April 2020

Artificial Intelligence Towards the Wireless Channel Modeling Communications in 5G

Saud Mobark Aldossari
University of South Florida

Artificial Intelligence Towards the Wireless Channel Modeling Communications in 5G

by

Saud Mobark Alhajaj Aldossari

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Kwang-Cheng Chen, Ph.D.
Nasir Ghani, Ph.D.
Randy Larsen, Ph.D.
Ismail Uysal, Ph.D.
Srinivas Katkoori, Ph.D.

Date of Approval:
April 1, 2020

Keywords: MmWave, Channel State Information, Data Driven, Path Loss, Neural Network

Dedication

I convey a special thanks to almighty God for the completion of this journey. This dissertation is dedicated to my parents, brothers, and sisters in all my endeavors. Special appreciation and dedication to my wife and my kids Ghenna, Sumu, and Mobark for all the encourage and love they have given me during this journey. To all of my friends and colleges, these hard days will not pass without you, God bless you all!.

Acknowledgments

I'm remarkably grateful to Professor Kwang-Cheng Chen for being my supervisor and this journey would not be achievable without the support, guides, and tremendous encouragement. Furthermore, I would like to express my deepest appreciation to my committee members Professor Nasir Ghani, Professor Randy Larsen, Professor Ismail Uysal and Professor Srinivas Katkoori for reshaping my dissertation. I also spread my sincere appreciation to my academic crew for being there with me during my doctoral journey.

Table of Contents

List of Tables

List of Figures

Abstract

Channel prediction is a mathematical predicting of the natural propagation of the signal that helps the receiver to approximate the affected signal, which plays an important role in highly mobile or dynamic channels. The standard wireless communication channel modeling can be facilitated by either deterministic or stochastic channel methodologies. The deterministic approach is based on the electromagnetic theories and every single object in that environment has to be known in that propagation space and an example of this method is ray tracing. While the stochastic modeling method is based on measurements that involve statistical distributions of the channel parameters and an example of this approach is Floating-Intrcept (FI) model. In other words, channel modeling uses mathematical parameters to obtain the effect of the channel medium. These effects cause the transmitted signal to be either destructive or constructive during the propagation. Where the main focus of this dissertation is how Artificial Intelligence will be used in channel modeling. Fifth-generation -5G- with massive MIMO, higher data rate, handover, and channel modeling become more and more complex with the new wireless generations than the traditional stochastic or deterministic approaches. Nowadays, traditional wireless communication channel modeling is considered an old fashion especially with new technologies era such as things that applies to MmWave. In this sense, researchers and academia looking forward to more effective methods that have less complexity and more accuracy. Emerging machine learning technology supplies a new direction to process big measurement data and traffic data toward the wireless channel. Thus, new novel strategies of channel learning are proposed to generate a model free of the wireless channel modeling by tacking these difficulties.

With the availability of high computational devices and data, Artificial Intelligence (AI) emerges to revolutionize system design for new radio 5G. The subcategories of AI involve machine learning, deep learning techniques such as supervise leaning methods will be used to predict the channel state information (CSI) of a variate of environments base on a certain dataset. The fundamentals of wireless communication systems concentrate on channel modeling particularly for new frequency bands such as MmWave. Machine learning can facilitate rapid channel modeling for 5G wireless communication systems due to the availability of partially relevant channel measurement data and models. When irregularity of the wireless channels leads to a complex methodology to achieve accurate models, appropriate machine learning methodologies explore to reduce the complexity and increase the accuracy. This dissertation demonstrates an introduction to alternative procedures beyond traditional channel modeling to enhance CSI prediction based on data-driven with the usage of AI techniques, to alleviate the dilemma of channel complexity and time-consuming process that the measurements take. An example of applying AI towards wireless channel modeling is applying regression techniques with measurement data of a certain scenario to successfully assist the prediction of the path loss model of a different operating environment.

