Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Communication Engineering Graduate program

Comparative Study of Machine Learning Techniques for Path Loss Prediction

By

Ademe Wondimneh

Advisor

Dr. Ing.Dereje Hailemariam

A Thesis Submitted to the School of Graduate Studies of Addis Ababa University in partial Fulfillment of the requirements for the degree of Master of Science in communication Engineering.

November, 2023

Addis Ababa, Ethiopia

Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Communication Engineering Graduate program

Comparative Study of Machine Learning Techniques for Path Loss Prediction

By

Ademe Wondimneh

Approved by Board of Examiner

________________________________ _____________

Chairman, School Graduate committee Signature

Dr. -Ing. Dereje Hailemariam _____________

Advisor Signature

_____________________________ _____________

Internal examiner Signature

_______________________________ ______________

External examiner Signature

Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for a degree in this or any other university, and all sources of materials used for the thesis have been fully acknowledged.

Students Name Signature

Ademe Wondimneh _________________

Place: Addis Ababa, Ethiopia

Data of submission: ______________________

As a university advisor, I approved the submission of this thesis for examination.

Name of Advisor: signature

Dr.Ing. Dereje Hailemariam __________________

Abstract

Path loss is the term used to describe the difference in signal strength between transmitted and received. Predicting this loss is a crucial task in wireless and mobile communication to gather data for resource allocation and network planning. Deterministic and empirical models are the two fundamental propagation models that are used to calculate path loss. There is a trade-off between accuracy and computing complexity between these models. Machine learning models reflect a classic conflict between accuracy and complexity and have significant potential in path loss prediction because they can learn complicated non-linear correlations between input properties and target values. This study investigates the application of machine learning techniques for path loss prediction in Addis Ababa LTE networks. Artificial neural networks (ANNs), random forest regression (RFR), and multiple linear regressions (MLR) are employed as machine learning models and compared with the widely used COST 231 empirical model. Data for training and testing is obtained through measurements from Addis Ababa LTE networks. The performance of the proposed models is evaluated using statistical metrics such as root mean squared error (RMSE), mean absolute error (MAE), and R-squared (R^2). The results demonstrate that the RFR model outperforms the other models in terms of prediction accuracy, achieving an MAE of 3.48, an RMSE of 5.35, and an R^2 of 0.77. The ANN model also exhibits satisfactory performance with an MAE of 4.19, an RMSE of 5.78, and an R^2 of 0.71. The Cost 231 model, on the other hand, exhibits lower prediction accuracy. In terms of computational complexity, ANNs are found to be the most computationally intensive, while MLR is the simplest model among the evaluated machine learning models. RFR falls between ANNs and MLR in terms of computational complexity.

Keywords: Propagation models, Path loss prediction, Machine learning, ANN, RFR, MLR.

Acknowledgment

I want to thank God first and foremost for guiding me through all of the challenges. He is the one who supported me throughout my entire life and allowed me to complete my thesis. I will continue to put my trust in you because I believe in you.

My sincere gratitude goes out to my advisor, Dr. -Ing. Dereje Hailemariam, who helped make my thesis feasible. This research would not have been possible without his direction, helpful criticism, unrestricted support, follow-up, encouragement, and dedication.

Additionally, I would like to thank my friend Mr. Abebaw Molla (Ethiotelecom staff) for his assistance with data collection, for training me on the Actix analyzer, and for everything he showed and told me about ethiotelecom-related issues. I also thank Mrs. Bethelhem Seifu, a PhD candidate, for her support during my thesis.

I also want to express my gratitude to my friends and family for their unwavering understanding and support throughout my research. My life is made easier by all of you.

Table of Contents

List of Figures

List of tables

List of abbreviations

AI	Artificial intelligent
ANN	Artificial neural network
CNN	Convolutional neural network
CSI	Channel state information
dB	Decibel
DL	Downlink
FBNN	Feedback neural network
FFNN	Feedforward neural network
LOS	Line of sight
LTE	Long term evolution
MAE	Mean absolute error
ML	Machine learning
MLPNN	Multilayer perceptron neural network
MLR	Multiple linear regression
MSE	Mean square error
NN	Neural network
PL	Path loss
RBFNN	Radial basis function neural network
RFR	Random forest regression
RMSE	Root mean square error
RNN	Recurrent neural network
UE	User equipment
UMTS	Universal mobile telecommunication systems

1. Introduction

1.1. Background

Path loss prediction is a fundamental aspect of wireless communication system design and optimization. It involves estimating the attenuation of signal strength as it propagates through the wireless medium, taking into account various factors such as distance, frequency, and environmental conditions [1,2]. Accurate path loss prediction is crucial for ensuring reliable signal coverage, minimizing interference, and optimizing network performance [3]. In this regard, conventional radio propagation models are categorized into two types: empirical and deterministic models [4].

Empirical models are derived from extensive measurement campaigns and statistical analysis of real-world data. These models capture the empirical relationship between path loss and the influencing factors based on measured data points [3]. One commonly used empirical model is the Cost 231 model, which provides a simplified formula for estimating path loss based on distance, frequency, and other parameters. Empirical models are straightforward to implement and can provide reasonably accurate predictions in specific scenarios where the underlying data is representative. However, they may lack the ability to capture the complexities and variations present in different environments and network conditions [3].

Deterministic models, on the other hand, are based on physical principles and mathematical equations that describe the propagation of electromagnetic waves. These models aim to model the exact behavior of the wireless channel by considering factors such as terrain, building structures, and antenna characteristics. Deterministic models require detailed information about the environment and precise modeling of various parameters, making them computationally intensive and often difficult to apply in practical scenarios [5]. Examples of deterministic models include the Ray-tracing and vector parabolic equations.

With the advancements in machine learning, researchers have explored the use of data-driven approaches to improve path loss prediction accuracy [6]. Machine learning algorithms, such as Artificial Neural Networks (ANN), Random Forest Regression (RFR), and Multiple Linear Regression (MLR), have shown promise in capturing complex relationships between input features and path loss. These algorithms can learn from large datasets, incorporating various factors influencing path loss and providing more accurate predictions compared to traditional models. They can be trained offline, by either measured or synthetic (simulated) data.

Artificial Neural Networks (ANN) is particularly effective in capturing nonlinear relationships and can adapt to different scenarios. By training on historical path loss data, ANN models can learn the complex mapping between input features (such as distance, frequency, and environmental conditions) and path loss, enabling accurate predictions.

Random Forest Regression (RFR) is an ensemble learning method that combines multiple decision trees to make predictions. It can handle complex relationships and interactions between input variables, making it suitable for path loss prediction. RFR provides robust predictions by aggregating the predictions of individual decision trees.

Multiple Linear Regression (MLR) is a simpler machine learning algorithm that assumes a linear relationship between input features and path loss. MLR estimates the coefficients of the linear equation that best fits the data, allowing for straightforward interpretation of the relationship between variables and path loss prediction.

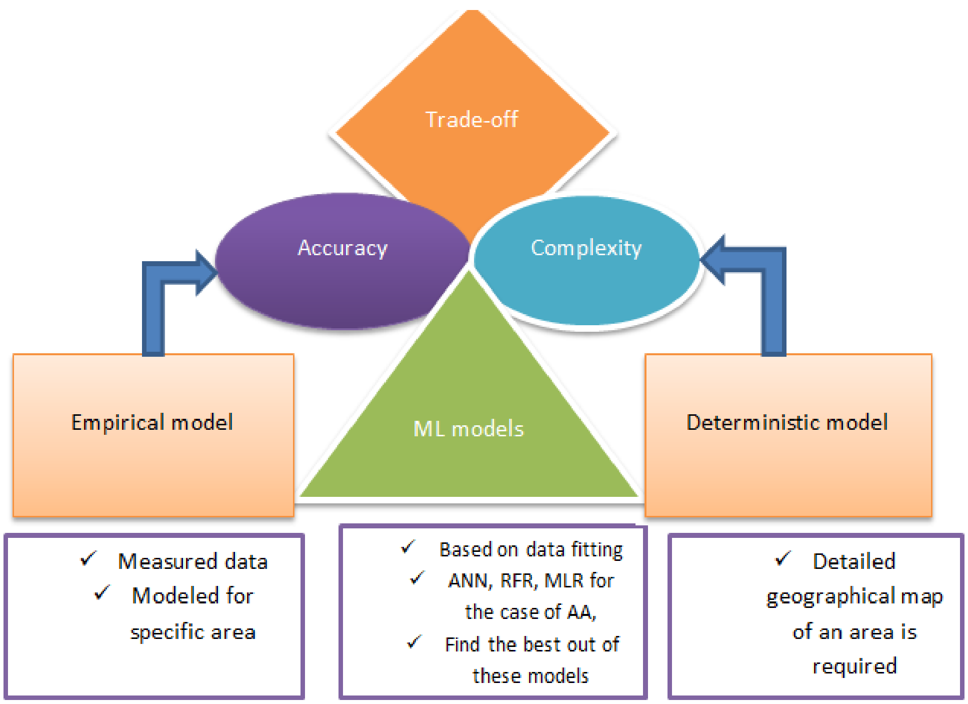
In this study, the performance of ANN, RFR, and MLR is compared with the widely used Cost 231 model, which represents an empirical approach to path loss prediction. The evaluation involves measuring the prediction accuracy of each model using metrics such as RMSE, MAE, and coefficient of regression. Additionally, the complexity of each model is considered to assess their practical implementation. By comparing the performance and complexity of these models, the study aims to provide insights into the effectiveness of machine learning algorithms in improving path loss prediction accuracy compared to traditional empirical and deterministic models.

1.2. Statements of the problem

Accurate path loss prediction is crucial for optimizing wireless communication systems, as it allows for estimating signal attenuation between a transmitter and receiver. While empirical models offer simplicity, their accuracy is often limited, particularly when applied outside of their intended environments. Conversely, deterministic models provide superior accuracy but at the expense of computational complexity, rendering them impractical for real-time applications. To address these challenges, ML models have emerged as a promising solution, offering a balance between accuracy and computational efficiency.

Within the context of Addis Ababa's unique environment, identifying the optimal ML model for path loss prediction is of paramount importance. The goal is to select a model that strikes a delicate balance between accuracy and computational efficiency. Three ML models, namely ANN, MLR), and RFR, were chosen for consideration.

Through a comprehensive comparative study, the performance of these ML models in predicting path loss will be evaluated and compared. Relevant metrics such as MSE, MAE, and determination coefficient (R^2) will be employed to assess the accuracy and predictive capability of each model. The ultimate objective is to determine which ML model is best suited for path loss prediction in the specific environment of Addis Ababa.



1.3. Objectives

1.3.1. General objective

The main aim of this thesis is to compare the performance of different machine learning models for path loss prediction and suggest the best model for the case of Addis Ababa environment

1.3.2. Specific objectives

- ✓ Understand different path loss propagation models
- ✓ Understand different flavors of machine learning algorithms

- ✓ Understand the way of data analysis using Python
- ✓ To know the way of data analysis using Actix analyzer
- ✓ Understand dataset preparation for machine learning.
- ✓ To know the different ray-tracing tools for path loss computations.
- ✓ Understand the ways of synthetic dataset preparation.
- ✓ To identify the different statistical performance measuring metrics

1.4. Literature review

Several researchers have proposed different techniques for path loss computations in various studies. Here are some of them:

A Feedforward Neural Network (FFNN) optimal model for path loss computations has been devised and shown under[7]. The data was gathered by means of drive test measuring instruments for the 1800 MHz frequency. Data analysis was done with Matlab. This paper used RMSE, MAE, MSE, R, and R2 as measuring metrics to test the accuracy of the FFNN model. The optimal accuracy using FFNN was achieved with 48-hidden neurons when the tangent function was employed as the activation function. A comparison has been made between the suggested model's performance and those of other empirical models (Cost-231, ECC-33, Hata and Egli). Plotting the measured path loss values against the distance revealed the matching projected path loss values. The results show that, over the distance range taken into consideration, the Egli model under predicts the path loss while the ECC-33 model over predicts it. Conversely, the FFNN model produced more accurate predictions than the Hata and COST-231 models, which were more cautious.

In[8], path loss computation with a machine learning model has also been studied. The goal was to demonstrate the ability of a machine learning algorithm to anticipate path loss with more accuracy and less complexity. The MLR model was the machine learning model applied in this instance. They used Matlab simulation dataset that are used by New York University in their wireless lab with some modification. Python was employed for data analysis. Using RMSE, MAE, and other metrics as measurement tools, the proposed machine learning model (MLR) was evaluated in this research for several environmental situations, including urban macro (Uma) and urban micro (UMi). It is clear from the outcome that the MLR model predicts Uma more accurately. Furthermore, the Multilinear Regression (MLR) model was evaluated with two sets of input features: four and seven. It's clear from the outcome that the MLR model with seven input features predicts more accurately.

The other researcher listed under [9] recommends path loss computation using machine learning algorithms. Utilizing drive test measurement data, the path loss of the received signal level was computed and investigated. This research employs multilayer perceptron neural network (MLPNN) and radial basis function neural network (RBFNN) as machine learning models. The RMSE was used to compare the two machine learning models. RBFNN was shown to have a smaller RMSE than MLPNN, indicating that it finds higher accuracy.

To address the lengthy simulation times associated with existing path loss prediction tools, researchers in [10] proposed RadioUnet, a CNN-based radio map estimation method. Their approach utilized model-based simulation and data-fitting techniques. They employed the U-net CNN architecture and prepared their dataset using measurements collected at 5.9 GHz from five different urban environments worldwide. WinProp software was also used for data generation. The number of buildings, the number of cars, and antenna locations were extracted and used as input features for the CNN. The proposed method, termed RadioUnet, was evaluated for both accuracy and computational complexity. Its accuracy was found to be comparable to that of ray-tracing solvers, while its computational efficiency significantly outperformed existing models. The authors noted that WinProp required 1 second under IRT-2 and 10 seconds under DPM for runtime, whereas RadioUnet achieved runtime of only 10^{-3} seconds.

Researchers in [11] investigated the use of ConvNet for path loss prediction to address the computational complexity of commercial path loss simulators based on ray-tracing models. They utilized both measurement and synthetic data for their study. The Volcano simulator was employed for synthetic data generation, while field data was collected through drive tests. The PLNet architecture was used for path loss prediction. The authors noted that different carrier frequencies typically result in varying signal attenuations and, consequently, path loss. The path loss predicted using ray-tracing simulation was compared to that predicted by PLNet. The results revealed that the predicted path loss matrix effectively captures the overall pattern of the ground truth. To augment the dataset size, data augmentation techniques such as reflection and rotation were applied. RMSE was used as the metric for evaluating prediction accuracy and was calculated by averaging over all pixels and samples. The authors trained PLNet using 728,574 samples and evaluated its performance using 96,724 samples. It took approximately 80 epochs for training to achieve a test performance of RMSE 7.48 dB. In terms of computational efficiency, the Volcano simulator required 30 seconds for processing, while PLNet completed the task in just 0.33 seconds.

