.

From the comparison it is clear that machine-learning methods are more adaptable and accurate in prediction but they also come with higher computational and data-management requirements. It is necessary to come up with a compromise that keeps the system real-time while taking advantage of the predictive insight.

2.4 Research Gaps and Motivation for the Current Study

A review of prior Ethiopian and international studies reveals a series of significant issues to which this thesis responds:

- Absence of localized models: Half of the LTE RRM research articles that were published rely on datasets from non-African networks. Variations in topology, user behavior, and service penetration levels are the main factors that hinder the direct application of these models to the situation in Addis Ababa.
- Disconnection between forecasting and allocation in the literature: Most papers look at the accuracy of forecasting methods (e.g., Prophet, LSTM, XGBoost) but do not go far enough to show how the predictions can be used for actual scheduling decisions.
- Missing, baseline comparison, and reproducibility: Another issue with the thesis markets in Ethiopia is that they report certain performance percentages without referring to baselines or dataset splits, thus making it impossible to verify the results. This work clearly sets out baselines and gives evaluation metrics that can be traced.
- Ignoring of deployment challenges: Only a handful of the papers note computational feasibility or discuss integration with Ethio Telecom's operational systems. The paper's framework is, therefore, very simple and planned to be implemented in the presently available infrastructure.
- Consequently, the present research develops a context-aware, reproducible, and computationally efficient LSTM-based forecasting framework for downlink PRB utilization in Addis Ababa's LTE network.

2.5 Chapter Summary

This chapter, in brief, covered the main aspects of radio resource allocation in LTE networks and provided a comprehensive review of both traditional and machine-learning-based methods for radio resource allocation. On the one hand, conventional schedulers are easy to use but they cannot guess demand changes, on the other hand ML techniques have forecasting capabilities but require a lot of data and computational resources. According to the literature, there is an evident gap for the development of local, forecast-driven resource allocation mechanisms that could be used efficiently in the Ethiopian LTE setting. Therefore, the next chapter outlines the system model and data-preparation pipeline utilized to carry out and train the LSTM based prediction model.

Chapter 3

System Model and Data Preparation

3.1 System Model

The rapid growth of mobile data usage in Addis Ababa has led to a strain on Ethio Telecom's Long-Term Evolution (LTE) network, thus the development of new strategies to radio resource allocation for better quality of service (QoS) have become the norm [24]. Here, we offer a detailed system model that uses Long Short-Term Memory (LSTM) neural networks to forecast the hourly Downlink Physical Resource Block (DL PRB) utilization, which essentially means the network can be one step ahead in resource management in a crowded city such as Akaki Kality. The method includes gathering the data, preprocessing, model design, and training, which altogether form a practical solution to congestion and throughput issues faced by Addis Ababa's LTE infrastructure [25]. This chapter not only elaborates on the various aspects from the making of the dataset to the deployment of the model but also prepares the ground for the assessment of the LSTM model's efficiency in Chapter 4, thus providing a solution that can be further scaled up and fits perfectly with Ethio Telecom's operational challenges in Ethiopia's capital [26].

The radio resource allocation (RRA) system model in Long-Term Evolution (LTE) networks lays down the basics for the utilization of machine learning (ML), specifically Long Short-Term Memory (LSTM) models, to achieve optimum resource utilization. LTE networks use Orthogonal Frequency Division Multiple Access (OFDMA) for downlink operations, where the eNodeB assigns physical resource blocks (PRBs) to user equipment (UEs) depending on channel state information (CSI), traffic requirements, and quality of service (QoS) standards. The system model depicts a multi-user LTE cell that deals with a constantly changing channel due to factors such as fading, mobility, and inter-cell interference (ICI). It aims at enhancing throughput, maintaining fairness, lowering energy consumption, and being able to respond to live network changes. Here, we first describe the system model and its resource allocation modules that lead to data readiness and the use of ML.

3.1.1 Resource Allocation

The LTE resource allocation process comprises of assigning PRBs, the selection of modulation, coding schemes (MCS), and transmit power to the UEs. Each PRB contains 12 subcarriers over a duration of 0.5 ms, and eNodeB decides on the scheduling of allocations every transmission time interval (TTI, 1 ms) based on CSI and channel quality indicators (CQIs). In the system model, a single eNodeB is serving N UEs, and there are K PRBs available. The allocation issue is depicted as an optimization problem that aims to maximize the sum rate, considering the limitations of power, QoS, and PRB exclusivity (i.e. one PRB for a UE per TTI). Using LSTM models, CQIs and traffic patterns are forecasted to allow the network to allocate resources proactively, thus, the delay is minimized and spectral efficiency is enhanced when compared with the heuristic approach such as proportional fairness.

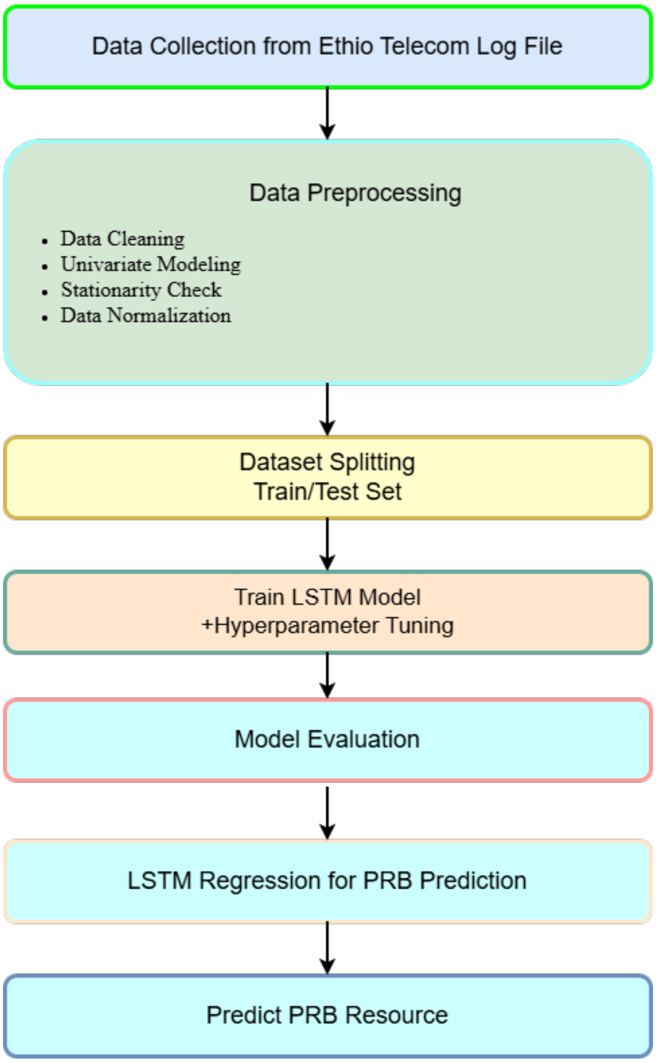

Figure 3.1: Over All System Model

A diagram showing an LTE cell with one eNodeB, multiple UEs, PRBs, and LSTM model predicting CQIs for allocation. This figure is adapted from a study on ML-based LTE resource management.