The irregularity of the wireless channel leads to a complex wayside to achieve accurate models where new technology is required to accomplish the precise results with new technologies. Machine learning algorithms involve channel modeling to reduce complexity and increase the accuracy that reduces the number of measurements. Furthermore, researchers explore machine learning methods that can link wireless channel modeling in different systems. Due to a large number of operations and extensive measurements, the researcher tends to perform machine learning to enhances the channel modeling prediction. The aim of using the machine learning algorithms is to develop alternative techniques to estimate the received signal that's usually get affected by the channel. Moreover, extract and develop useful information from channel measurement data in the wireless communication system. Lacking using machine learning (ML) techniques for mobile wireless channel models is overcomes throughout this dissertation. By applying ML algorithms such as classification techniques will be used to investigate the wireless channel modeling and compare the result of each model by using the interpretation of performance measures such as accuracy, precision, and the number of misclassifications.

Artificial intelligence (AI), particularly machine learning (ML), is widely studied to enable a system to learn of intelligence, predict and make an assessment instead of the needs of humans. Switching the traditional channel modeling to machine learning channel modeling still in its early stage. One of the main issues in current communication is accurate of prediction the channel parameters, whereas using machine learning techniques could enhance the prediction and reduce the complexity. ML can be used to predict and estimate the wireless channel parameters and examine large and small-scale fading including parameters such as path loss, delay path loss exponent, frequency, Doppler spread and random variable that explains the large scale fading.

Usually, the supervised learning can be divided into two main subjects which are the regression and classification learning. The regression method is considered continuous values whereas the classification is a discrete value. Both of these two subjects are useful in estimating the channel parameters such as the path loss component and the large-scale random variable. The error can be minimized if optimization techniques are involved or by modifying some machine learning algorithms. One of the topics in this dissertation is to assist base stations to select the optimum signal based on the availability of the CSI data. This approach emerges as an attractive technique in the radio access network (RAN) and link selection to result in the strongest propagated link becomes the critical technology to facilitate RAN using mmWave. By taking advantage of existing operating data and apply appropriate artificial neural networks (ANN) algorithms to alleviate severe fading in the mmWave band. Additionally, we applied classification techniques using ANN with multilayer perception to predict the path loss of multiple transmitted links and base on a certain loss level, and thus execute effective relay selection, which also recommends the handover to an appropriate path. ANN with multilayer perception is compared with other ML algorithms to demonstrate effectiveness for relay selection in 5G-NR. Thus, machine learning (ML) is a new way that will change the design, standardization, and optimization of the communication systems. ML techniques such as supervise leaning and unsupervised methods will be used to estimate the wireless channel parameters by inferring CSI based on data-driven since the propagation signal of communication systems fundamentals is focusing on channel modeling particularly for a new technology era such as MmWave.

Millimeter-wave supplies alternative frequency bands of wide bandwidth to better realize pillar technologies of enhanced mobile broadband (eMBB) and ultra-reliable and low-latency communication (uRLLC) for 5G - new radio (5G-NR). With the usage MmWave frequency band, propagated signals become weaker due to fading through the channel and there must be a way to predict the strongest signal in a quick way to avoid the delay and to assist the coverage. With the usage of AI techniques and based on data-driven, predicting the strong signal can be achieved.

This dissertation focuses on predicting channel state information based on data-driven and elaborates on how to overcome some wireless issues in the new era 5G using Artificial Intelligence. Thus, based on our investigation in this dissertation, we can conclude that applying artificial intelligence towards wireless channel modeling is a promising technique and should be implemented in current and future wireless communication systems.

Chapter 1: Introduction

The main aim of this dissertation is to contribute to the existing literature on the field of wireless communication systems by serving readers with aggregating materials and providing more novel works by introducing a new idea to the wireless communication systems. Involving and combining Artificial Intelligence (AI) to wireless channel modeling is one of the promising methods to reach the 5G requirements. Starting from this section, which describes the background and motivation of this dissertation. The following chapters of this dissertation can answer these pursuing questions.

- The classical wireless channel modeling is time consuming, complex and inaccurate, how to overcome these drawbacks?
- How to enhance the accuracy of prediction the channel state parameters such as path loss and other channel state information?
- How to meet the five generation (5G) and new radio (NR) requirements?
- How artificial intelligence (AI) can improve wireless communications systems?
- How machine learning can be involved in the wireless communication systems?
- How to assist base stations to predict the acceptable received signals?
- 5G and beyond wireless generations have a range of frequency bands, how can a base station be assisted to predict a specific frequency band?