Studies cited in [12] have tested the radio propagation prediction and the building occupancy calculation. They predicted path loss using FNN and estimated building occupancy using CNN. The dataset was created using measurement data, meaning that training was done using the measured data. It has undergone LTE network testing. However, it has been noted in a number of publications that determining the path loss at a particular place takes time and resources, which results in high costs. The use of solver-based generating data was the solution to this problem.

Path loss estimation with both linear and nonlinear regression methods has been applied and evaluated in [13]. In this instance, the mM wave for an urban setting has been taken into account. For dataset preparation, they used the channel state information (CSI) obtained from an open-source Matlab simulation. Their dataset preparation does not take environmental aspects into account. According to reports, employing seven features in multiple linear regression yields better results than using fewer features in nonlinear regression.

A comparison of K-nearest neighbor (KNN), RF, and support vector regression (SVR) has been done in [14]. WinProp with LTE technology was employed for data simulation modeling. They found that the suggested strategy produces better results when compared to the Cost 231 and Wolfish-Ikegami models. Furthermore, it has been noted that KNN yields more accurate results than the other two.

Deep learning has helped to increase the cellular network's capacity and coverage under [15]. The compass area is used as the study area for this research project. They tested at 2.6GHz and analyzed the ray-tracing using the Wireless Int. tool. Convolutional neural networks and concatenated neural networks have been applied to feature extraction and prediction, respectively. The suggested method's prediction accuracy was evaluated, and the results were positive.

The researchers listed below employ two deep learning strategies for path loss prediction improvement: Generalized Regression Neural Network (GRNN) and Multi-Layer Perceptron

Neural Network (MLP-NN)[16]. The dataset used in this work was created using measurements obtained from several BTSs in Nigeria. The received power (dBm) is measured using the Cellular Mobile Network Analyzer (SAGEM OT 290). From the findings, the Hata-Okumura model underestimates the path loss whereas the cost-231 model overestimates the path loss in all BTSs. The best accuracy is found using GRNN model.

Under [17], not only is the path loss at a point predicted, but also the path loss distribution in an area. The VGG-16 architecture CNN algorithm is employed. CNN requires RGB images with a size of 224 × 224 x 3, a batch size of 6, and a learning rate of 0.0001. The input parameters were 2D aerial/satellite pictures. They prepared the dataset by going through the following stages. collect a collection of satellite photos, compile matching 3D models, use ray-tracing with these models to determine PL at a dense receiver location grid, then compute the combined region's PL-histogram and image-histogram. The authors utilize Sketch up with Place Maker Plugin to create 3D models based on satellite images, and they use Wireless InSite to import each 3D model and simulate ray tracing. Ray-tracing was performed using a flat terrain model. There are two distinct frequencies in use: 900MHz and 3.5GHz. It was employed in both urban and suburban settings.

In the case of Addis Ababa, Ethiopia, [18] looked into outdoor propagation route loss models for UMTS networks using FFBPNN. Finding a different path loss prediction model for UMTS networks was the goal. Low level documentation and drive tests were employed to measure the data used for training and testing the suggested model. Data analysis was performed using Matlab. To evaluate the accuracy of the FFBPNN model, performance metrics such as RMSE, MSE, MAE, R, and R^2 were employed. It has also been done to compare the suggested approach with other empirical models. In addition to the cost 231, the FFBPNN model was also compared with the ECC-33 and SUI. It is clear from the outcome that the FFBPNN model predicts better than the others. The authors suggested that further research be done on the performance of machine learning algorithms for path loss computations in order to determine which model is the best.

1.5. Methodology

The methodology employed in this thesis encompasses literature review, data collection, data preprocessing, model development, and model evaluation. The literature review includes extensive reading of books, journals, articles, reports, and other relevant resources related to this thesis has been conducted. The data is collected from Addis Ababa LTE networks using drive test tools. In addition to the measured data, the important input features are collected from ethio telecom low-level design documents. Due to the potential presence of noise, outliers, missing values, corrupted data, and incorrect data in the collected dataset, a rigorous preprocessing step was implemented. Actix analyzer, MS-Excel and python are used for data extraction and further preprocessing. For this thesis, the distance between the transmitter and mobile terminal, altitude, longitude, base station antenna height, operating frequency, transmitting power , azimuth, mechanical tilt and electrical tilt are used as an input features and path loss is used as the output or target variable. The preprocessed data was categorized into training and testing datasets. The train_test_split technique is used to split the datasets. An ANN, RFR and MLR models are selected for path loss prediction. These models are developed using their own characteristics and the training and testing of each models are done. The developed models are evaluated to assess their performance. Statistical performance metrics, including, MAE and coefficient of determination are employed to measure the models performance relative to empirical pathloss models.

The overall methodology for this thesis is as follows:

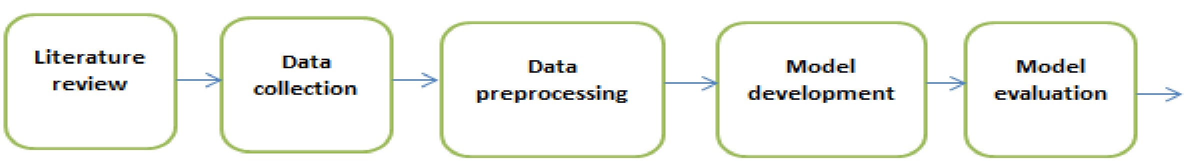


Figure 1: Methodology flow diagram

1.6. Scope and limitation of the study

The performances of ANN, RFR, MLR and cost 231 models are compared for path loss prediction. The evaluation is based on measured data from LTE networks in Addis Ababa, Ethiopia. The primary objective is to assess the prediction accuracy and computational complexity of each model. The study is limited to a specific geographical area (Addis Ababa, Ethiopia) and may not be directly applicable to other regions with different environmental characteristics. The data used for training and testing the models are collected from only LTE network and the evaluation is based on this network type (LTE) and may not accurately represent path loss behavior in other network generations (e.g., 5G). The study focuses on the performance of the models in predicting path loss under ideal conditions and does not consider the impact of factors such as clutter loss and climate on signal transmission. The computational complexity assessment is based on the specific implementation of each model and may vary depending on hardware and software configurations.

1.7. Contribution

The performance of four different path loss prediction models namely ANN, RFR, MLR, and Cost 231 are compared in this thesis. This comparison provides valuable insights into the relative strengths and weaknesses of each model. The real-world data from LTE networks in Addis Ababa, Ethiopia is used for training and testing the models. Therefore, the findings are more likely used for the environments that have similar characteristics with the area where the measurements are taken. The study found that RFR is the best model for path loss prediction in the context of LTE networks in Addis Ababa, Ethiopia. This finding is useful for network planners and operators who need to make accurate predictions of path loss.

1.8. Thesis organization

There are six sections in this thesis. The introduction is covered in the first chapter. Radio wave propagation is defined in the second part. There will be descriptions of the various path loss propagation models, radio wave propagation methods, and varieties of path loss propagation models. An overview of machine learning and the suggested models for path loss prediction is provided in the third part. The experimental setup, together with additional data gathering techniques for training and testing, will be covered in section four along with an explanation of the model development utilized for path loss prediction. The fifth portion will provide a description of the results and discussion. The conclusion and recommendations will be covered in the final section.

2. Radio wave propagation

Radio propagation, the behavior of radio waves as they travel from a transmitter to a receiver, plays a pivotal role in shaping the performance and effectiveness of wireless communication systems. It significantly influences the coverage area, capacity, and overall quality of service (QoS) that these systems can deliver.

- Coverage: Radio propagation determines the extent to which radio signals can reach from the transmitter to the receiver. Understanding propagation characteristics allows network designers to plan the placement of base stations and optimize antenna configurations to ensure adequate coverage for users within a given area.
- Capacity: Radio propagation also influences the maximum data rate that can be transmitted over a wireless link. Signal attenuation, interference, and other propagation effects can limit the capacity of wireless channels, impacting the overall network throughput.
- Quality of Service (QoS): Radio propagation characteristics affect the overall quality of service experienced by users. Signal strength, delay, and bit error rate (BER) are crucial parameters that determine the reliability and performance of wireless communication links.

Understanding radio propagation principles and modeling techniques is essential for designing and optimizing wireless communication systems. Accurate path loss prediction, interference analysis, and coverage planning are all dependent on a thorough understanding of radio propagation phenomena.

In the dynamic world of mobile communication, the movement of mobile stations, or users, introduces a range of factors that affect radio propagation. These factors include the distance traveled by the mobile station from the transmitter and the presence of various objects, both natural and man-made, along the propagation path.

Natural objects, such as atmospheric conditions, hills, water bodies, and terrain variations, can influence signal transmission. Similarly, man-made objects, such as buildings, cars, roads, trees, and vegetation, can also impact signal propagation. Additionally, time-dependent factors like the time of day and the season of the year can further influence radio propagation behavior.

2.1. Radio wave propagation mechanisms

In wireless and mobile communication, the signal transmitted from a transmitter to a receiver can follow various paths, depending on the presence of obstacles between the two points. When there are no obstructions, the signal takes the direct route, known as line-of-sight (LOS) propagation. This path offers the shortest transmission time and delivers the strongest signal to the receiver.

However, when obstacles like buildings, hills, or vegetation impede the direct path, the signal is forced to take alternative routes, resulting in non-line-of-sight (NLOS) propagation. In NLOS scenarios, the signal can follow three main mechanisms[19][20][21]

Reflection: When radio waves encounter a surface, such as a building, wall, or the Earth's surface, they can bounce off the surface and change direction. This phenomenon is called reflection. Reflected waves can reach the receiver and contribute to the overall signal strength. Reflections from multiple surfaces can create multiple paths and cause multipath propagation [22]

Diffraction: Diffraction occurs when radio waves encounter an obstacle or an edge of an object, such as a building corner or a mountain. The waves bend or spread around the obstacle, reaching areas that are in the shadow of the obstacle. Diffraction allows signals to propagate beyond the line-of-sight and improves coverage in certain scenarios [22].

Scattering: Scattering happens when radio waves interact with small objects or irregularities in the propagation medium, such as dust particles, foliage, or rough surfaces. The waves scatter in different directions, leading to signal dispersion. Scattering can cause signal loss, fading, and changes in signal polarization [22].

The figures shown below illustrates the mechanisms of radio wave propagation

Figure 2: Radio wave propagation mechanisms [18]

As it can be observed from the figure above, these propagation mechanisms interact with each other and impact the overall behavior of radio waves in a given environment. The complex combination of LOS, reflection, diffraction, and scattering contributes to signal strength, coverage, and quality variations in wireless communication systems.

Understanding these mechanisms is crucial for designing and optimizing wireless networks, accounting for signal propagation characteristics and mitigating the effects of fading, interference, and path loss[18].

2.2. Path loss propagation models

In order to estimate a propagation pathloss in a certain environment based on varying characteristics like frequency, distance, and the obstacles in the path between transmitter and receiver, pathloss propagation models are required [1]. The models are mathematical formulas built from experimental data. The models are designed to help forecast path behavior and path losses in situations when it is not possible to measure all of the actual parameters due to a variety of difficult conditions. There are multiple pathloss estimation models.

There are two types of radio propagation models: large-scale models and small-scale models. Large-scale models are found at distances of thousands of meters. In addition to being utilized in network capacity forecast and link budget calculations, it also records the loss in received signal intensity as a function of distance. Localized, small-scale models exist in time intervals of a few seconds or in space time intervals of a few meters. They are distinguished by abrupt changes in the strength of the received signal across brief travel distances or periods of time. Small-scale models capture local constructive and destructive multipath effects, which have an impact on modulation schemes, equalization tactics, and physical layer link design. The focus of this thesis is on a large-scale path loss model.

In the context of large-scale modeling, the transmitter and receiver can be situated in environments characterized by line-of-sight (LOS) or non-line-of-sight (NLOS) conditions. These environments often feature obstructions such as buildings, trees, and other objects.

In LOS scenarios, the receiver may receive a direct attenuated signal, known as the line-of-sight signal, from the transmitter. However, in NLOS scenarios, the receiver receives indirect signals, referred to as non-line-of-sight signals, due to various physical phenomena such as reflection, refraction, diffraction, and scattering. These indirect signals can arise from interactions with the surrounding environment and objects.

In addition, the direct and indirect signals can interfere with each other, leading to further variations in the received signal strength. The interference between LOS and NLOS signals can significantly impact the overall signal quality and coverage.

Regarding large-scale modeling, it can be categorized into three main types. These are deterministic, statistical and empirical models.

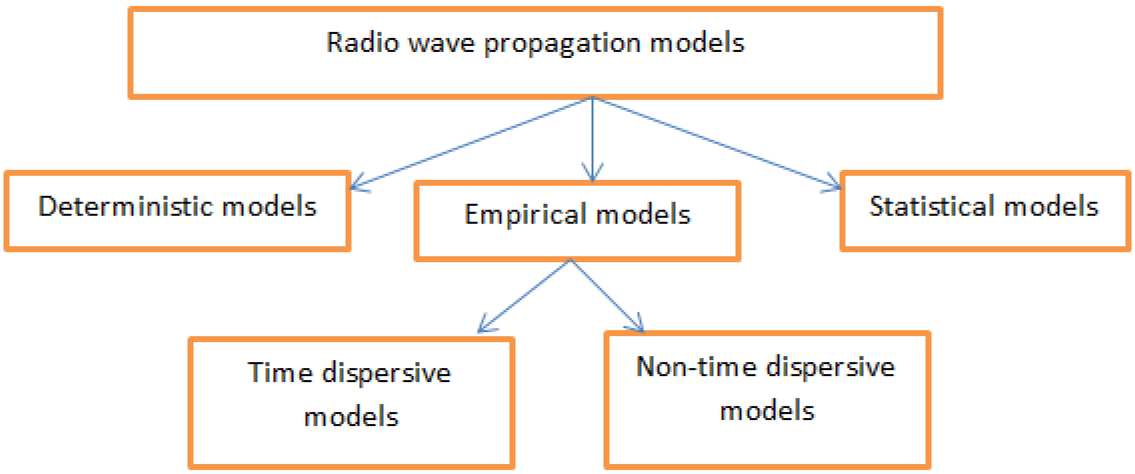


Figure 3: types of path loss propagation models

2.2.1. Deterministic models

Deterministic models in path loss prediction are based on the fundamental laws governing electromagnetic wave propagation. These models utilize the principles of physics rather than statistical outcomes derived from experiments. By incorporating the physical principles of electromagnetic wave propagation, deterministic models aim to predict the received signal power at specific locations.