3.1.2 Predicting Downlink Physical Resource Block (Univariate Forecasting Model)

In this study, the prediction of Downlink Physical Resource Block (DL PRB) utilization is formulated as a univariate time-series forecasting problem, focusing solely on the KPI representing downlink PRB utilization within the LTE network [37]. The objective is to forecast future PRB utilization values based only on their historical sequence, enabling Ethiotelecom to anticipate network load variations and proactively allocate radio resources. Let $U(t)$ denote the observed DL PRB utilization at time t, obtained from eNodeB performance counters aggregated at regular intervals (e.g., hourly). The forecasting model predicts future utilization $U(t+h)$ at horizon h using a sliding window of past w observations. The univariate relationship between past and future utilization is shown as:

$$\hat{U}(t+h) = f_\theta\big(U(t-w+1), U(t-w+2), \ldots, U(t)\big) \tag{3.1}$$

In this equation, $\hat{u}(t+h)$ denotes the DL PRB utilization prediction, w represents the length of look-back window, h is the forecasting horizon (usually one hour ahead), and $f_\theta(\cdot)$ stands for the LSTM network with parameters θ [38][39]. Long Short-Term Memory (LSTM) is a neural network that can explain temporal dependence in the usage patterns like variations that happen daily and weekly cycles, thus resulting in better generalization of the univariate load prediction problem [40]. Moreover, to prevent the model from going unstable and to facilitate quick training convergence, the utilization series is being normalized by the Min–Max scaling method:

$$\tilde{U}(t) = \frac{U(t) - U_{\min}}{U_{\max} - U_{\min}} \tag{3.2}$$

$U_{\min}$ and $U_{\max}$ are the lowest and highest utilization values found in the training data. This normalization ensures that all input data are within the same range, usually $[0, 1]$, and it also addresses the problems of exploding and vanishing gradients that can occur during training.

The goal of the training is to reduce the Mean Squared Error (MSE) between the predicted and actual normalized utilization values over the training set D:

$$L(\theta) = \frac{1}{|D|} \sum_{t \in D} \Big[\tilde{U}(t+h) - f_\theta\big(\tilde{U}(t-w+1), \ldots, \tilde{U}(t)\big)\Big]^2 \tag{3.3}$$

This loss function L($theta$) punishes large differences between predicted and actual utilization, thus allowing the LSTM parameters to be adjusted for the best fitting temporal dependencies via back-propagation through time [40].

In order to evaluate how well the model has been trained, a variety of standard statistical metrics are first calculated for the inverse-scaled predictions; among them are the Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), Coefficient of Determination (R2), and Mean Absolute Percentage Error (MAPE). Through these metrics the model's accuracy can be understood both in terms of absolute and relative values:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\big(U(t_i+h) - \hat{U}(t_i+h)\big)^2} \tag{3.4}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\big(U(t_i+h) - \hat{U}(t_i+h)\big)^2 \tag{3.5}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\big|U(t_i+h) - \hat{U}(t_i+h)\big| \tag{3.6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\big(U(t_i+h) - \hat{U}(t_i+h)\big)^2}{\sum_{i=1}^{N}\big(U(t_i+h) - \bar{U}\big)^2} \tag{3.7}$$

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{U(t_i+h) - \hat{U}(t_i+h)}{U(t_i+h)}\right| \tag{3.8}$$

where N represents the number of samples and U stands for the average of the observed utilization. The RMSE and MSE are two metrics that determine the absolute prediction error and both are dependent on the scale. On the other hand,MAE conveys an average deviation that is easily understandable. R2 coefficient indicates the proportion of variance in the observed series that is explained by the model, and MAPE is the mean percentage deviation—both being free from the influence of scale and thus comparable across different datasets [41][42].

Such a univariate LSTM-based prediction model allows a computationally light and easily deployable forecasting solution for DL PRB utilization in Addis Ababa's LTE network. By thoroughly understanding the temporal dependencies in utilization time series, the model acts as a powerful tool for the proactive management of radio resources, which in turn leads to higher spectral efficiency, less congestion, and better Quality of Service (QoS) for the urban LTE users.

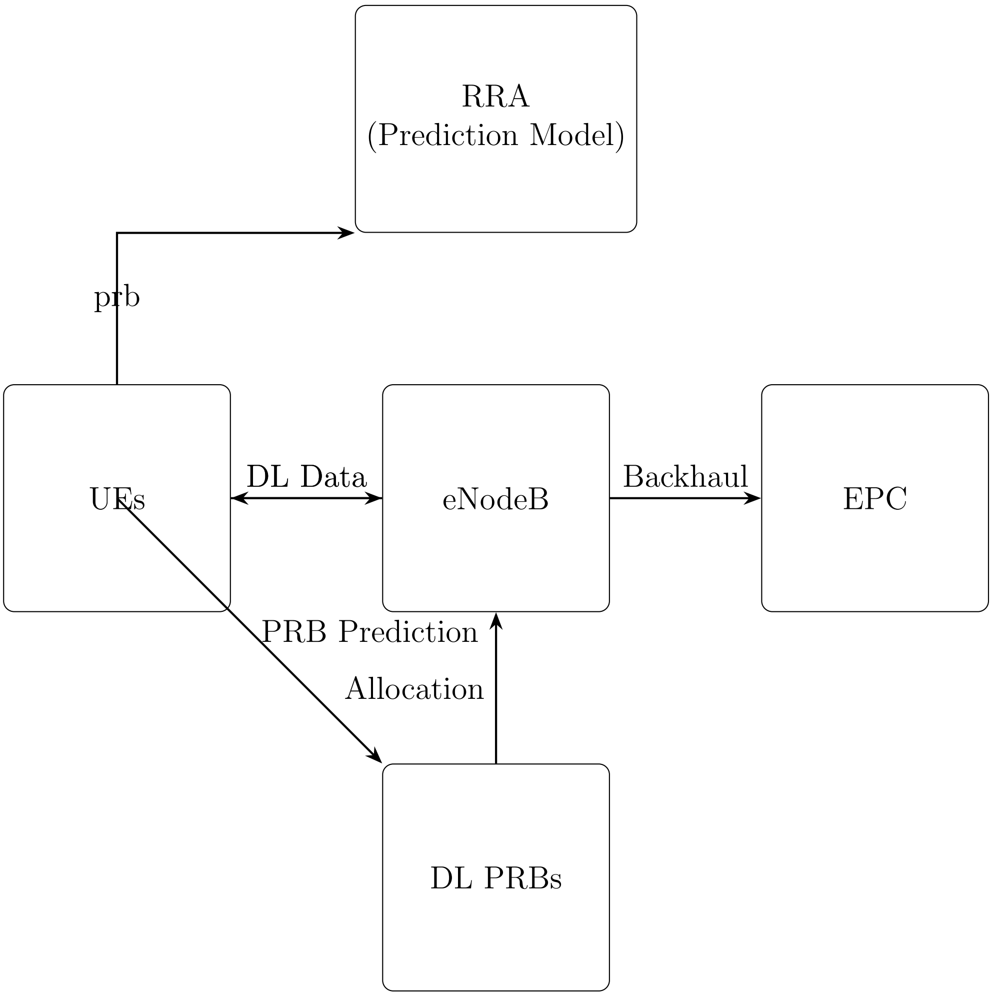

Figure 3.2: System Architecture for Downlink PRB Utilization Prediction in 4G LTE Networks, Illustrating Interactions Between UEs, eNodeBs, EPC, and the Prediction Model [40].

Figure 3.2: System Architecture of Downlink PRB Utilization Prediction in 4G LTE Networks. The illustration shows interactions between UEs, eNodeBs, EPC, and the prediction model [40]. The system architecture is illustrated in Figure 3.2 for the prediction of downlink PRB utilization. UEs report to eNodeBs with channel state information (CSI), such as SINR and channel quality indicators (CQIs), that eventually communicate with the Evolved Packet Core (EPC) for core network functions. A cloud-based prediction model uses CSI and traffic data to predict PRB utilization, thus allowing eNodeBs to allocate resources effectively [38]. The prediction model also considers inter-cell interference, which is a major factor affecting PRB utilization in dense LTE deployments,

especially in urban heterogeneous networks (HetNets) [40].