1.1 Motivation

During the past two decades, academia and industrial researchers have been attracted to investigate and mimicking wireless channel fading using mathematical models to predict the wireless channel information. Nowadays, these investigations have achieved a level with no further improvement to reduce the complexity of characterizing the wireless channel medium. With the usage of the high computational device and the availability of data, artificial intelligence techniques are involved to enhance the prediction and reduce the complexity to reach reliable communication. With the usage of AI towards wireless systems, we are now able to reach the 5G-NR requirements that are based on three pillars which are enhanced mobile broadband (eMBB), ultra-reliable, low latency communications (URLLC), and massive machine-type communications (mMTC).

1.2 Outline of Dissertation

In this subsection, the structure of the dissertation will be illustrated:

- Chapter 2: The introduction of the theoretical background and the state of art of wireless channel modeling.
- Chapter 2: Overview of some artificial intelligence methods.
- Chapter 3: Applying artificial intelligence techniques towards wireless channel modeling.
- Chapter 4: Predicting the path loss of wireless channel models using machine learning techniques in mmWave urban communications.
- Chapter 5: Predicting optimum propagated link for 5G new radio via artificial neural networks.

- Chapter 6: Investigation and Prediction of the Wireless Channel Modeling for High Frequency Bands Using Artificial Intelligence Methods.

- Chapter 7: Future works and conclusion

1.3 Dissertation Contribution

Comparable to every field in real life, the wireless communication field faces issues especially with the revolution of wireless communication in the current decade. This section will cover some of the case studies that we have investigated and how it will be resolved. Additionally, an overview of the wireless channel will be addressed.

1.3.1 Complexity and Inaccuracy

The traditional wireless channel modeling using Deterministic or Empirical approaches is complex [1]. This complexity is due to a large number of operations and large conducting measurements. The irregularity and complexity of the wireless channel are in need of new methodologies to meet the new wireless generation requirements. To be more specific, the complexity and the accuracy of the path loss models can be varying with many factors such as environments, interference levels, energy, distance, etc. Thus, Artificial Intelligence based on data-driven is proposed in full details in this dissertation to overcome this issue.

1.3.2 Link Selection

Wireless communication systems are facing a problem with selecting the strongest propagated signals. Base station needs to be assisted to select the optimum received signal. However, we're proposing a method to avoid the fading is by allowing the base station to select the strongest link with the help of machine learning techniques to improve the reliability of the wireless communications based on data driven.

1.3.3 Frequency Bands Prediction

Predicting the frequency bands is not considered an issue in the previous wireless generations, whereas nowadays, it's becoming more critical due to the availability of the bands ranging from 700 MHz to 100 GHz in 3G to 5G respectively. The base station has no strategies to make such a prediction suddenly. Thus, we are proposing new methodologies to assist base stations to has the ability to predict the frequency bands to make the communications more reliably to meet the 5G pillars.

1.4 Background

The main purpose of the wireless channel modeling in the area of cellular systems is to support basic physical layer communication. Channel modeling, intends to mathematically describes the multipath components (MPCs) characterization and can be implemented by either Deterministic or Stochastic manner [2]. The electromagnetic waves propagated through a channel and how these waves get affected by the surrounding environments require a way to represent it. A mathematical representation is used to describe the channel impulse response in the time domain or by channel transfer function in the frequency domain and this called a model-based approach. While a second option is a model-free approach, which is based on data driven to estimate and infer the CSI and that would be the main goal of this dissertation. Traditional wireless communication channel modeling based on either Deterministic or Stochastic models and the irregularity of the wireless channel's behavior leads to a complex and time-consuming methodology to achieve accurate models. The fifth generation of mobile communication (5G) and beyond systems suggest further complicated tasks to obtain appropriate channel models. Classical wireless procedures are becoming more complex handling big data and it's the time for an alternative procedure such as Machine Learning (ML) [3]. Also, wireless channel modeling prediction is a critical mission in the wireless communication systems. A disruptive approach to more effectively tackle the technological dilemma in channel modeling is to take advantage of existing measurement data and apply appropriate machine learning (ML) algorithms to reduce complexity and increase accuracy. After intuitive application of regression techniques to channel modeling [4], we demonstrate applying classification techniques in ML to predict the path loss of channel modeling, while showing how the complexity such as the required number of measurements can be greatly reduced. Consequently, appropriate applying ML can assist more efficient development of channel modeling in new scenarios for 5G - new radio (5G-NR)..