Unlike empirical models that rely on statistical analysis of measured data, deterministic models are built upon the understanding of wave behavior and the interaction with the propagation environment. These models establish a simplified relationship between the characteristics of the environment and the propagation of signals[23].

The success of a deterministic model depends on the information available about the propagation environment. The model's ability to accurately map the real-world propagation environment into the model's representation is crucial in achieving reliable predictions. Deterministic models tend to require more computational resources due to the complexity of the calculations involved. However, they generally offer higher accuracy compared to empirical models.

One example of a deterministic model is a ray tracing model. Ray tracing models simulate the propagation of electromagnetic waves by tracing the paths of individual rays as they interact with objects and surfaces in the environment. These models require extensive computations but provide more accurate predictions of signal propagation [23].

2.2.2. Empirical models

Empirical models in path loss prediction are constructed based on observed and measured data using the general path loss formula. These models use regression techniques to derive mathematical expressions that describe the path loss. They are primarily employed for predicting path loss in specific areas where measurements have been conducted but may require adaptation for different locations [18]. The accuracy of empirical models depends on the similarity of the propagation environment in which the model was developed to the environment where it is being applied [18].

One common challenge in using empirical path loss models is applying them in areas with significantly different propagation environments from where the model was originally developed. For instance, the Okumura model attempts to address this issue by incorporating correction factors, but discrepancies persist unless the characteristics of the environment, such as "urban," "suburban," and "open," closely resemble those in the region where the measurement data was collected. Despite this limitation, empirical models are widely used due to their simplicity and computational efficiency[18].

Empirical models can be further classified as time dispersive and non-time dispersive models. Time dispersive models provide information about the channel's time-varying behavior, specifically the multipath delay spread. Non-time dispersive models, on the other hand, use data sets and statistical analysis to construct curves that represent the average path loss as a function of parameters such as antenna heights, height above average terrain, terrain roughness, and local clutter (e.g., foliage and buildings). Examples of non-time dispersive models include the SUI, Hata, and COST-231 Hata models [18].

2.2.3. Statistical models

Statistical models in path loss prediction utilize random variables to characterize the propagation environment by assigning probability distributions to channel parameters such as delay spread or small-scale fading factors [24]. Similar to empirical modeling, propagation experiments and statistical analysis are conducted to validate the goodness-of-fit of each parameter distribution. Correction factors can be introduced based on the fitness of the data. These models are commonly employed in the planning of green field or open area deployments and evaluations, as they rely on simple parameters. They provide simple single-slope log-distance expressions to predict the mean path loss at a given distance, d, from the transmitter.

In the following sections, the COST 231 model, which is deployed in the AA environment, will be presented. This model is used to compare the proposed machine learning models used for path loss prediction. To facilitate a comprehensive understanding, we will begin with the simplest propagation model, known as the free space model. Additionally, we will discuss the baseline models that are used as references for the COST 231 model, namely the Okumura models and Hata models.

A. Free space path loss model

It is the simplest model and formulated as follows [17]

$$FSPL = 20\log(f_c) + 20\log(d) + 20\log\left(\frac{4\pi}{c}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 2.1$$

Where c stands for speed of light in vacuum (m/s), d is the distance between the transmitter and the receiver in km and f_c is the frequency in GHz. The constant term in equation 2.1 becomes -27.55 and replacing this in the above equation becomes

$$FSPL = 20\log(f_c) + 20\log(d) - 27.55 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 2.2$$

We can see from Equation 2.2 that the path loss depends only on the frequency of the signal and the distance between the receiver and the transmitter showing how simple this model is, this is of course ignoring many details and characteristics that affect the real measurements.

B. Okumura model

The Okumura model, one of the most renowned models in the field, serves as the foundation for other models specialized in open areas, as it has limitations in dealing with areas with high-blocking buildings[25] . This model is well-suited for free areas without blockage from trees or buildings, as well as suburban areas such as villages or scattered houses with minimal blockage near the transmitter, and urban areas with medium blockage, such as towns with two or more-story buildings[17].

The Okumura model can be applied to frequencies ranging from 150 to 1920 MHz in urban areas. The transmitter's height can vary between 30 to 1000 meters, while the receiver's height is limited to less than 10 meters. To calculate the path loss using the Okumura model, the median path loss is determined using the following equation[17].

$$PL = L_F + A_{mu}(f,d) - G(h_{te}) - G(h_{re})_{G_{AREA}} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots . 2.3$$

where L_F stands for free-space path loss, median attenuation relative to free space is $A_{mu}(f,d)$, transmitter altitude gain is $G(h_{te})$, $G(h_{re})$ is the receiver altitude gain and G_{AREA} is environmental gain such as urban, rural, etc. Furthermore, the equations below are important to calculate the Okumura model path loss for that environment:

$$G(h_{te}) = 20\log\left(\frac{h_{te}}{200}\right) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 2.4.$$

and this can be applied for transmitter height between 30m and 1000m, while for receiver altitude less than 3m we can use Equation (2.5), but for higher altitudes for the receiver we use Equation (2.6),

$$G(h_{re}) = 10\log\left(\frac{h_{re}}{3}\right) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots .. 2.5$$

$$G(h_{re}) = 20\log\left(\frac{h_{re}}{3}\right) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots ..2.6$$

Correction factors like terrain related parameters can be added using a graphical form to allow for street orientation as well as transmission in suburban, open areas, and over irregular terrain. There are four types of irregular terrain: general sloping terrain, isolated mountains, rolling hills, and mixed land-sea paths. The different correction factors need to be determined by evaluating the terrain related parameters.

C. Hata model

We now look at Hata model, which is a limited version of Okumura model as it is limited for quasi-smooth terrain [25], with limited transmitter altitude between 30 - 200 meters and transmission distance of less than 20 km. The Hata model equation to estimate path loss for urban areas is shown as follows [26]

$$PL = 69.55 + 26.16\log(f_c) - 13.82h_{te} - a(h_{re}) + (44.9 - 6.55\log(h_{te}))\log(d) \ldots \ldots \ldots 2.7$$

where f indicates the frequency in MHz, h_{te} and h_{re} are the transmitter and receiver height in meters, respectively, the distance between transmitter and receiver is d, and $a(h_{re})$ is the correction factor, which is 0 for 1.5m height.
For small and medium cities, $a(h_{re})$ is as follows:

$$a(h_{re}) = (1.1\log(fc) - 0.7)h_{re} - 1.56\log(fc) + 0.8 \ldots \ldots \ldots \ldots \ldots \ldots \ldots ..2.8$$

while for large cities:

$$a(h_{re}) = \begin{cases} 3.2(\log(11.75h_{re}))^2 - 4.97, fc > 300 \\ 8.29(\log(1.54h_{re}))^2 - 1.1, fc \leq 300 \end{cases} \qquad \ldots\ldots 2.9$$

Some studies have suggested that the path loss experienced at 1845 MHz is approximately 10 dB larger than those experienced at 955 MHz [25]. This is why Cost231-Hata model [27] was introduced to help with this deviation from the Okumura model.

D. Cost 231 mode

It is an extension of Okumura_Hata model used for the frequency range from 1500MHz to 2000MHz. This model is used for the separation distance up to 20km and the transmitter and receiver antenna height considered being from 30m-200m and 1m-10m respectively[27]

$$P_L(dB) = 46.3.3 + 33.9\, logf_c - 13.82\, logh_t - a\big(h_{m)}\big) + (44.9 - 6.55\, logh_t)\log d + cf, \ldots\ldots 2.10$$

Where $a\big(h_{m)}\big) = 3.2(log11.75h_r)^2 - 4.97\ for\ urban\ and\ f > 400MHz$ ………..……2.11

$$a\big(h_{m)}\big) = (1.1\ logfc - 0.7)hr - (1.5logf - 0.8)\quad for\ suburban\ and\ rural \ldots\ldots\ldots 2.12$$

$$cf\ is\ the\ correction\ factor\ = 0\ dB\ for\ suburban\ and\ 3\ dB\ for\ urban$$

3. Machine learning models

Machine learning is an application of artificial intelligent (AI) which gives devices the ability to learn from their experience and improve their self without doing any program[28]. It can be also defined as a field which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

3.1. Types of machine learning

Based on the type of input and output data, machine learning algorithms can be classified into three types: Supervised, Unsupervised and Reinforcement learning[29]

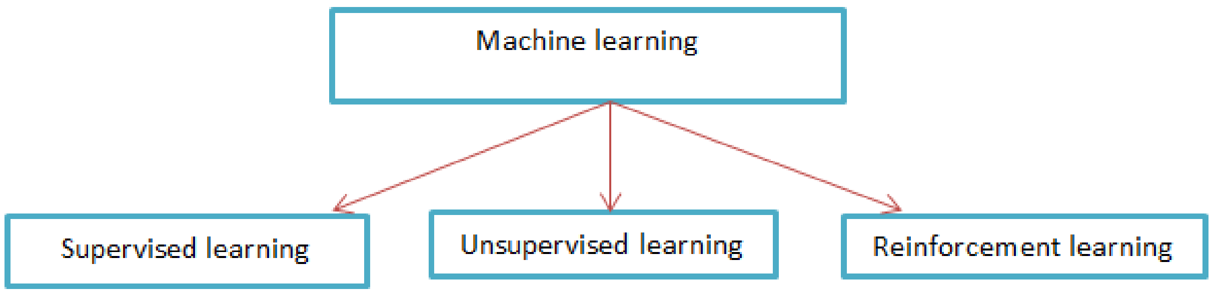


Figure 4: types of machine learning

3.1.1. Supervised learning

In this kind of machine learning, labeled data is used to train the machine. Both the target values and the input features are known. Once the computer has been trained with this data, it assists in predicting results for data that has not yet been seen. Because the dataset must be manually categorized by an individual, this is an extremely expensive operation, particularly when handling big amounts of data. Supervised learning comes in two varieties. Regression and classification are these two. While the target values in regression are continuous, in classification supervised learning the output values are discrete

Path loss prediction uses input features such as separation distance, operating frequency, longitude, antenna height, and so on. Path loss is the target value, and these are all known. Consequently, this falls under the category of learning algorithms that use regression

3.1.2. Unsupervised learning

In this type of learning, the user need not supervised the model instead it allows the model to work on its own to find patterns and information that was previously undetected. It is based on unlabeled data. Clustering and classifications types of problems are grouped under this type of learning.

3.1.3. Reinforcement learning

It is a type of machine learning where the training of machine learning models to make a sequence of decisions. It learns from mistakes and makes a serious of decisions based on the rewards and feedback they receive for their actions. The machine learns from its own experience when there is no training dataset present.

In the next, the key terms that are used related with machine learning algorithms will be described.

Features: These are the fields that are used as an input. One column of the data in the input set can be considered as a feature. These are independent variables and that could be fed to the model. In the case of path loss prediction, the distance, height of transmitter, the location of base stations, the operating frequencies, transmitted power, etc are taken as features.

Label: These are the output values that we get from the model after training. The predicted values are the labels. This are also called target or dependent values. In the case of path loss prediction using the power of machine learning, the path loss values are the target values.

Model: The relationship between features and the label is called a model. A machine learning model is a mathematical model that generates predictions by finding patterns in the given data.

3.2. Regression-based algorithms used for path loss prediction

Based on the underlying algorithm they used and the type of task they are designed to solve, there are different regression machine learning models. The choice of each model to be selected depends on many factors like the problem statement and kind of output you want, type and size of the data, the available computational time, number of features, and observations in the data.

So many scholars proposed so many machine learning models for path loss prediction. Out of these, the ANN, RF and MLR model are tested in this thesis for the case of AA and compared with the well-known cost 231 propagation models. Therefore, the detail of these machine learning models will be explained in the next section.

3.2.1. Artificial neural network

An artificial neuran networks are programs designed to solve any problems by trying to mimic the structure and functions of human nervious system. [30]These are constructed and implemented to model the human brain. It performs varios tasks such as pattern matching, classification, optimizations, predictions, etc. ANN possess a large number of processing elements called neurons. Thses neurons are woking on parallal and are connected with each others by conncetion link. Each link is associated withweights which contain information about the input signal. Each neuron has an internal state ot its own which is a function of inputs that neurons receives.

I. Elements of neuran network

The following digarma shows the elementart structure of neural networ. There different components that are used in representation of neurons. These are inputs, weights, activation fucntions, outputs, etc.

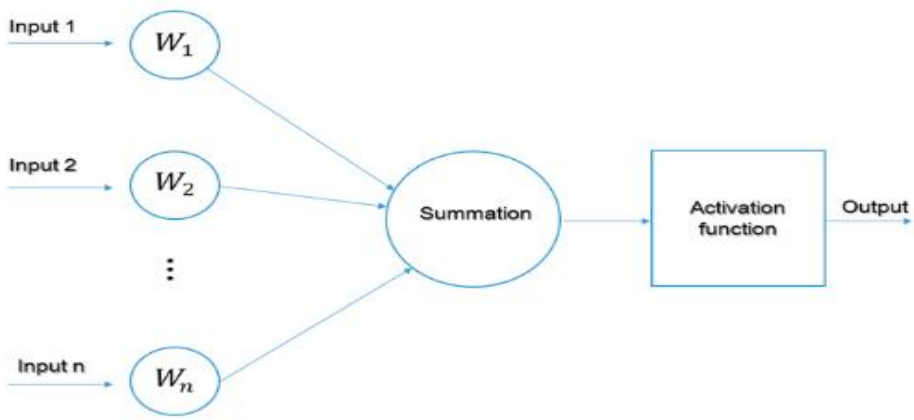


Figure 5: Elementary structure of a single neuron in a Neural Network

The summetion is artificial neurons. As it can be observed in the figure above, multiple inputs and bias are going into the artificila neurons. This artificial neuron gives an output.

- Inputs

These are collection of data that are used as an inputs for training the model. They are also called features. For example, the factors the contributes to the path loss like separation distanbce, operating frequencies, height of transmitter and reciever, antenna orenations, environment, etc can be taken as an inputs.

- Interconnections

These refers to connctions between individual neurons in the network. They are represented by weights. They helps to facilitate communication between neurons and allows to pass infromation from one neuron to the other. These weights are numerical values that are assigned to the connection between neurons. They are assigned with an initial small random values. They can be adjusted using rules like backpropagation algorithms . This algorithm works obased on the input and output values. The algorithm calculates the gradient of the error with respect to the weights. This gradient will be used to undate the weights in such a way that reduces the error.

- Learning rules

It's a process by which a NN adapts itself to a stimulus by making proper parameter adjustments, resulting in the production of desired response. Two kinds of learning

a) Parameter learning:- connection weights are updated
b) Structure Learning:- change in network structure

- Learning

The process of modifying the weights in the connections between network layers with the objective of achieving the expected output is called training a network.