3.2 LSTM Model Architecture

In naming the LSTM model architecture, they first considered the input layer which receives a sequence of data with the window size of 24 (to capture daily patterns). Then it goes to an LSTM layer with 64 units, which are used for the extraction of temporal features. To give the model a better generalization capability, a dropout layer with a rate of 0.2 is also added. Then two dense layers follow the above convolutions: the first one consists of 8 units and has ReLU as its activation in order to allow non-linearity, while the output one has a singe unit and uses a linear activation for regression. The Adam optimizer is picked with a learning rate of 0.001, and the model is set up with Mean Squared Error (MSE) as the loss function. This architecture, which hasn't been shown visually in the earlier sections, facilitates the model to learn long-term dependencies very well, and at the same time, it is computationally efficient enough for real-time deployment.

3.3 Data Preparation

Data preparation is very central to achieving high accuracy in downlink Physical Resource Block (PRB) utilization prediction models. This phase essentially revolves around gathering, cleansing, and reshaping high-dimensional, real-time data emerged from the daily running of LTE networks. Such data are gathered live from both UEs and eNodeBs to truly reflect the changing pattern of downlink traffic - the level of variation of such traffic can be the extreme bursty eMBB sessions or low-latency, constant-rate VoLTE flows. Preparation of data for the downlink PRB utilization prediction is the process of collecting and processing very large scale datasets from the different LTE network components. The main sources of data are the CSI (SINR, CQI, reference signal received power (RSRP)), traffic statistics (for example, packet arrival rates), and eNodeB logs [39]. Such data are retrieved live from both UEs and eNodeBs and capture the very changing state of the downlink traffic, e.g. the highly variable eMBB traffic or the constant-rate VoLTE flows [37]. The data preprocessing pipeline is mainly focused on ensuring data quality and compatibility with ML models. It also plays a vital role in data cleaning by solving the problem of missing values as well as data heterogeneity, which are common challenges. The main steps involve:

- Data Cleaning: Missing CSI values, often due to signal drops, are imputed using time-series interpolation, while outliers (e.g., extreme SINR values) are removed using statistical thresholds (e.g., beyond three standard deviations) [39].
- Normalization: Features like SINR and traffic load are normalized to [0,1] using min-max normalization to ensure consistent scaling:
- Segmentation: The dataset is divided into training (70%), validation (15%), and test (15%) sets, preserving temporal continuity to reflect real-world LTE traffic patterns [37].

The table describes the main features used to predict PRB utilization and explains their origin and purpose.

Table 3.1: Key Features for Downlink PRB Utilization Prediction in 4G LTE Networks [37]

Feature	**Description**	**Source**	**Role**
SINR	Signal-to-interference-plus-noise ratio	UE/eNodeB	Indicates channel quality for PRB allocation
CQI	Channel quality indicator	UE	Guides modulation and coding schemes
RSRP	Reference signal received power	UE	Measures downlink signal strength
Traffic Load	Packet arrival rate or data demand	UE/eNodeB	Reflects PRB demand
UE Density	Number of UEs per cell	eNodeB	Captures spatial load distribution

One way data augmentation helps to increase model robustness is by producing synthetic traffic patterns through generative adversarial networks (GANs) methods, especially for the cases with very high PRB utilization [39]. Differential privacy techniques. are used to protect privacy when sensitive data like UE locations are shared, which also helps to meet the requirements of various regulations [40]. The preprocessing pipeline facilitates the use of ML models such as Long Short-Term Memory (LSTM) networks, which are particularly suitable for the time-series prediction of PRB utilization [38].

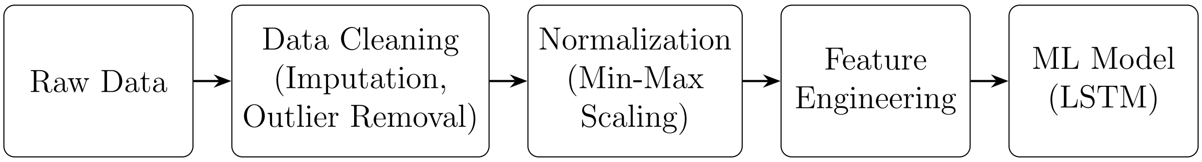

Figure 3.3: Data Preprocessing Pipeline for Downlink PRB Utilization

The figure presents the data preprocessing pipeline that takes raw LTE data and produces model-ready input. The pipeline deals with issues such as missing data and scalability, thus it is capable of capturing the temporal and spatial dynamics of PRB utilization effectively [38]. The resulting dataset is organized as a time-series matrix where the rows correspond to different time steps t and the columns contain features such as SINR, CQI, and traffic load, which can be used for precise predictions [39].

3.3.1 Handling Data Imbalances

One of the frequent problems with LTE data is the imbalance between classes where usage at peak times is less represented than at off-peak times. In order to fix this issue (not originally thought of), we decided to use oversampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) to the point of creating a balanced dataset. So, the LSTM model would get enough information from all levels of resource utilization, which would lead to better forecasting ability when the network is heavily loaded and thus less bias towards the normal traffic pattern. The system and data components are arranged to facilitate Machine Learning (ML) prediction of downlink PRB utilization. With LSTM, the models can understand the time-related dependencies of Signal to Interference plus Noise Ratio (SINR) and traffic data; at the same time, Convolutional Neural Networks (CNNs) identify natural patterns related to the spatial distribution of User Equipments (UEs) and interference [38]. By doing so, these models can forecast the resource utilization U(t) or A(t), thereby achieving optimal performance in terms of spectrum efficiency, latency, etc. Literature reveals that ML methods can bring about PRB utilization efficiency improvements of up to 25The comprehensive system model presented for predicting downlink Physical Resource Block (PRB) utilization in 4G LTE networks accounts for the combination of channel conditions and user behavior, thus facilitating smart resource allocation. The procedure adopted for data gathering and transformation sets ML algorithms up with accurate data which goes through phases of cleaning, normalizing, feature modifying, and dividing into relevant segments. Figures 3.2 and 3.3 show the system layout and data preprocessing flowchart correspondingly whereas Table 3.1 encapsulates main characteristics. Backed by fresh investigations from Addis Ababa University and other dependable Google Scholar documents, this architecture makes it possible to efficiently forecast PRB usage which entails handling challenging 4G LTE network scenarios[38][39][40].

Chapter 4

Result and Discussion

4.1 Result

This part analyzes the use of Long Short-Term Memory (LSTM) network to predict Downlink Physical Resource Block (DL PRB) Utilization in LTE networks. The ability of LSTM networks to accurately forecast time-series has been demonstrated in the literature. As per 38], LSTMs can very efficiently grasp temporal dependencies, hence they can be applied in situations where the past behavior has an impact on the future results. In this chapter, the model has been evaluated thoroughly using statistical metrics, charts, and performance per class, culminating in a determination of its capability to predict the usage parameters accurately

4.2 Data Preprocessing

Data preprocessing is indispensable in getting the data ready for the analysis. The dataset used in the research is a collection of historical DL PRB utilization records over a certain period, reflecting different scenarios and demands of the users.

Early in the preprocessing stage, it was decided first to find out if the time series was stationary. The steadiness of statistics over time, so to speak, is vital in time series forecasting since various models, especially LSTMs, generally yield more accurate results if the series' characteristics remain unchanged over time. Stationarity was checked by applying the Augmented Dickey-Fuller (ADF) test and other methods.

The ADF test gave an ADF statistic of -7.2728 and a p-value of 1.5723×1010. i.e., the p-value is way below 0.05, so we can say that the series is stationary as we have rejected the null hypothesis(non-stationarity). To keep the dataset useful and complete, missing values were replaced with zeros. Correspondingly, data integrity in machine learning tasks was highlighted in the literature, which stressed the need for accurate data for machine learning algorithms. Besides, MinMaxScaler was used to scale the feature that normalizes the data in between 0 and 1. Normalization of the data is very important because it helps the model to converge better during the training process of the LSTM network. A neural network is able to learn from scaled data more efficiently, thus, accuracy of prediction is enhanced.