This is achieved through

- Supervised learning
- Unsupervised learning
- Reinforcement learning

- Activation function

Activation function: it is the weighted sum of the summing unit, and the output is generated based on the signal from this activation value. There are different types of activation functions[31]. The performance of an ANN is also determined by using the best activation function. Each activation function has different performance for different problems.

- ✓ Identity function:

It returns the input as the output without any transformation.

$f(x) = x \; for \; all \; x$

- ✓ Sigmoid function

It maps the input to a value between 0 and 1, which can be interpreted as a probability. It is represented mathematically as follows:

$$S(x) = \frac{1}{1 + e^{-z}}$$

The graph of sigmoid function is shown below.

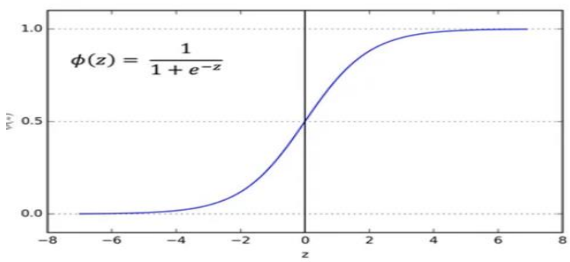


Figure 6: Sigmoid function [31]

The gradient of sigmoid function is defined everywhere, and that its output is conveniently between 0 and 1 for all x. The logistic sigmoid function is easier to work with mathematically, but the exponential functions make it computationally intensive to compute in practice and so simpler functions such as ReLU are often preferred

- ✓ ReLu

There are a number of widely used activation functions in machine learning. One of the simplest is the rectified linear unit, or ReLU function, which is a piecewise linear function that outputs zero if its input is negative, and directly outputs the input otherwise [31]

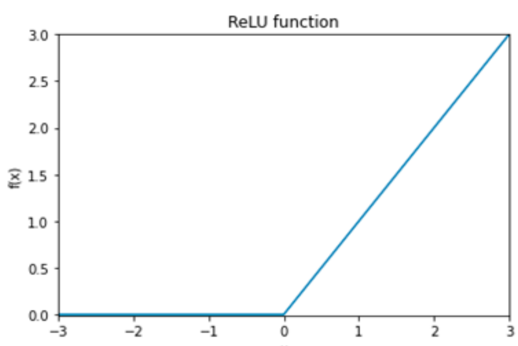


Figure 7: ReLu function[31]

The ReLu function returns the input if it is positive, and 0 otherwise

Mathematically [19]

$f(x) = \max(0, x)$

- ✓ Outputs

These are the values that are predicted by the models. They are also called target value. They can be contnious or discret, they can be single or multiple. For example in the case of path loss prediction using the power of machine learning, the path loss values are the outputs / target values.

II. Structure of NN

Based on the interconnection between neurons, there are two types of ANN. These are feedforward, feedbackward.

A. Feedforward neural network

In this type of netowk, there is only one direction of computation. It is a type of non-recurrent neural network that consists of input, output, and hidden layers. The network processes input data by passing signals in one direction through the layers. Each processing unit in a layer receives input data and performs calculations based on a weighted total of its inputs. The resulting values are then used as inputs for the subsequent layer. This process continues until the output is determined after passing through all the layers. Examples of feed-forward networks include the Perceptron (linear and non-linear) and Radial Basis Function networks. The single-layer perceptron is the most basic type of neural network, with a single layer of output nodes that receive inputs directly via a set of weights. Each node calculates the total of the products of the weights and inputs. This architecture was one of the first and simplest neural network structures to be developed.

The feedforward ANN has two types. The single layer feedforward ANN and multi layer feedfroward ANN.

a. single layer feedforward neural network

A neural network in which the input layer of source nodes projects into an output layer of neurons but not vice-versa is known as a single layer feedforward network [32]. In other words, a hidden layer is absent or not present in the network and there is no feedback from the output or hidden layer. In the single-layer network, 'single-layer' refers to the output layer of computation nodes as shown in Figure below.

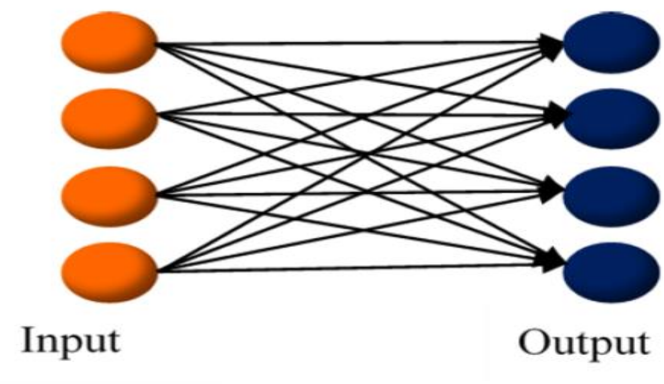


Figure 8: single layer feed-forward network [32]

b. Multi layer feedforward ANN

This type of network consists of one or more hidden layers, whose computation nodes are called hidden neurons or hidden units. The function of hidden neurons is to interact between the external input and network output in some useful manner and to extract higher-order statistics. The source nodes in the input layer of the network supply the input signal to neurons in the second layer (1st hidden layer). The output signals of the second layer are used as inputs to the third layer and so on [33].

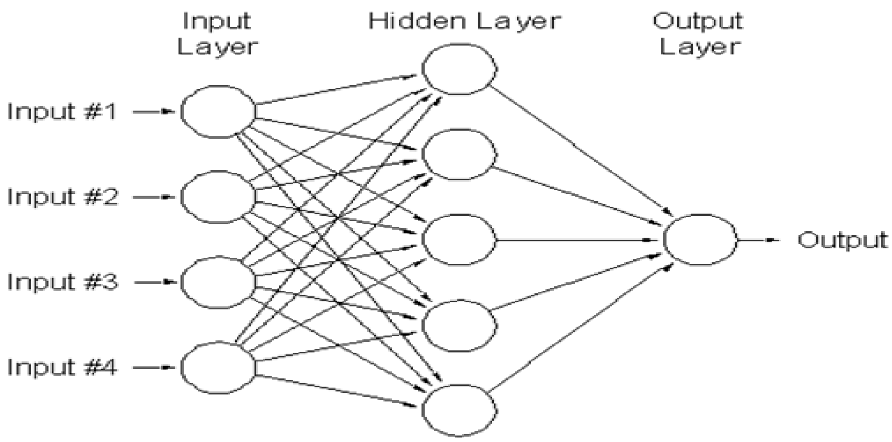


Figure 9: Multi-layer FFNN

B. Feedback Neural Networks

A feedback network, such as a recurrent neural network (RNN), is characterized by feedback paths that allow signals to travel in both directions through loops [18]. Unlike feed-forward networks, neuronal connections in RNNs can be made in any way, leading to a non-linear dynamic system that evolves continuously during training until it reaches an equilibrium state.

The feedback paths in RNNs enable the network to model sequences of data, making them useful for tasks such as speech recognition, language translation, and time-series analysis. During training, the network learns to adjust the weights and biases of its connections to optimize its performance on the task at hand.

RNNs have been shown to be effective in a variety of applications, including natural language processing, image and speech recognition, and robotics [29] However, they can be challenging to train due to the vanishing gradient problem, which can make it difficult for the network to learn long-term dependencies. The following figure shows the structure of FBNN.

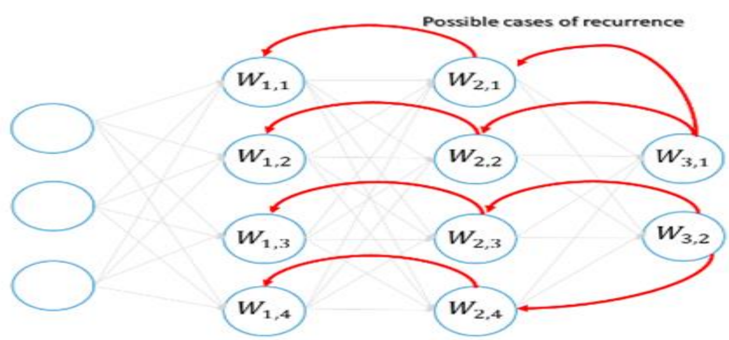


Figure 10: structure of FBNN [29]

3.2.2. Random forest regression

Random forest regression is a machine learning algorithm used for regression tasks that involve predicting a continuous output variable based on one or more input variables. It is an ensemble learning method that combines multiple decision trees to make predictions [34]. In random forest regression, each decision tree in the ensemble is trained on a random subset of the data and a random subset of the input features. This helps to reduce over fitting, which can occur when a model is too complex and fits the noise in the data rather than the underlying patterns [34].

To make a prediction using a random forest regression model, the input data is passed through each decision tree in the ensemble, and the output from each tree is averaged to produce the final prediction. The resulting prediction is typically more accurate and less prone to over fitting than the prediction from a single decision tree [35]. The trees in random forests run in parallel, meaning there is no interaction between these trees while building the trees.

A. Ensemble learning

An ensemble method refers to a strategy that amalgamates predictions from multiple machine learning algorithms to enhance predictive accuracy beyond what each individual model can achieve[34]. The resulting composite model, formed by combining various models, is termed an ensemble model.

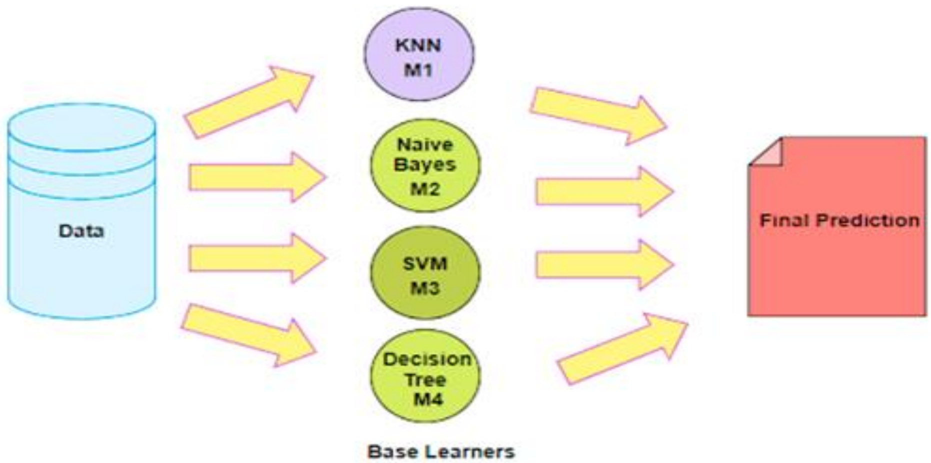


Figure 11: Ensemble learning method[34]

There are two types of ensemble learning. These are boosting and bagging.

i. Boosting

Boosting is a set of algorithms that leverage weighted averages to transform weak learners into more robust ones. It operates on the principle of collaboration, where each model's output influences the features emphasized by the subsequent model. The term "boosting" reflects the idea of one model learning from another, leading to an iterative improvement in the learning process[36]

ii. Bagging

It is alternatively known as bootstrap, which involves random sampling with replacement. This technique enables a more in-depth analysis of bias and variance within a dataset. Bootstrap entails the random extraction of a small subset from the dataset[36]. Bagging, a broader method, is employed to mitigate variance, particularly in algorithms like decision trees with high variance. In bagging, each model operates independently, and their outputs are aggregated without favoring any specific model.

B. Working principles of Random forest regression

Random forest operates by constructing a multitude of decision trees at training time and outputting the class that's the mode of the classes (classification) or mean prediction (regression) of the individual trees.

A random forest is a meta-estimator (i.e. it combines the result of multiple predictions), which aggregates many decision trees with some helpful modifications:

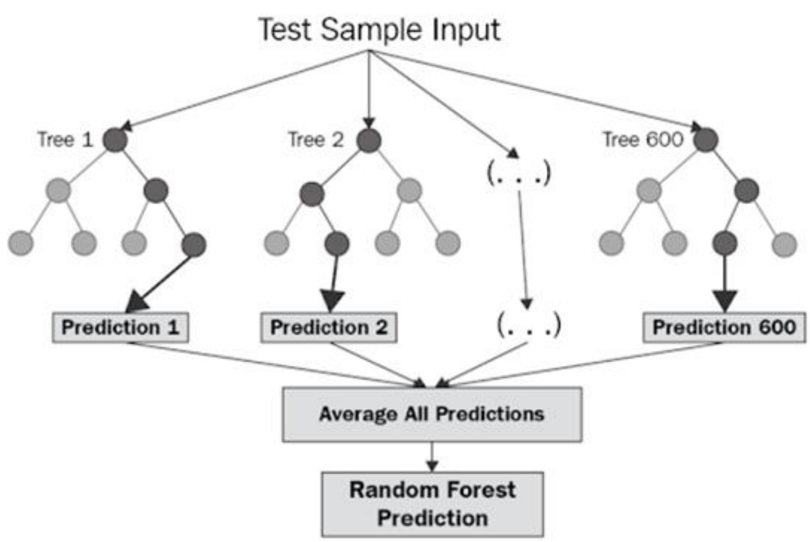


Figure 12: Random Forest samples[37]

C. Advantages and disadvantages of random forest regression

i. Advantages

- ✓ Random forest is one of the most accurate learning algorithms available. For many data sets, it produces a highly accurate classifier.
- ✓ It runs efficiently on large databases.
- ✓ It can handle thousands of input variables without variable deletion.
- ✓ It gives estimates of what variables are important in the classification.
- ✓ It generates an internal unbiased estimate of the generalization error as the forest building progresses.
- ✓ It has an effective method for estimating missing data and maintains accuracy when a large proportion of the data are missing.

ii. Disadvantages

- ✓ Random forests have been observed to over fit for some data sets with noisy classification/regression tasks.
- ✓ For data including categorical variables with different numbers of levels, random forests are biased in favor of those attributes with more levels. Therefore, the variable importance scores from random forest are not reliable for this type of data.

3.2.3. Linear regression

A strong statistical method for simulating the relationship between a dependent variable and one or more independent variables is regression analysis. Simple Linear Regression (SLR) and Multiple Linear Regression (MLR) are two essential types of regression that provide insights into data patterns and connections. This section discusses the concepts, presumptions, and applications of SLR and MLR.

1. Simple linear regression(SLR)

With simple linear regression, a linear equation is used to predict a dependent variable y based on a single independent variable x [34].

The simple linear equation takes the form of :

$$y = \beta_0 + \beta_1 x + \epsilon$$

Where β0 and β1 are the y-intercept and the slop respectively and ε as the error term.