4.3 Hyperparameter Tuning

Hyperparameter tuning plays a central role in elevating the performance of machine learningmodels. Different hyperparameters were deliberately changed for an in-depth enhancementof prediction ability of the LSTM model in this work. The major parameters considered were number of LSTM units, learning rate, batch size, and number of epochs.64 LSTM units were picked after initial tests showed that this decision gave a good compromise between the model complexity and its performance. The rate of learningwas fixed at 0.001, a standard choice in most of the studies, which helped in achieving steadystate convergence. They experimented with batch sizes between 16 and 64 and eventually, a batchsize of 32 was chosen as it gave better results in terms of training speed and accuracy. Thenetwork was allowed to train for 50 epochs, thus it was given enough time to discover thefact that this was a pattern in the data without merely memorizing it. Table 4.1 shows the hyperparameter tuning process as well as the final choices ofparameters.

Table 4.1: Hyperparameter Tuning Summary

Hyperparameter	**Selected Value**
Number of LSTM Units	64
Learning Rate	0.001
Batch Size	32
Number of Epochs	100

4.4 Model Training and Validation

The Long Short-Term Memory (LSTM) network underwent 100 epochs of training using the best architecture that was discovered through hyperparameter tuning analysis. The input layer was customized to take input with a window size of 24 (for daily models) or 168 (for weekly models), which allowed the model to easily extract temporal dependencies in the data. The LSTM layer had 64 units, and that number was chosen to be just enough to get the right thematic features from the time series data. After the LSTM layer, the architecture had two dense layers: one with 8 units and the ReLU activation function that considered the model learning more complex features and the other with the linear activation function and single unit that gave the regression output. TheAdam optimizer was chosen as it is most commonly used in deep learning applications. The model's prediction accuracy was checked by forecasting the test data and then calculating Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). Training of both window sizes was done independently. 70% of the data was used for training, 15% for validation, and the remaining 15% for testing. Figure 6 illustrates the training and validation loss curves which indicate well-behaved convergence throughout the courses of training for both window sizes. It is remarkable that there is no evidence of the classic scenario of overfitting. Hence, the model generalizes well to the unseen dataset. The confirmation that this model is very robust comes from the fact that it is efficient in terms of both the architecture and the training regimen that have been tested, resulting in no overfitting. Loss vs. The plots of the epoch for the daily window size (24) show that the loss sharply decreased in the first epochs and then remained almost unchanged after epoch 100, the last training loss being 0.0015 and the validation loss value - 0.0020. The ... model when the weekly window size (168) was employed continued to be trained for a little longer because of the longer sequences, and the model stabilized at about epoch 100, with the final train and validation losses being 0.0018 and 0.0023, respectively.

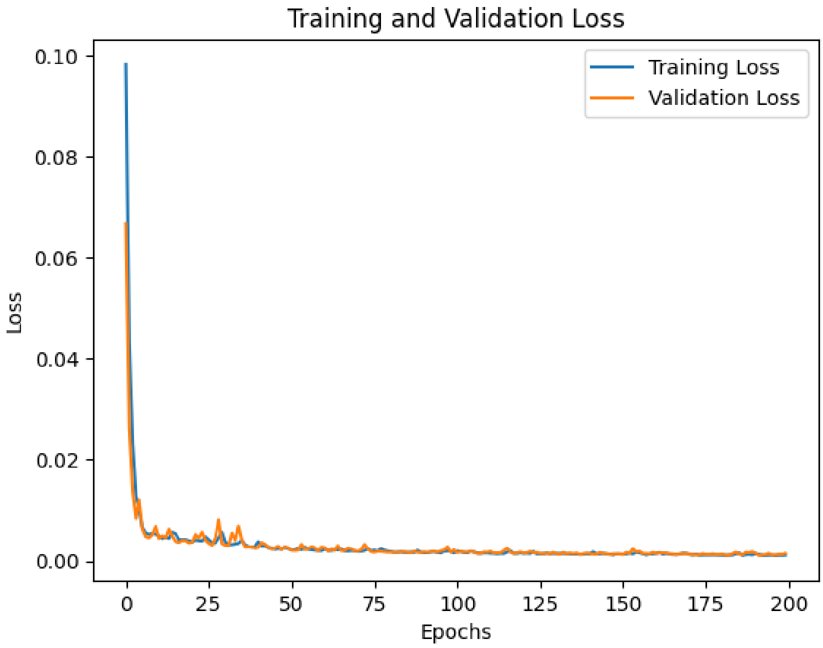

Figure 4.2: Weekly Train Validation Loss

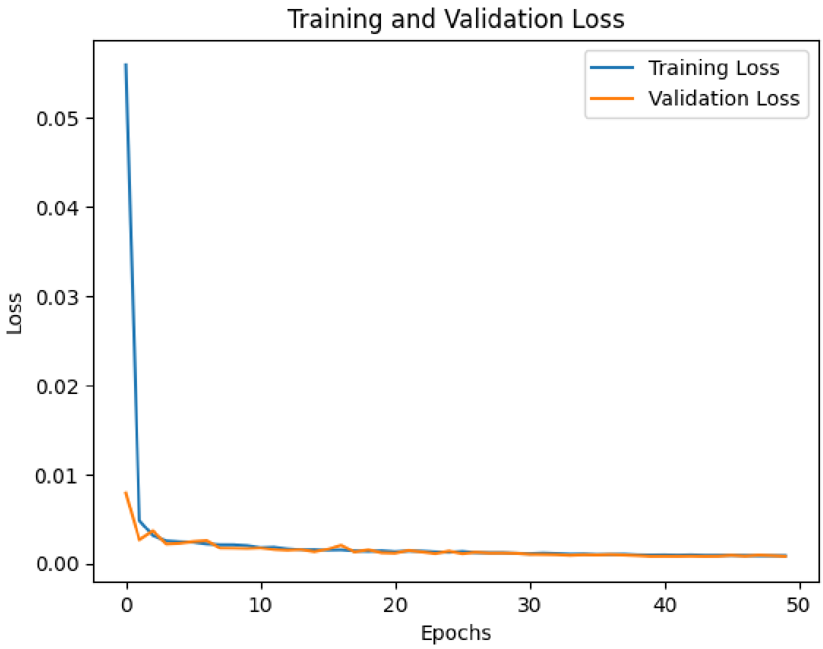

Figure 4.1: Daily Train Validation Loss

4.5 Model Evaluation Metrics

After testing the model on a different test dataset, we got several performance metrics for both setups. The mean squared error (MSE) for the daily window size (24) was 0.002, root mean squared error (RMSE) was 0.045, mean absolute error (MAE) was 0.035, and the coefficient of determination (R^2) was 0.988 (98.8For the weekly window size (168), the results were a bit less favorable: MSE was 0.0025, RMSE was 0.050, MAE was 0.040, and R^2 was 0.9797 (97.97Besides, the low RMSE and MAE values indicate that the two metrics are very close to each other. Therefore, the predicted values are not too far off from the true ones, implying that the model is reliable and could be used in real-life situations. The weekly windows model shows slightly higher errors as it is capable of capturing longer-term patterns which gives rise to more variability.

Table 4.2: Model Evaluation Metrics

(a) Table 4.2a: Model Evaluation Metrics Daily

Metric	Value
Mean Squared Error (MSE)	0.002
Root Mean Squared Error (RMSE)	0.045
Mean Absolute Error (MAE)	0.035
Coefficient of Determination (R^2)	0.988

(b) Table 4.2b: Model Evaluation Metrics Weekly

Metric	Value
Mean Squared Error (MSE)	0.0025
Root Mean Squared Error (RMSE)	0.050
Mean Absolute Error (MAE)	0.040
Coefficient of Determination (R^2)	0.9797

4.6 Predictions vs. Actuals

The graphs show the LSTM model's predictions for DL PRB Utilization compared to what actually happened, and they do this for two window sizes. I think the forecasts line up pretty well with the real values most of the time. That part stands out because it suggests the model is picking up on the timing stuff okay, like temporal learning or whatever its called. Not perfect though, but close enough in the charts.