1.1.Working principle of SLR

The goal of simple linear regression is to find the values of β_0 and β_1 that best fit the data points. This is done by minimizing the sum of squared errors (SSE), which represents the total difference between the predicted values of y (based on the model) and the actual values of y

There are two main methods for finding the line of best fit:

Least squares method: This is the most common method and involves solving a set of equations to find the values of m and b that minimize the SSE.

Gradient descent: This is an iterative algorithm that starts with an initial guess for m and b and then adjusts them in small steps until the SSE is minimized.

1.2.Multiple Linear Regression (MLR)

Multiple linear regression (MLR) builds upon the principles of simple linear regression but extends them to analyze the relationship between a dependent variable and multiple independent variables. Here's a breakdown of its working principles:

Multiple Linear Regression extends SLR to predict y based on multiple independent variables[34]

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

where y_i is the dependent variable/ outputs/ predictions/ target variables β_0 is the intercept, and $\beta_1, \beta_2 \; and \; \beta_p$ are the coefficients for the independent variables $x_{i1}, x_{i2}, \ldots x_{in}$

1.2.1. Working principles

Similar to simple linear regression, MLR aims to find the values of β1, β2, ...βp, and β_0 that best fit the data. This involves minimizing the sum of squared errors (SSE), the total difference between the actual and predicted values of y

Least squares is the primary method for finding optimal coefficients. It involves solving a system of equations based on the data and the model equation. Alternatively, iterative algorithms like gradient descent can be used, where initial guesses for the coefficients are gradually updated to minimize the SSE.

Once the optimal coefficients are determined and the model's performance is validated, it can be used to make predictions for new sets of independent variable values. Simply plug the values into the model equation and solve for the predicted value of y

1.3.Common Assumptions for SLR and MLR

There are common assumptions for both SLR and MLR models. These are:

- Linearity:

Relationship: There must be a linear relationship between the independent variable or variables and the dependent variable. This indicates that the relationship between the variables can be represented by a straight line.

Verification: The link can be seen and its linearity evaluated using scatterplots. It may be necessary to use alternative models or transformations (such as logarithmic) for non-linear patterns.

- Constant variance:

homoscedasticity All values of the independent variable(s) should have a constant variance, or the difference between the actual and predicted values.

Verification: The distribution of the errors can be seen using residual plots, which can also be used to spot any heteroscedasticity (non-constant variance). To address it, methods such as weighted least squares can be used.

- ❖ Error independence:

Errors without correlation: There should be no correlation between the errors for any given observation. This implies that one observation's result shouldn't affect the result of another.

Verification: The independence of errors can be evaluated using tests such as the Durbin-Watson statistic. Autoregressive integrated moving average (ARIMA) models are one type of approach that may be necessary when dealing with autocorrelation (correlated errors).

- ❖ Error normality:

Standard distribution: There ought to be a normal distribution of errors. Although less important than others, this assumption may have an impact on the reliability of particular statistical tests.

Verification: The normality of mistakes can be evaluated using Q-Q plots and normality tests. If the assumption is not met, robust regression techniques or transformations can be applied.

- ❖ Not multicollinear (limited to MLR):

Independent variables: The degree of correlation between the independent variables shouldn't be very high. They must so offer distinctive data regarding the dependent variable.

Confirmation To determine multicollinearity, use the Variance Inflation Factor (VIF). One way to deal with it is to use strategies like dimension reduction or removing strongly linked variables.

1.4.Applications of MLR

MLR is widely used in many fields, including finance, economics, social sciences, and engineering. It can be used for both explanatory and predictive modeling, and it is useful for identifying the most important factors that contribute to the dependent variable. In this thesis, we are going to use this model for path loss prediction.

4. Experimental setup and model development

4.1. Experimental setup

This section presents the specific decisions and setups made in relation to the model construction, hyperparameter tuning, preprocessing methods, evaluation measures, and dataset used in path loss prediction. These decisions delineate the methodical methodology employed to guarantee the study's validity and dependability, facilitating equitable comparisons of various machine learning approaches.

4.1.1. Data collection

The collection of data is very important for accuracy of machine learning models. The data used for machine learning models used for path loss predictions can be obtained through simulation or via measurement. For this case, the collection is done through measurement. An extensive measurement campaign was conducted in the LTE networks of Addis Ababa, Ethiopia to obtain samples. Each sample should include the path loss values and the corresponding input features. The measured input feature includes the longitude, latitude, separation distance between transmitter and receiver and the received power. Nemo Handy, a popular measurement tool, was used for data collection process. Other important features like mechanical tilt (M-tilt), electrical tilt (E-tilt), azimuth, transmitter power and height of each base station were obtained from low-level documents provided by Ethio Telecom. The area where measurements were conducted is shown in figure below.

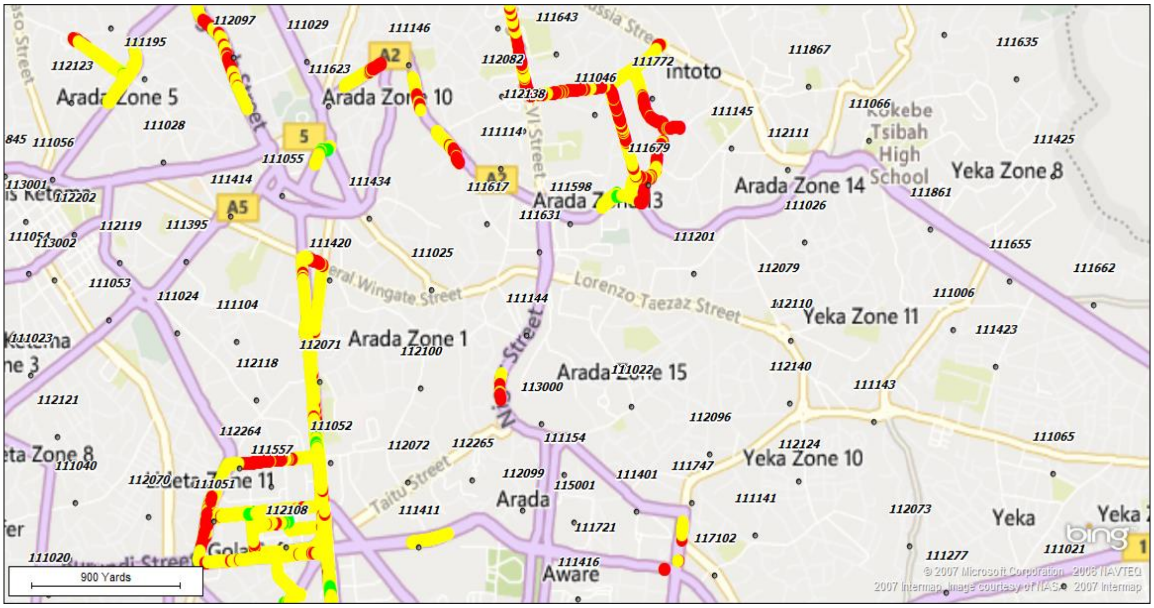


Figure 13: The area where measurements are taken

4.1.2. Data preprocessing

The data may need to be prepared before it can be used to train and test the machine learning algorithms. This may involve cleaning the data, removing outliers, and transforming the data into a format that is compatible with the machine learning algorithms.

The data obtained from Nemo handy is in the form of raw data. This raw data was extracted using Actix, a software tool commonly used for extracting information from measurement tools. The data was exported from Actix and further analyzed using Excel.

There were missing data points. Any missing data points were carefully examined using an elimination technique.

4.1.3. Feature selection

Propagation path loss is influenced by various parameters, some of which used as input features while the path loss itself is treated as the target or output value for this thesis. The input features that are utilized for this thesis are separation distance between transmitter and receiver, longitude and latitude of the transmitters, transmitter antenna height, operating frequency, transmitting power, azimuth, M_tilt and E_tilt.

To minimize complexity and focus on the most relevant features, a correlation analysis was performed between the input features and the target variable (path loss). It was observed that azimuth and E-tilt had low correlation and were subsequently removed from the dataset. The performances of each machine learning techniques were tested using all given input features and by removing the features that low correlation. The change was insignificant.

4.1.4. Dataset classification

The dataset was divided into training and testing sets. A suitable ratio or percentage of data was allocated to each set to ensure an adequate amount of data for training and evaluating the machine learning algorithms. From the given dataset 70% were used for training the machine learning models and the remaining 30% were used for testing the model. The train_test_split technique is used for classifying the dataset into training and testing.

4.2. Model development

An ANN, RFR and MLR are the selected models for path loss prediction of Addis Ababa LTE network. From thus models the best model will be provided. Before implementing the machine learning algorithms, parameter tuning was performed to optimize their performance. An ANN and RFR were fine-tuned by adjusting their respective hyper-parameters.

4.2.1. Hyper Parameter tuning of ANN

Hyperparameters are those parameters whose values are predetermined before the learning process starts. These hyperparameters for an ANN consist of the quantity of hidden layers, the number of neurons in each layer, the kind of activation functions, and the optimizer selection. To maximize path loss prediction performance and efficacy, the right hyperparameter values must be chosen.

Several ANN hyperparameter setups are investigated in this work. We design and train ANN models with different numbers of hidden layers: 1, 2, 3, 4, 6, and 7. The test data is used to evaluate each model after it had been trained using the training set. Performance evaluation metrics such as RMSE, MAE, and R^2 are used to assess the models' performance.

Furthermore, the effect of various activation functions on the performance of the ANN is investigated. Activation functions that are tested include sigmoid, ReLU, and a mix of ReLU and linear. Every variation is created and put to the test to see how it affected the accuracy of the predictions.

The performance of the ANN is also greatly enhanced by the optimizer selection. Various optimizers are tested, such as Adam, AdaGrad, and RMSprop. Performance metrics are compared and the ANN model with four hidden layers, ReLU activation function, and corresponding optimizer is assessed. The ANN with the Adam optimizer produced the best outcomes, according to the findings.

Additionally, the number of neurons employed in each ANN layer is taken into account. Since 48 neurons per hidden layer is a frequently used number in pertinent literature, it is employed in this instance. The decision is made to maximize the accuracy of the model by allocating a certain number of neurons per layer, which has a significant effect on the ANN's performance.

In conclusion, using the Adam optimizer and ReLU activation function, the best ANN model architecture for path loss prediction is found to be an ANN with four hidden layers, each with 48 neurons. This model's distinctive architecture is seen in the illustration below.

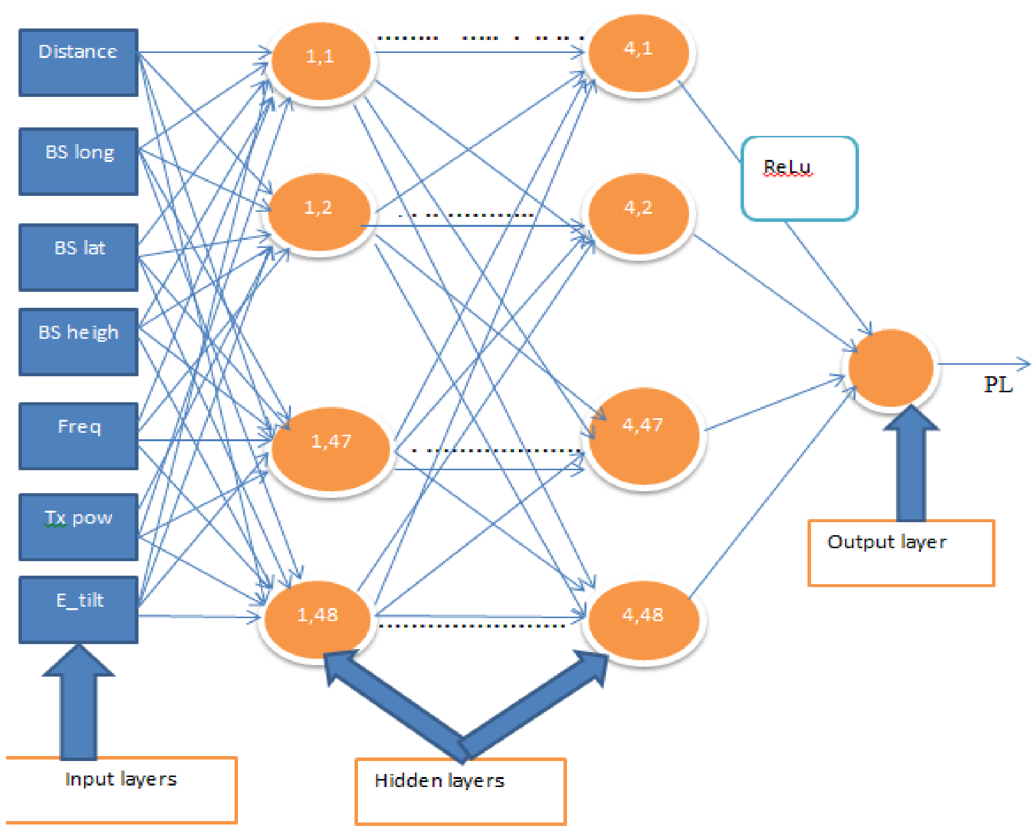


Figure 14: model design and development architecture

An ANN with a dense layer is employed in this thesis. An ANN consists of convolutional and dense layers. They each have benefits and drawbacks of their own. Convolutional layers are typically utilized for image-based problems, while dense layers are used for problems requiring measured or CSV-type data, as well as for problems with limited datasets. Convolutional layers are really also utilized with CSV datasets. Thus, a dense layer ANN is constructed for this particular scenario

4.2.2. Hyper Parameter tuning of RFR

The performance and complexity of RFR with different number of estimators, the minimum number of samples required being at a leaf node and the minimum number of samples required to split an internal node during the construction of a tree are different. Finding an optimum combination of thus parameters is crucial for accurate path loss prediction.

The number of estimator is the maximum depth of each tree in RFR. This value determines the prediction accuracy of RFR. RFR with larger value brings best prediction accuracy but becomes more complex. There is a trade-off between model complexity and prediction accuracy. For this thesis, RFR with 100 numbers of estimators were used.

Min_samples_split: This parameter represents the minimum number of samples required to split an internal node during the construction of a tree. The min_samples_split having values of 5 were used

Min_samples_leaf: This parameter represents the minimum number of samples required to be at a leaf node. The leaf node with 3 samples is used for this thesis.

Hyper-parameters	Values
Number of trees	100
Min_samples_split	5
Min_samples_leaf:	3

Table 1: summery of hyper-parameters with their values for RFR

4.3. Model evaluation

The performance of machine-learning-based path loss models is measured by samples in the test dataset, which do not appear in the model training process. The evaluation metrics include prediction accuracy and complexity.

In terms of evaluating the accuracy, performance indicators like MaxPE, MAE, ESD,CF), RMSE, and MAPE are commonly used. But for this thesis, RMSE, MAE and R^2 are used.