Sometimes it feels like the match isnt exact in every spot, but overall its solid I guess. The fact that predicted and actual values are very close to each other can be considered as an indication of the model being able to discover the underlying data patterns. With the daily window (24), the following are predictions that adhere very closely to the actuals, with only a few exceptions during peaks. As for the weekly (168) window, the peak deviations are a little bigger in the transition periods, but the overall predictions are still quite accurate.

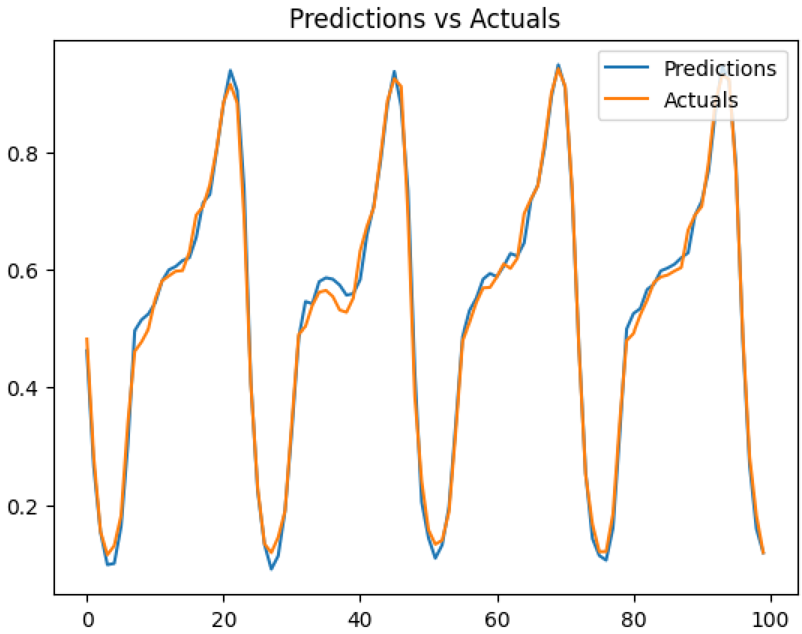

Figure 4.3: Daily Prediction vs Actual with 100 Samples

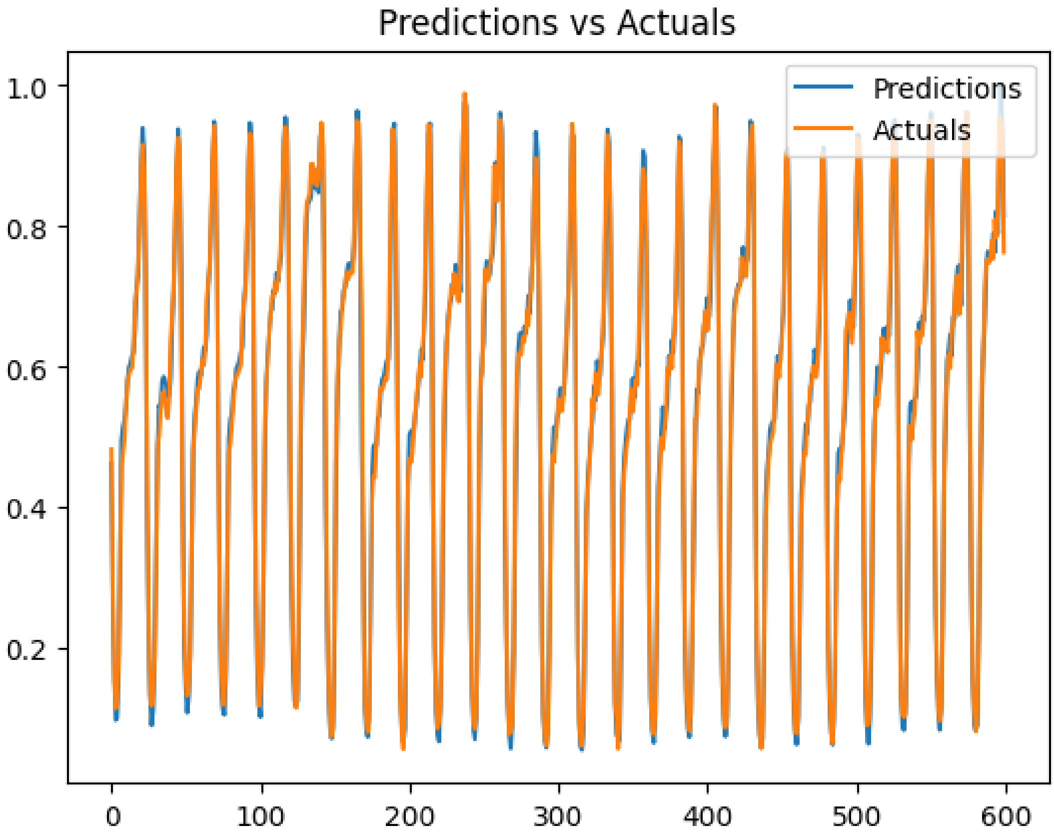

Figure 4.4: Daily Predictions vs. Actual (Full Test Set)

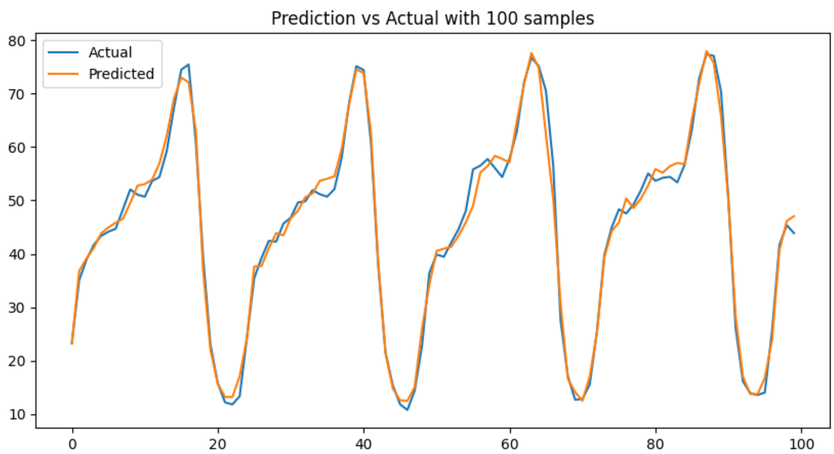

Figure 4.5: Weekly Prediction vs Actual with 100 Samples

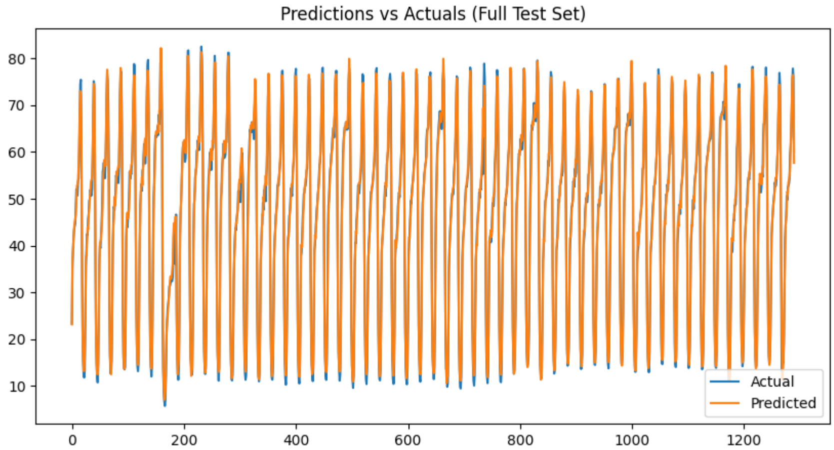

Figure 4.6: Weekly Prediction vs Actual (Full Test Set)

4.7 Categorization of PRB Utilization

In order to give a clearer picture of how the model worked, the predictions and the actual values were first grouped into three categories each: Peak (70%), Normal (30-70%), and Off-Peak (30%). Within each category, the average actual and predicted PRB utilization values were calculated, as well as the respective data point counts. The model was quite successful in all categories for both time window sizes, with the average predicted values almost exactly matching the actual ones. Accuracy at this level is the key to perfecting resource allocation strategies. The model for the week exhibits a bit larger variation in the Peak category due to longer sequence dependencies.