Computational complexity is the measure of the number of computing resources (time and space) that a particular algorithm consumes when it runs. Mainly there are two types of computational complexities. These are time complexity and space complexity.

The time complexity that is the amount of time that a ML algorithm takes for training and testing is used as a performance measuring metric.

4.3.1. Selection of evaluation metrics

For evaluating the performance of the machine learning algorithms, three commonly used metrics were selected. These are Root mean square error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2).

Root Mean Square Error (RMSE): RMSE measures the average deviation between the predicted path loss values and the actual path loss values. It provides an indication of the overall accuracy of the predictions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(y_{m,i} - y_{p,i})^2}$$

Where $y_{m,i}$ *is the measurted path loss at pint i*

$y_{p,i}$ *is the predicted path loss at poin*

Mean Absolute Error (MAE): MAE represents the average absolute difference between the predicted and actual path loss values. It provides a measure of the magnitude of the errors.

$$MAE = \frac{1}{n}\sum_{i=0}^{n}|y_{m,i} - y_{p,i}|$$

Where $y_{m,i}$ *is the measurted path loss at pint i*

$y_{p,i}$ *is the predicted path loss at point i* And n is the number of points

Coefficient of Determination (R^2): R^2 measures the proportion of the variance in the path loss values that can be explained by the input features. It indicates the goodness of fit of the regression models. It ranges from 0 to 1. R^2 having values of 1 indicates, the machine learning algorithm can predict the target values from the input features without any error.

5. Result and discussion

In this section the results found using the proposed models will be presented. The comparison of each model with cost 231 will be described.

5.1. Performance evaluation of ANN under varying hyper-parameter configurations

The performance of ANN is different with different combination of elements. Hence, combining these elements to for better performance is crucial. In this section, the performance of ANN using different setting will be described.

5.1.1. ANN with different hidden layers

Experiments with ANN utilizing varying numbers of hidden layers are carried out in the results section to assess their effect on prediction accuracy. An ANN with 1, 2, 3, 4, 5, 6 and 7 hidden layers is examined, and its performance parameter are the RMSE, MAE and R-squared.

The outcomes showed that better prediction accuracy is achieved by expanding the ANN's number of hidden layers. The accuracy increase gradually up to the fourth hidden layers. Nevertheless, the accuracy gain petered out after 4 hidden layers.

As we can see from the table, the RMSE is decreasing as the number of hidden layers increases from 1 to 4. This suggests that adding hidden layers up to a certain point improves the model's ability to reduce the average prediction error. After the fourth hidden layer, the RMSE remains relatively stable, with slight fluctuations. This indicates that adding more hidden layers beyond four does not have a significant impact on further reducing the average prediction error.

This result implies that adding more hidden layers will eventually lead to a point when prediction accuracy diminishes. The accuracy of the ANN model for path loss prediction is not significantly increased by adding more layers above a certain point.

Overall, these findings show how important it is to choose an ANN's hidden layer count properly. Up to a point, adding more hidden layers can improve accuracy; however, after that, there may be no discernible advances in accuracy and the addition of layers may increase computing complexity without yielding appreciable improvements in prediction performance.

Metrics	Models						
	ANN-1	ANN-2	ANN-3	ANN-4	ANN-5	ANN-6	ANN-7
RMSE	7.04	6.38	6.21	5.78	5.78	5.84	5.79
MAE	5.45	4.76	4.58	4.19	4.22	4.35	4.22
R^2	0.59	0.67	0.69	0.71	0.72	0.72	0.72

Table 2: Comparison of ANN using different hidden layers

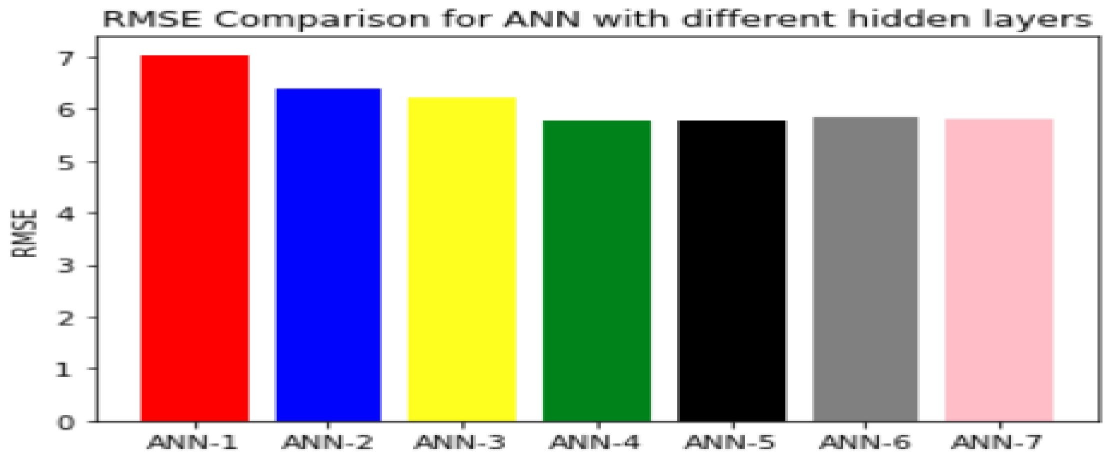


Figure 15: RMSE plot for ANN using different hidden layers

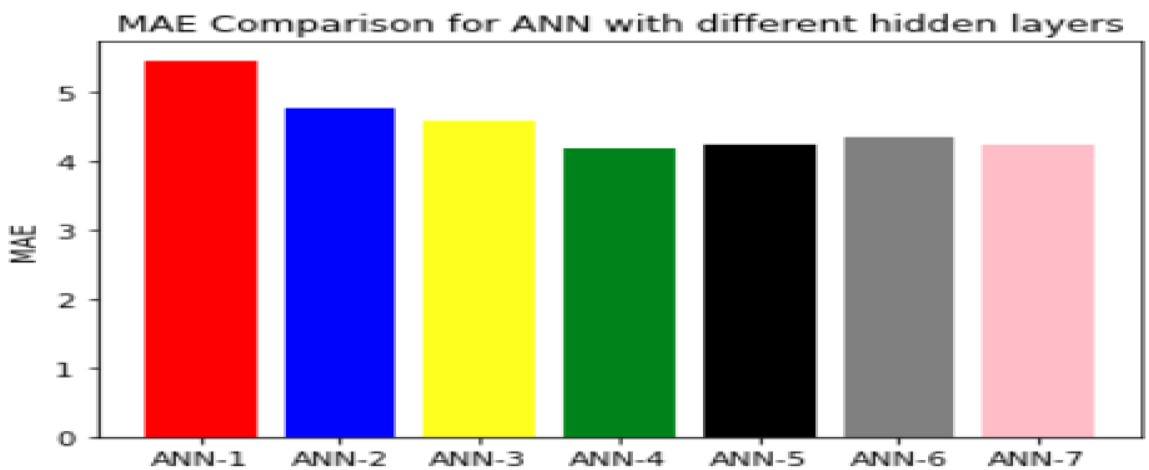


Figure 16: MAE plot for ANN using different hidden layers

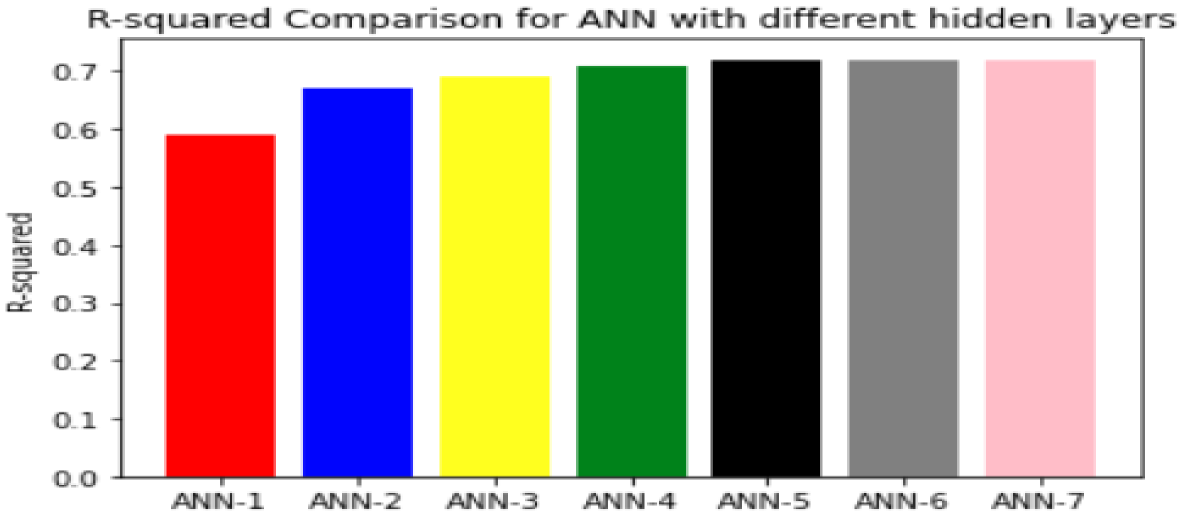


Figure 17: R-squared plot for ANN using different hidden layers

5.1.2. ANN with different activation function

Several activation functions, notably ReLU, sigmoid, and ReLU+linear, are tested in the thesis on path loss prediction using ANNs. Utilizing the RMSE, MAE and R^2 statistic, the effectiveness of the ANN models is assessed. Comparing the ANN model with the ReLU activation function to the other models, the findings shows that the ReLU model performed better, as evidenced by reduced RMSE values.

It is well known that the Rectified Linear Unit (ReLU) activation function may effectively model non-linear relationships in data. By setting positive inputs to their original values and mapping negative inputs to zero, it creates non-linearity. Because of this characteristic, ReLU is a good choice for path loss prediction when capturing intricate patterns.

Metrics	Activation functions		
	ReLu	Sigmoid	ReLu+ Linear
RMSE	5.78	8.81	6.01
MAE	4.19	7.05	4.46
R^2	0.71	0.37	0.70

Table 3: Performance of ANN with different activation functions

5.1.3. ANN with different optimizers

Several optimizers, namely Adam, Adagrad, and RMSprop, were examined in this study on path loss prediction utilizing ANNs. The RMSE metric was used to assess the accuracy of the ANN models. The ANN model with the Adam optimizer has the lowest RMSE of 5.78, according to the results, followed by RMSProp at 6.43 and AdaGrad at 8.42.

These results indicate that, when it comes to path loss prediction using ANNs, the Adam optimizer outperforms both RMSprop and Adagrad. Adaptive learning rates are used by the Adam optimizer for every parameter, facilitating effective convergence and handling of various data types. Its greater success in this particular task is probably influenced by this adaptability. In comparison to Adagrad, RMSprop performs somewhat better and additionally adjusts for learning rates. However, as seen by the smaller RMSE value, Adam outperforms it in terms of accuracy. Of the three optimizers that were examined, Adagrad has the highest RMSE because it modifies learning rates according to the historical gradients of each parameter. This implies that Adagrad might not be as good at capturing the intricate non-linear relationships needed to anticipate path losses accurately.

Metrics	Models(ANN using different optimizer)		
	Adam	RMSpro	AdaGrad
RMSE	5.78	6.43	8.42
MAE	4.19	4.87	6.77
R^2	0.71	0.66	0.42

Table 4: Performance of ANN with different optimizers

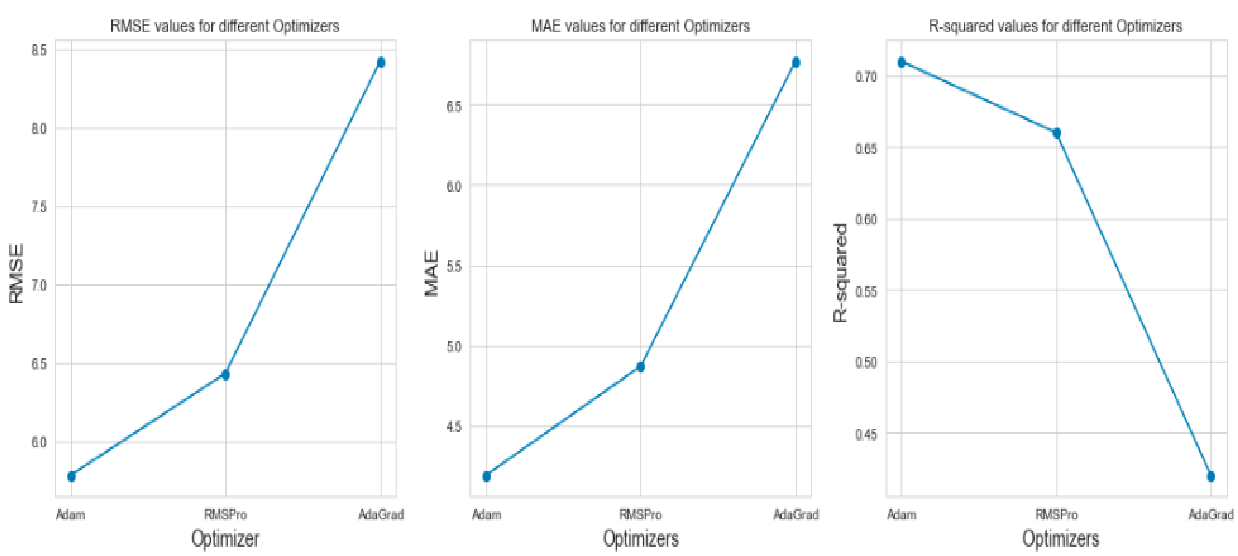


Figure 18: Comparison of ANN with different optimizers using RMSE, MAE and R^2

5.2. Prediction accuracy analysis of proposed models

In this section, the performance of four path loss prediction models using quantitative measuring metrics will be presented. In addition to that, the path loss versus di stance plot of measured path loss against the predicted path loss using each model will be discussed.

5.2.1. Quantitative evaluation of models performance

In this section, quantitative analysis of the four path loss prediction models, namely, ANN, RFR, MLR, and cost 231 models will be discussed. The analysis is done using the results found from python. The path loss predicted by the developed models against measured path loss will be compared, using RMSE, MAE and R^2 as quantitative measurement.

The table 4 and graph 19 shows the RMSE, MAE, and R^2 values for the four models.:

Metrics	Models			
	Cost -231	RFR	ANN	MLR
RMSE	10.43	5.35	5.78	9.02
MAE	8.37	3.48	4.19	7.29
R^2	0.114	0.77	0.71	0.34

Table 5: Comparison of Different models using statistical measuring metrics

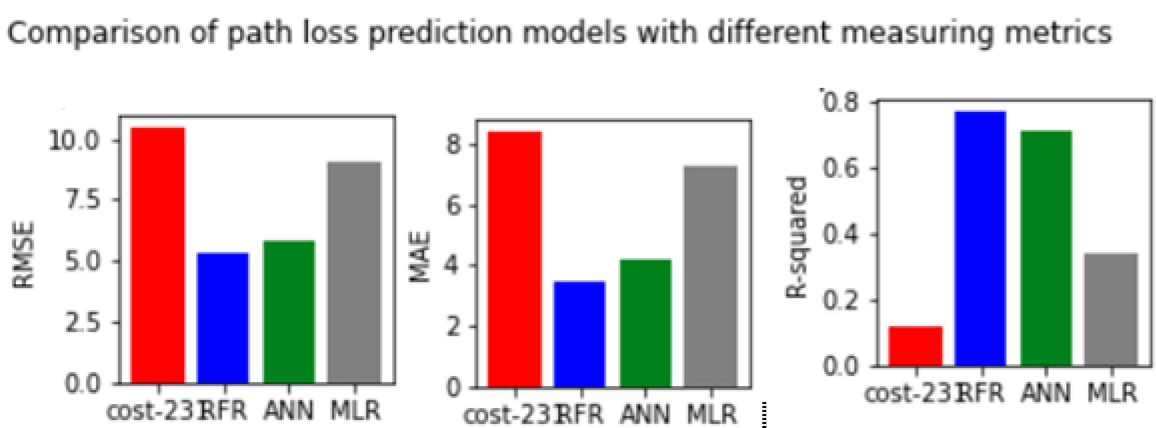


Figure 19: Comparison of proposed models using different measuring metrics

From the result the following points can be mentioned. The ANN and RFR models perform the best, with RMSE values of 5.78 and 5.35, respectively. This indicates that these models are able to make more accurate predictions of path loss than the MLR and COST 231 models. The MLR and COST 231 models exhibit higher RMSE values, of 9.02 and 10.43, respectively. This suggests that their predictions are less accurate than the ANN and RFR models. This is likely due to the limitations of these models in capturing complex nonlinear relationships in the data. The ANN and RFR models also have higher R^2 values, of 0.71 and 0.77, respectively. This indicates that these models are better able to explain the variance in the path loss data than the MLR and COST 231 models.

In general, the results show that the ANN and RFR models are more effective than the MLR and COST 231 models for path loss prediction for the area where measurements are taken.

5.2.2. Models performance evaluation

In this section, the models comparison based on path loss vs distance will be described.

The measured path loss versus distance is shown in the figure below.

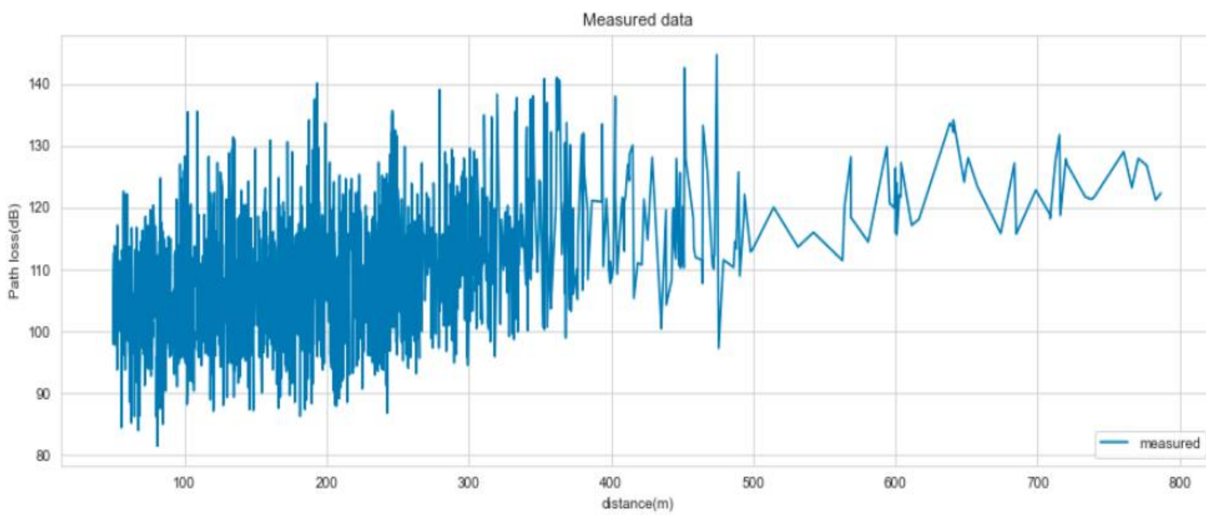


Figure 20: Measured path loss vs distance

The figure above clearly demonstrates that path loss increases as distance increases, adhering to the general principle of path loss theory. However, the increment is not linear, suggesting that other factors play a significant role in determining path loss. From the graph, it can be observed that distance is the primary factor influencing path loss, but it is not the sole determinant

The following figure shows the measured path loss of a radio signal as a function of distance, compared to the COST 231 path loss model.

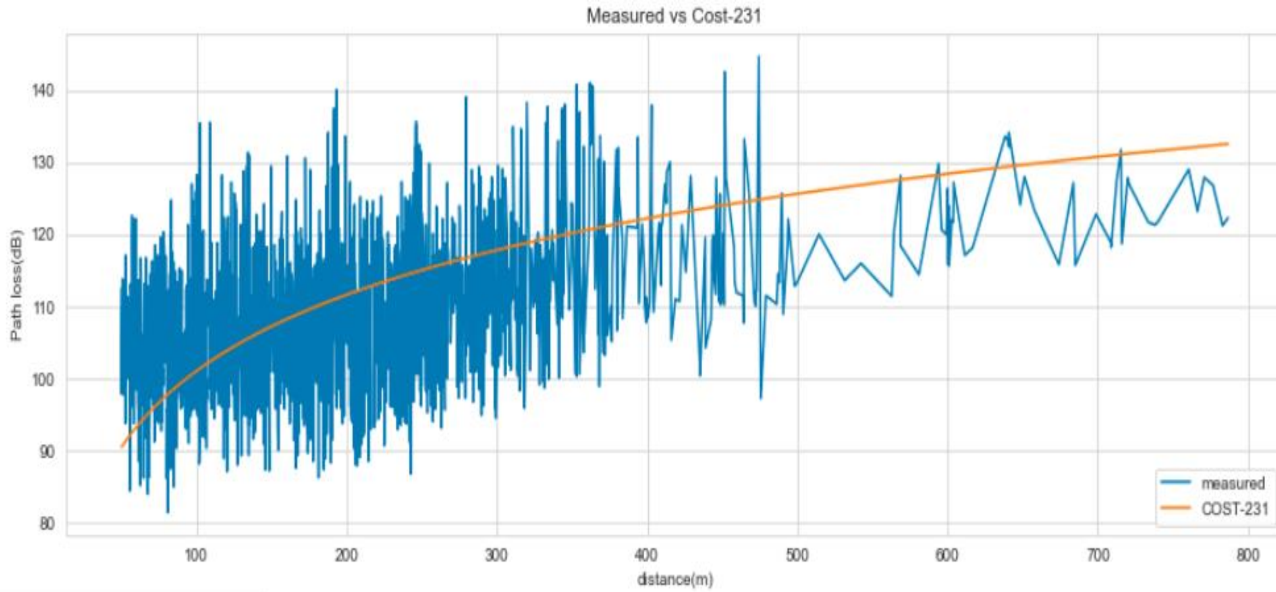


Figure 21: comparison of meaasured path loss with cost-231 model

The graph shows that cost 231 model averages the measured path loss. The finding indicates that the measured path loss is slightly higher than the COST 231 model at short distances. This is because the COST 231 model does not account for the effects of near-field propagation, which can be significant at short distances.

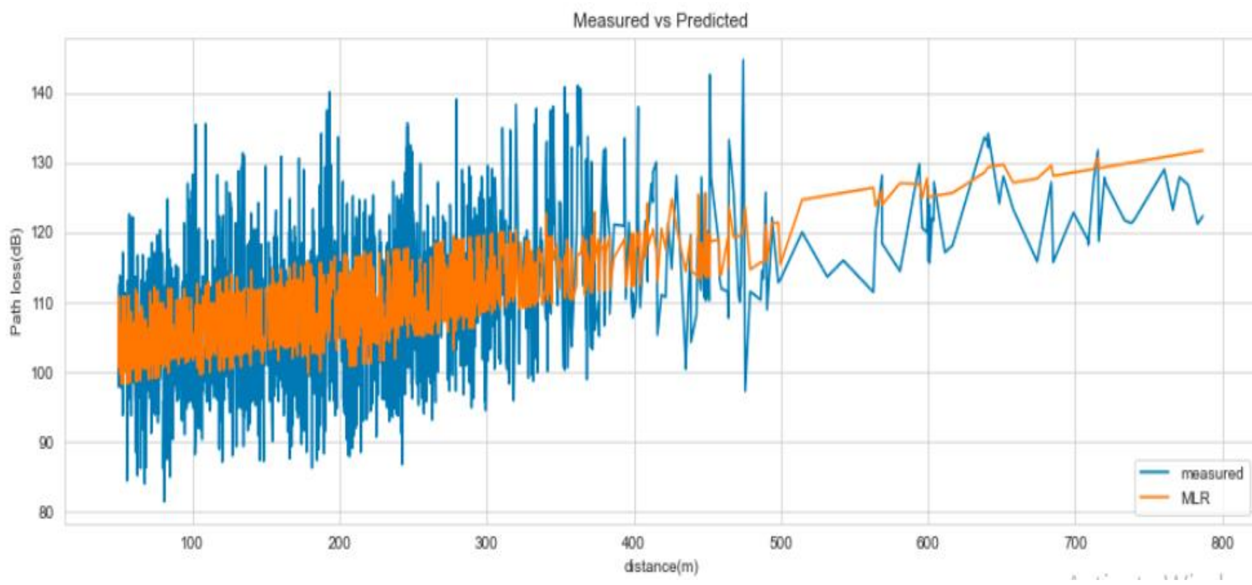


Figure 22: Comparison of measured path loss with path loss predicted by MLR

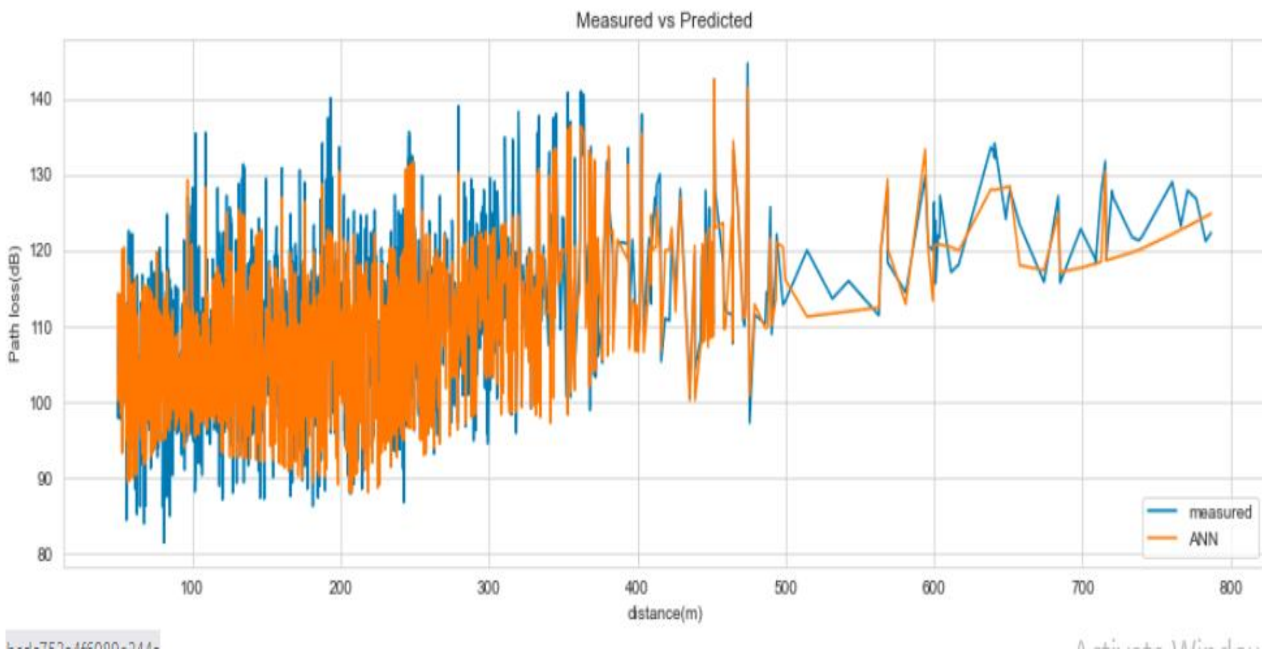


Figure 23: Comparison of measured path loss with path loss predicted by ANN

The measured path loss in the plotted graph against distance closely matches with the predicted path loss obtained from ANN. This indicates a successful performance of the ANN model in predicting the path loss. The graph serves as a validation of the ANN model's ability to generalize well beyond the training data. It shows that the model can accurately predict path loss values for distances that were not explicitly used during the training phase.

The high similarity between the measured and predicted path loss values suggests that the ANN model can reliably estimate the path loss for an AA LTE networks.