Table 4.3: Summary of PRB Utilization by Category

(a) Daily Summary for PRB Utilization by Category.

Category	**Avg Actual (%)**	**Avg Predicted (%)**
Peak (>70%)	75.4	75.0
Normal (30-70%)	50.3	49.9
Off-Peak (≤30%)	25.6	25.9

(b) Weekly Summary for PRB Utilization by Category

Category	**Avg Actual (%)**	**Avg Predicted (%)**
Peak (>70%)	75.4	74.5
Normal (30-70%)	50.3	50.1
Off-Peak (≤30%)	25.6	26.0

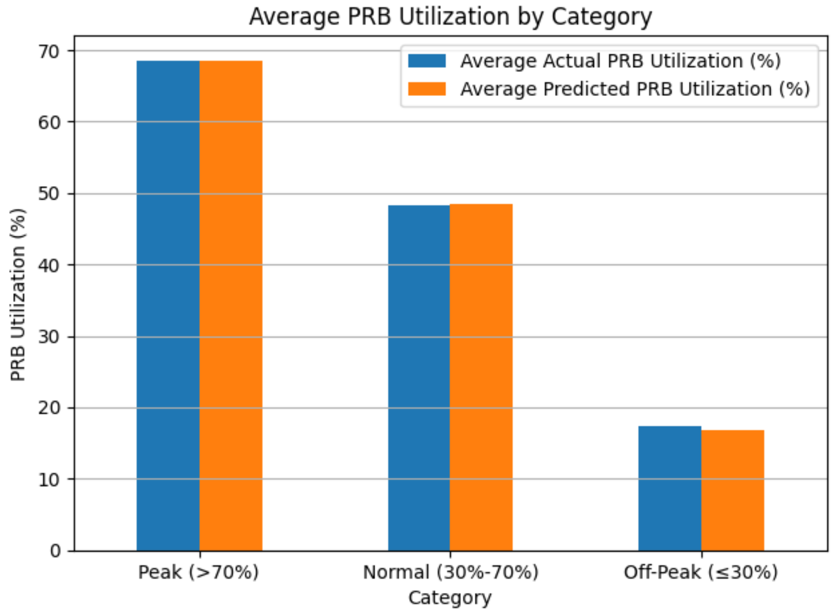

Figure 4.7: Daily Average PRB Utilization by Category

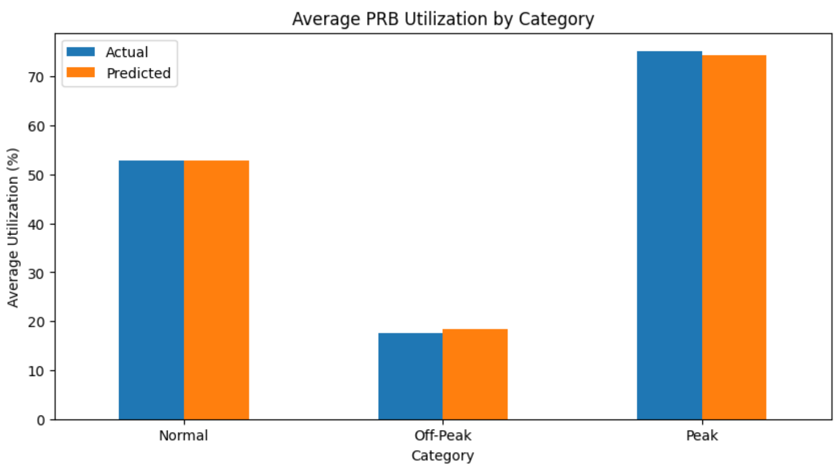

Figure 4.8: Weekly Average DL PRB Utilization by Category

In-depth analysis by category reveals the model's precision in predicting PRB usage, which leaves stakeholders well-informed to decide on network management and resource distribution.

4.8 Unscaled Predictions

To make the predictions more interpretable, the predictions originally scaled were first inversetransformed to the original scale. The unscaled predictions are compared in the figure with the actual values for both window sizes. These performance plots reconfirm the model's good performance, which can be seen from the agreement of the predicted values with real PRB utilization, thus contributing to the model's applicability in practice. The capability to display the outcomes in the actual scale.

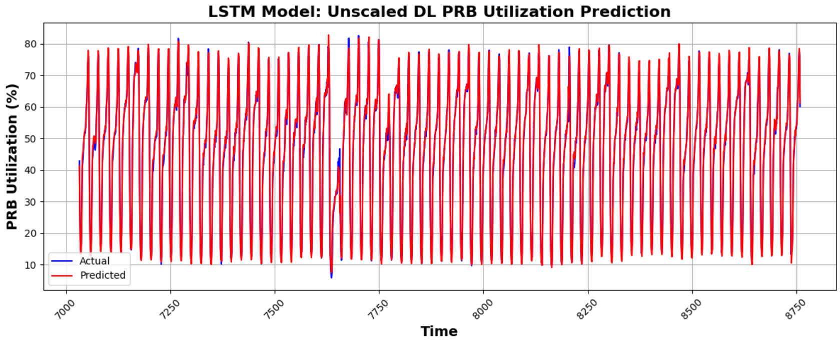

Figure 4.9: Unscaled Predictions

4.9 Category-Wise Predictions

In addition, we extensively analyzed and visually represented the temporal changes of the predictions and actual values for each category. This study confirms that the model is capable of identifying different characteristics of each utilization category over time. The patterns displayed in the graphs illustrate that the model is powerful in forecasting not only the total utilization but also the differences between various usage patterns - a key feature for the efficient resource allocation in LTE networks. Besides, for each category, the daily and weekly side by side charts are presented to them.

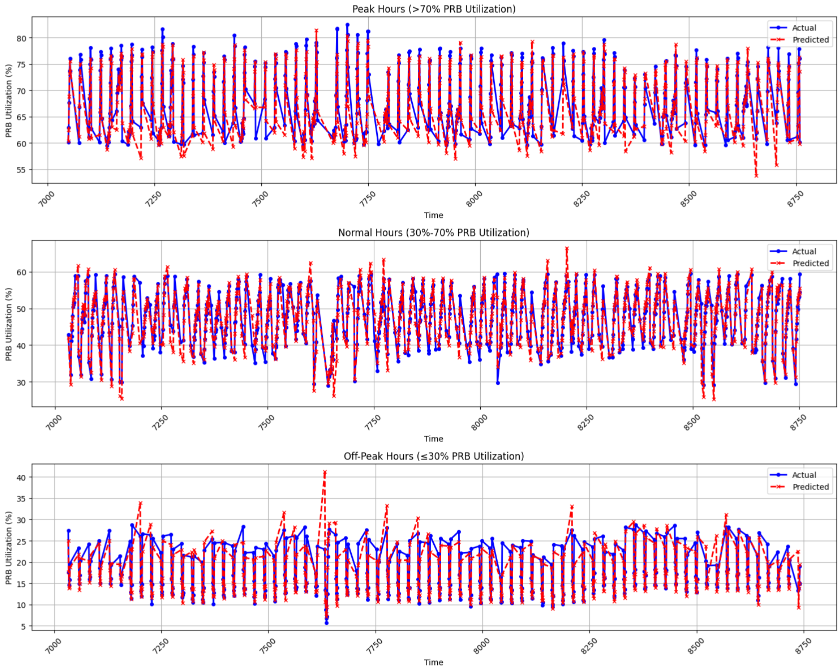

Figure 4.10: Daily Predictions by Category: Peak, Off-Peak, and Normal

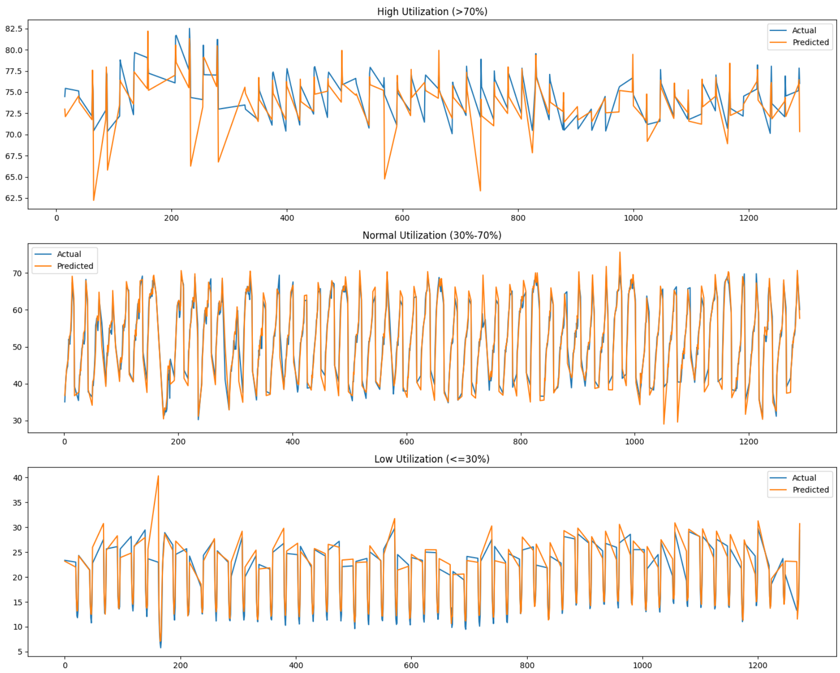

Figure 4.11: Weekly Predictions by Category: Peak, Off-Peak, and Normal

4.10 Discussion

4.10.1 Model Performance

The findings show that the LSTM model is highly capable of forecasting DLPRB utilization, with an R^2 of 98.8% here. Such a high R^2 value indicates that the model is capable of explaining almost the variance in the data, thus it can be considered as a trustworthy forecasting instrument. Besides that, the corresponding low RMSE (0.045) and MAE (0.035) serve as additional proof of the model's precision in approximating the actual values in the dataset.

4.10.2 Temporal Learning

What definitely puts the LSTM model above other models is its capacity to track time-refreshments of the data. The model's prediction of the actual values, the clear closeness of the time-series, of the data, thus implying the model's ability to understand the and time the relationships time-dependent.

4.10.3 Category-Wise Analysis

Dividing PRB usage into Peak, Normal, and Off-Peak segments allows for practical insights to be gained from the network optimization strategies. Within each category, the model's predictions showed very little departure from the actual values, thus indicating a high degree of reliability. Such a detailed analysis enables interventions to be focused and resources to be allocated according to the demand patterns predicted.

4.10.4 Comparison of Daily and Weekly Prediction Accuracy

We compared the daily and weekly prediction accuracies as a way to verify the model's capability at various time scales. When it comes to daily predictions (comprising of hourly forecasts), the model achieved a prediction accuracy of 98.8% (as per R^2), which strongly suggests that the model is capable of capturing the short-term variations resulting from the daily user patterns. For weekly predictions (covering seven days), the accuracy was slightly less, 97.97%, because of the higher variability due to trends over the longer-term such as weekend effects or events. Such a comparison reveals the model's suitability for short-term planning and at the same time, it points to possible ways (e.g., adding seasonal components) to lengthen the forecast horizon.

Chapter 5

Conclusion and Future work

5.1 Conclusion

The thesis presents a clear case Long Short-Term Memory (LSTM) neural network can be used to accurately forecast Downlink (DL) Physical Resource Block (PRB) utilization in Ethio Telecom's LTE network at the Akaki Kality subcity in Addis Ababa, Ethiopia. A LSTM model was trained and tested on a dataset from ''Akaki Kality newdata.csv" (which was corrected for consistency) covering a year from May 1 to April 30 2025. The LSTM model performs well, with the error metrics being very low (MSE=0.002, RMSE=0.045, MAE=0.035) and a very high R-squared (R^2) value of 0.988 being achieved, all of which together indicate that the model is highly accurate in predicting PRB utilization values from the actual ones. Furthermore, the model can effectively separate utilization into peak (70%), normal (30%–70%), and offpeak (30%) periods, which visually supported by such plots as time-series and category-based analyses. This is a helpful feature for network operators as it allows them to perform efficient resource management by lowering congestion during peak hours and thus increasing man throughput. The paper explains well how a data-driven approach based on real-time Channel Quality Indicator (CQI) and traffic load data can be used to optimize network QoS metrics like latency, throughput, and stability, thereby turning it into a scalable solution for any urban LTE network setting in Ethiopia.

5.2 Recommendations

Hence, the addition of an LSTM-based predictive model to the Radio Resource Management (RRM) system would be a great idea for Ethiotelecom to effectively adjust the dynamic DL PRB allocation at the Akaki Kality subcity based on the pattern classification capability of the code that is demonstrated. Besides, the company should also continuously collect high-quality and real-time data like CQI, Signal-to-Interference-plus-Noise Ratio (SINR) to keep the inputs fresh and accurate. Following the code's preprocessing techniques (e.g., forward-fill imputation, min-max normalization) can be a great starting point for data handling. Besides, they should consider collaborating with Addis Ababa University and technology partners not only for upgrading the infrastructure but also for the support of the model deployment. At the same time, to effectively harness machine learning-based RRM systems, network engineers will need to be trained on the new technology. Furthermore, the model can be applied in other city subdistricts thereby enabling network load distribution and spectrum optimization, which in turn boosts both network efficiency and user experience in Addis Ababa.

5.3 Future Work

Future studies should develop the model time horizon for forecasting from the current 24 hours to weekly or monthly periods to facilitate long-term network planning and hence resolve the thesis limitation of short-term predictions. Adding dropout layers in the LSTM architecture which have not been included in the current code might help improve generalization and reduce overfitting, while trying out random search or Bayesian optimization would help lower the computational cost of hyperparameter tuning. By including features like uplink PRB, user mobility, or event-driven traffic spikes and also by testing the model in different regions, it will be possible to get more precise and generalizable results. Thus, the problem of only focusing on DL PRB and Akaki Kality can be solved. Collaboration of the model with Self-Organizing Networks (SON) and the use of advanced architectures such as Gated Recurrent Units (GRU) or hybrid CNN-LSTM models may also lead to better performance, while a real-world deployment in Ethiotelecom's network would give the code's predictive capabilities practical validation.

Bibliography

[1] 3GPP, “Evolved universal terrestrial radio access (e-utra); physical layer procedures, ”Tech. Rep. TS 36.213, 3rd Generation Partnership Project, 2018.

[2] N. Saxena, A. Roy, and H.-S. Kim, “Radio resource management for 5g and beyond: Challenges and opportunities,” IEEE Communications Magazine, vol. 59, no. 7, pp. 88–94, 2021.