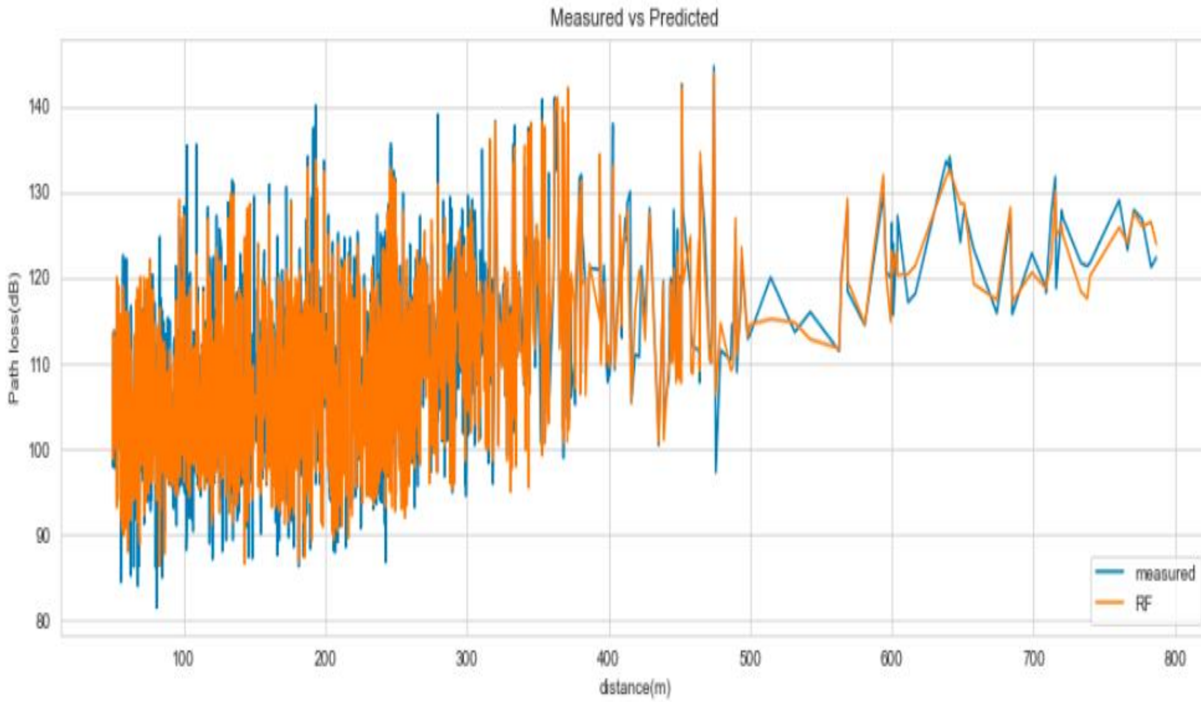


Figure 24: Comparison of measured path loss with path loss predicted by RFR

The path loss predicted by the RFR model is more similar to the measured path loss compared to the path loss predicted by the ANN model. This indicates that RFR is better in predicting the path loss for AA LTE network. The RFR model captures the underlying patterns and relationships in the data more accurately, resulting in a higher degree of similarity between the predicted and measured path loss values.

The close agreement between the predicted path loss by the RFR model and the measured path loss provides a higher level of confidence in the model's outputs. This confidence can be leveraged in making decisions related to network planning, optimization, and performance evaluation.

In general, from figure 21, 22, 23 and 24, the following conclusions can be made.

The comparison between the different models demonstrates that the RFR model outperforms the other models in terms of accurately predicting path loss. It exhibits the highest degree of similarity to the measured values, indicating its superior performance in capturing the underlying patterns and relationships in the data.

The ANN model, although not as accurate as the RFR model, still provides reasonably similar predictions compared to the measured path loss. This suggests that the ANN model captures some of the key factors influencing path loss, but it may not capture the more complex and nonlinear relationships as effectively as the RFR model for the given data and environment.

The MLR model, although less accurate than the RFR and ANN models, still shows a reasonable level of similarity to the measured path loss. This suggests that the MLR model captures some of the key factors influencing path loss, but it may not capture the more complex relationships as effectively as the RFR and ANN models.

The COST 231 model exhibits the lowest level of similarity with the measured path loss among the models considered. It averages the measured path loss.

5.3. Computational complexity analysis

The computational complexity of machine learning models is evaluated in this study. This section will present the time complexity results for the proposed machine learning models.

5.3.1. Time complexity

In this study, the time complexity of three machine learning models: RFR, and MLR are analyzed. Time complexity refers to the amount of time required for training and testing these models.

The primary objective is to assess the computational efficiency of each model and compare their training and testing times.

Figure 25 shown below indicates the complexity comparison of proposed machine learning models.

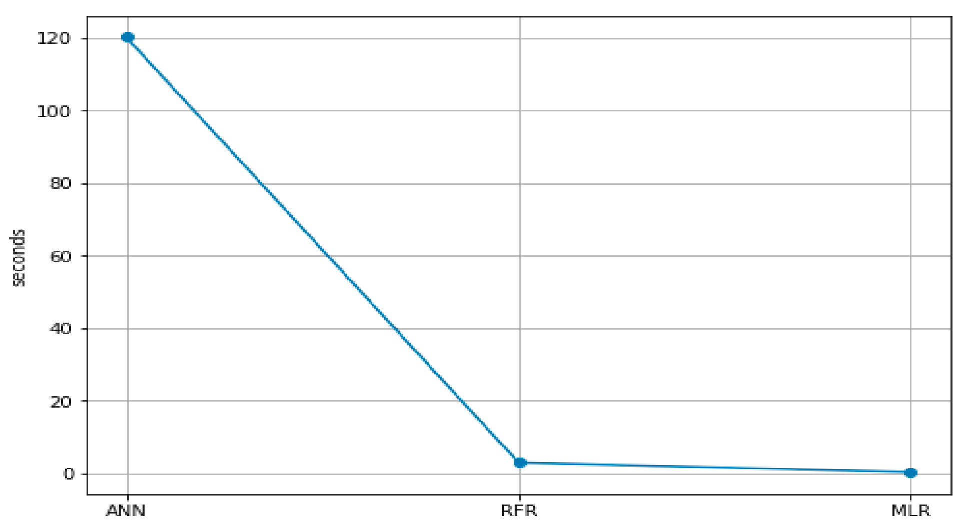


Figure 25: Complexity comparison of ML models

The result in figure 25 and table 5 indicated that ANN exhibited high computational time, meaning it required a significant amount of time for both training and testing processes.

On the other hand, MLR demonstrated the best training and testing times among the three models, indicating its computational efficiency. MLR is known for its simplicity and assumption of a linear relationship between input features and the target variable, resulting in faster computations.

Models	Execution Time (seconds)
ANN	120.07
RFR	3.04
MLR	0.36

Table 6: Execution time of ML models in second

Comparatively, RFR showcased a time complexity that was less than that of ANN but greater than that of MLR. RFR involves an ensemble of decision trees, which adds some computational overhead but is still more efficient than ANN

6. Conclusion and recommendation

6.1. Conclusion

Precise path loss prediction is critical to the effective and dependable operation of wireless communication systems. Usually, deterministic and empirical models are used to forecast path loss. When estimating the attenuation of electromagnetic signals as they propagate through the environment, empirical models make use of measurable data. When these models are used in locations other than the ones for which they were designed, they might not be as accurate. Deterministic models, on the other hand, are more accurate but require a lot of computing resources.

Path loss computation is being done using machine learning models as a solution to these problems. These models are designed to capture the intricate correlations between target values and input factors by learning from massive quantities of measured data. In order to estimate path loss using measured data, the three regression-based machine learning models are examined in this study: ANN, MLR, and RFR.

These models' performance is assessed in terms of computational complexity and prediction accuracy. The predictions of the machine learning models are compared with the well-known Cost 231 model in order to evaluate how accurate they are. Prediction accuracy is measured using statistical performance indicators including R^2, MAE), and RMSE.

From the results, it can be concluded that RFR achieves the highest prediction accuracy, followed by ANN, MLR, and the Cost 231 model. This demonstrates the power of machine learning models in improving path loss prediction compared to empirical models. Additionally, in terms of computational complexity, ANN exhibits the highest complexity, while MLR is the simplest model. RFR falls in between, offering a favorable balance between prediction accuracy and computational requirements.

In summary, RFR emerges as the most suitable model for path loss prediction, delivering both high prediction accuracy and manageable computational complexity. This study highlights the potential of machine learning models as alternative solutions for path loss estimation, outperforming empirical models in terms of accuracy.

6.2. Recommendation

The performance of the machine learning models is dependent on the kind of datasets that are used to train and evaluate them. Therefore, it is essential to do additional research with various datasets and in various settings. This aids in the development of the model that can be used to various settings and circumstances.

Furthermore, other researchers have proposed and evaluated other machine learning models for path loss prediction, such as support vector machines (SVR) and decision tree regression. Therefore, experimenting with different machine learning model types may result in a superior model that is appropriate for the particular context and dataset type.

For additional study, the hybridization of ML models for path loss estimates can be looked into. Furthermore, the combination of ML models with other methods, like ray-tracing approaches may be investigated.

References

[1] Y. Salem and L. Ivanek, “Propagation Modelling of Path Loss Models for Wireless Communication in Urban and Rural Environments at 1800 GSM Frequency Band,” *Adv. Electr. Electron. Eng.*, vol. 14, no. 2, Jun. 2016.

[2] M. Bin Roslee and K. F. Kwan, “Optimization of Hata propagation prediction model in suburban area in malaysia,” *Prog. Electromagn. Res. C*, vol. 13, pp. 91–106, 2010.

[3] Y. Zhang, J. Wen, G. Yang, Z. He, and J. Wang, “Path Loss Prediction Based on Machine Learning: Principle, Method, and Data Expansion,” *Appl. Sci.*, vol. 9, no. 9, p. 1908, May 2019.

[4] Z. J. K. K. T. K. Sarkar, “A Survey of Various Propagation Models for Mobile Communication,” in *Smart Antennas*, Wiley, 2003, pp. 239–307.

[5] A. Seretis and C. D. Sarris, “An Overview of Machine Learning Techniques for Radiowave Propagation Modeling,” *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3970–3985, Jun. 2022.

[6] E. Ostlin, H. Zepernick, and H. Suzuki, “Macrocell Path-Loss Prediction Using Artificial Neural Networks,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, Jul. 2010.

[7] S. I. Popoola, E. Adetiba, A. A. Atayero, N. Faruk, and C. T. Calafate, “Optimal model for path loss predictions using feed-forward neural networks,” *Cogent Eng.*, vol. 5, no. 1, p. 1444345, Jan. 2018.

[8] S. Aldossari and K.-C. Chen, “Predicting the Path Loss of Wireless Channel Models Using Machine Learning Techniques in MmWave Urban Communications,” in *2019 22nd International Symposium on Wireless Personal Multimedia Communications (WPMC)*, IEEE, Nov. 2019, pp. 1–6.

[9] N. Faruk *et al.*, “ANN-based model for multiband path loss prediction in built-up environments,” *Sci. African*, vol. 17, p. e01350, Sep. 2022.

[10] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, “RadioUNet: Fast Radio Map Estimation With Convolutional Neural Networks,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 6, pp. 4001–4015, Jun. 2021

[11] X. Zhang, X. Shu, B. Zhang, J. Ren, L. Zhou, and X. Chen, “Cellular Network Radio Propagation Modeling with Deep Convolutional Neural Networks,” in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, NY, USA: ACM, Aug. 2020, pp. 2378–2386.

[12] K. Inoue, K. Ichige, T. Nagao, and T. Hayashi, “Radio propagation prediction using deep neural network and building occupancy estimation,” *IEICE Commun. Express*, vol. 9, no. 10, pp. 506–511, Oct. 2020.

[13] A. M. Aldosary, S. A. Aldossari, K.-C. Chen, E. M. Mohamed, and A. Al-Saman, "Predictive Wireless Channel Modeling of MmWave Bands Using Machine Learning," *Electronics*, vol. 10, no. 24, p. 3114, Dec. 2021.

[14] N. Moraitis, L. Tsipi, and D. Vouyioukas, "Machine Learning-Based Methods for Path Loss Prediction in Urban Environment for LTE Networks," in *2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, IEEE, Oct. 2020, pp. 1–6.

[15] J. Thrane, D. Zibar, and H. L. Christiansen, "Model-Aided Deep Learning Method for Path Loss Prediction in Mobile Communication Systems at 2.6 GHz," *IEEE Access*, vol. 8, pp. 7925–7936, 2020.

[16] Hiraki F. et.al, "Improved Path Loss Prediction Using Deep Learning Models".

[17] O. Ahmadien, H. F. Ates, T. Baykas, and B. K. Gunturk, "Predicting Path Loss Distribution of an Area From Satellite Images Using Deep Learning," *IEEE Access*, vol. 8, pp. 64982–64991, 2020.

[18] Mekonen.Y, "Outdoor Propagation Pathloss Model for UMTS Networks using Feedforward Backpropagation (FFBP) Neural Network: Case of Addis Ababa, Ethiopia," Addis Ababa institute of technology, Addis Ababa, Ethiopia.

[19] C. Braithwaite, "UMTS Network Planning and Development: Design and Implementation of the 3G CDMA Infrastructure."

[20] Xi Li, *Radio Access Network Dimensioning for 3G UMTS*,. .

[21] Jukka Lempiäinen and Matti Manninen, *UMTS Radio Network Planning, Optimization and QOS Management*. Boston: Kluwer Academic Publishers, 2004. doi: 10.1007/b105973.

[22] S. Sun, T. S. Rappaport, M. Shafi, P. Tang, J. Zhang, and P. J. Smith, "Propagation Models and Performance Evaluation for 5G Millimeter-Wave Bands," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8422–8439, Sep. 2018.

[23] J. Thrane, D. Zibar, and H. L. Christiansen, "Comparison of Empirical and Ray-Tracing Models for Mobile Communication Systems at 2.6 GHz," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, IEEE, Sep. 2019, pp. 1–5.

[24] M. S. Mollel and M. Kisangiri, "Comparison of Empirical Propagation Path Loss Models for Mobile Comparison of Empirical Propagation Path Loss Models for Mobile Communication," Jan. 2014.

[25] Y. Singh, "Comparison of Okumura, Hata and COST-231 Models on the Basis of Path Loss and Signal Strength," *Int. J. Comput. Appl.*, vol. 59, no. 11, pp. 37–41, Dec. 2012.

[26] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.

[27] D. J. Cichon and T. Kürner, “Propagation prediction models,” 1995, pp. 115–207.

[28] T. M. Mitchell, “Machine Learning,” USA: McGraw-Hill, Inc., 1 ed, 1997.

[29] K. Mason, J. Duggan, and E. Howley, “Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks,” *Energy*, vol. 155, pp. 705–720, Jul. 2018.

[30] M. J. Somers and J. C. Casal, “Using Artificial Neural Networks to Model Nonlinearity,” *Organ. Res. Methods*, vol. 12, no. 3, pp. 403–417, Jul. 2009.

[31] M. Alnatoor, M. Omari, and M. Kaddi, “Path Loss Models for Cellular Mobile Networks Using Artificial Intelligence Technologies in Different Environments,” *Appl. Sci.*, vol. 12, no. 24, p. 12757, Dec. 2022.

[32] S. Vani, T. V. Madhusudhana Rao, and C. Kannam Naidu, “Comparative Analysis on variants of Neural Networks: An Experimental Study,” in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, IEEE, Mar. 2019, pp. 429–434.

[33] Ajith Abraham, *Artificial Neural Networks*, vol. 66. Ed. United States of America: John Wiley & Sons Ltd, 2012.

[34] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 103. New York, NY: Springer New York, 2013.

[35] L. Breiman, “Random Forests," Machine Learning",” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001,

[36] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York, NY: Springer New York, 2009.

[37] O. Bamisile, A. Oluwasanmi, C. Ejiyi, N. Yimen, S. Obiora, and Q. Huang, “Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions,” *Int. J. Energy Res.*, vol. 46, no. 8, pp. 10052–10073, Jun. 2022.