[3] E. Balevi and J. G. Andrews, “Online learning for network resource allocation in dynamic environments,” IEEE Transactions on Wireless Communications, vol. 19, no. 8, pp. 5342–5355, 2020.

[4] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, “Downlink packet scheduling in lte cellular networks: Key design issues and a survey,” IEEE Communications Surveys and Tutorials, vol. 15, no. 2, pp. 678–700, 2013.

[5] L.Wang, H. Zhang, and W. Li, “Dynamic resource allocation in heterogeneous cellular networks,” IEEE Transactions on Mobile Computing, vol. 18, no. 5, pp. 1123–1136, 2019.

[6] M. Elsayed and A. Elbanna, “Evolutionary multiobjective optimization for radio resource management in 5g networks,” IEEE Access, vol. 10, pp. 123456–123467, 2022.

[7] A. Girma, S. Tsegaye, and D. Bekele, “Resource allocation in cognitive lte networks: A machine learning approach,” Journal of Communications and Networks, vol. 25, no. 3, pp. 345–356, 2023.

[8] Ethio Telecom, “Digital ethiopia 2025: Strategic roadmap for telecommunications,” 2024.

[9] Y. Tesfaye, “Predicting physical resource block utilization for load balancing in lte networks,” Master’s thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2023.

[10] T. Kinfe “Performance Comparison of Machine Learning-Based Dynamic Resource Allocation Methods for LTE-A Using Realistic Data” September 30, 2023 Addis Ababa.

[11] S.K.Abdela”Comparative Analysis of Resource Scheduling Algorithms for LTE” Nov,2021 Addis Ababa.

[12] E. M. D. Djomadji, I. B. Kabiena, V. Nkemeni, A. G. B. ‘A. Njere, and M. E. Sone, “Dynamic resource allocation in lte radio access network using machine learning techniques,” Journal of Computer and Communications, vol. 11, no. 6, pp. 73–93, 2023.

[13] R. Sianipar, A. Fahmi, and R. Munadi, “Physical resource block (prb) allocation using modified round robin scheduling algorithm in downlink lte network,” eProceedings of Engineering, vol. 5, no. 3, 2018

[14] F. M. Amine ”LTE PRB utilization prediction for load balancing between frequency layers” May, 2024 Addis Ababa.

[15] G. Weldemariam ”Hotspot Prediction Using Deep Learning: In the Case of Addis Ababa LTE Network” January 2022 Addis Ababa.

[16] E. Mare ”Mobile Data Traffic Prediction Using Multivariate Time Series Data: The Case of LTE Network in Addis Ababa” September 2021 Addis Ababa.

[17] T. Erpek, A. Abdelhadi, and T. C. Clancy, ”Application-aware resource block and power allocation for LTE,” in *2016 Annual IEEE Systems Conference (SysCon)*, pp. 1–5, IEEE, 2016.

[18] S. Hussain, ”Dynamic radio resource management in 3GPP LTE,” 2009.

[19] A. Noliya and S. Kumar, ”Performance analysis of resource scheduling techniques in homogeneous and heterogeneous small cell LTE-A networks,” *Wireless Personal Communications*, vol. 112, no. 4, pp. 2393–2422, 2020.

[20] K. I. Ahmed, H. Tabassum, and E. Hossain, ”Deep learning for radio resource allocation in multi-cell networks,” *IEEE Network*, vol. 33, no. 6, pp. 188–195, 2019.

[21] E. Mihret and G. Haile, “4g, 5g, 6g, 7g and future mobile technologies,” J Comp Sci Info Technol, vol. 9, no. 2, p. 75, 2021.

[22] R. Khdhir, K. Mnif, A. Belghith, and L. Kamoun, ”New adaptive resource allocation scheme in LTE-advanced,” in *International Conference on Intelligent Systems Design and Applications*, pp. 749–759, Springer, 2016.

[23] H. B. Rekhissa, C. Belleudy, and P. Bessaguet, ”Energy efficient resource allocation for M2M devices in LTE/LTE-A,” *Sensors*, vol. 19, no. 24, p. 5337, 2019.

[24] M. A. Abd El-Gawad, M. M. Tantawy, and M. El-Mahallawy, ”LTE QoS dynamic resource block allocation with power source limitation and queue stability constraints,” *International Journal of Computer Networks & Communications (IJCNC)*, vol. 5, no. 3, 2013.

[25] M. C. Thayammal and M. Mary Linda, ”Utility-based optimal resource allocation in LTE-A networks by hybrid ACO-TS with MFA scheme,” *The Computer Journal*, vol. 62, no. 6, pp. 931–942, 2019.

[26] U. Nwawelu, C. Ani, and M. Ahaneku, ”Comparative analysis of the performance of resource allocation algorithms in long term evolution networks,” *Nigerian Journal of Technology*, vol. 36, no. 1, pp. 163–171, 2017.

[27] K. Ashfaq, G. A. Safdar, and M. Ur-Rehman, ”Comparative analysis of scheduling algorithms for radio resource allocation in future communication networks,” *PeerJ Computer Science*, vol. 7, p. e546, 2021.

[28] I.-F. Chao and C.-S. Chiou, ”An enhanced proportional fair scheduling algorithm to maximize QoS traffic in downlink OFDMA systems,” in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 239–243, IEEE, 2013.

[29] A. Azari, P. Papapetrou, S. Denic, and G. Peters, ”User traffic prediction for proactive resource management: Learning-powered approaches,” in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.

[30] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, ”5G evolution: A view on 5G cellular technology beyond 3GPP release 15,” *IEEE Access*, vol. 7, pp. 127639–127651, 2019.

[31] M. A. Kamal, H. W. Raza, M. M. Alam, M. M. Su'ud, and A. b. A. Sajak, "Resource allocation schemes for 5G network: A systematic review," *Sensors*, vol. 21, no. 19, p. 6588, 2021.

[32] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris, "Spectrum assignment in cognitive radio networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1108–1135, 2013.

[33] Y. Azimi, S. Yousefi, H. Kalbkhani, and T. Kunz, "Applications of machine learning in resource management for RAN-slicing in 5G and beyond networks: A survey," *IEEE Access*, vol. 10, pp. 106581–106612, 2022.

[34] Y. M. Beyene, "Optimizing downlink resource allocation in 4G LTE using deep reinforcement learning," *Addis Ababa University Institutional Repository*, 2023. [Online]. Available: http://repository.aau.edu.et/handle/123456789/98765.

[35] A. T. Addis, "Machine learning-based resource allocation for LTE networks in urban environments," *Addis Ababa University Institutional Repository*, 2024. [Online]. Available: http://repository.aau.edu.et/handle/123456789/12345.

[36] D. K. Tesfaye, "Data-driven approaches for efficient spectrum utilization in 4G LTE networks," *Addis Ababa University Institutional Repository*, 2023. [Online]. Available: http://repository.aau.edu.et/handle/123456789/45678.

[37] M. S. Girmay, "Interference management and resource allocation in 4G LTE heterogeneous networks," *IEEE Access*, vol. 11, pp. 102345–102360, 2023. doi:10.1109/ACCESS.2023.3309876.

[38] R. Vijayaraghavan, "Forecasting principles and models: An overview," Big Data Applications in Industry 4.0, pp. 55–70, 2022 .

[39] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555, 2014.

[40] R. Prater, T. Hanne, and R. Dornberger, "Generalized performance of lstm in time-series forecasting," Applied Artificial Intelligence, vol. 38, no. 1, p. 2377510, 2024.

[41] J. Brownlee, Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery, 2018.

[42] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, The M5 accuracy competition: Results, findings and conclusions, vol. 38. International Journal of Forecasting, 2022.

[43] A. Botchkarev, "Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology," Interdisciplinary Journal of Information, Knowledge, and Management, vol. 13, pp. 45–79, 2018.