1.5 Channel Modeling Literature Review

The main aim of the channel modeling in the area of cellular systems is to support basic physical layer communication. Channel modeling [5], intends to mathematically describe the multipath components (MPCs). Channel modeling can be implemented by either Deterministic or Stochastic manner, to the reorientation of the propagation waves through a channel and how these waves get affected by the channel surrounding. Then, a mathematical representation describes the channel impulse response in the time domain or by channel transfer function in the frequency domain. In parallel with the measurement and theoretical research, the simulation tool is building to reduce the complexity and improving the performance.

Channel modeling of wireless communication [6] can be represented mathematically by the time-variant CIR impulse response $h(t, \tau)$ in the time domain or in the frequency domain by using time-variant channel transfer function $H(t, f)$. The channel impulse response can be obtained by sounding the channel periodic pulses with PN sequences or pulse at the transmitter while in the frequency domain, the channel can be sounded with multi-tone signals such as vector signal generator (VSG). Both time and frequency domain describe the behaviors of the signals where the fourier transform is used to obtain the transfer response in the frequency domain. The mathematical models can include parameters such as carrier frequency, shadow, phase, Doppler, delay, distance, path loss and others where most of these parameters are varied randomly [7].

Fading is identified as the time variation of the received multipath components MPCs power that went through the medium channel effects where it can be categorized into two parts as Small-Scale Fading and Large-Scale Fading. Propagation in wireless channel models results in either large or small scale fading base on the variations of the signal. Large-scale fading can be either path loss or shadowing due to shadowing by blockers that are larger than the carrier wave. While small-scale fading is due to amplitude variations due to multipath time delay or Doppler spread.

Before applying machine learning techniques, we need to understand the fundamental of the wireless channel and how to model it. There are two types of channel modeling as mentioned previously, which are Deterministic and Stochastic Channel Models. Others may divide the wireless communication channel models into three categories as geometry based deterministic models such as Ray tracing, non-geometrical stochastic models as the empirical models and geometry-based stochastic models (GBSMs) which is a combination of deterministic and empirical models. Deterministic channel modeling: The main idea of this is to reproduce the communication by using Maxwell's equations. The CIR will be produced using such as the ray-tracing model with geographical data. The disadvantage of this approach is every single location has different surfaces that require different designs especially with urban areas that require a different scenario and it demands a highly precise geographical database. In this approach, all objects such as buildings, cars, trees, and others are divided into a small slot that gathered in the ray tracing. There are other models such as finite-difference time-domain (FDTD). The time-variant CIR is expressed for frequency selective channel as:

$$y(t) = u(t) * h(\tau, t) = \int_{-\infty}^{\infty} h(\tau, t) u(t - \tau) d\tau \tag{1.1}$$

where y(t) is the received signal, u(t) is the transmitted signal, * is the convolution sign and $h(\tau, t)$ is the channel function with respect to delay and time.

$$h(\tau, t) = \sum_{K=1}^{N(t)} a_k(t) e^{j(2\pi f \tau_k(t))} \delta(\tau - \tau_k(t)) \tag{1.2}$$

Furthermore, alpha is the amplitude, K is the MPC at time t and τ is the delay. While stochastic channel modeling approach can be divided into two classifications, geometry-based stochastic models (GSCMs) and non-geometry based stochastic models. The Non-geometry models can either a narrowband stochastic channel models or wideband stochastic channel model [6]. Where the narrowband stochastic model usually focuses on the characterization of the fading statistics as well as the Doppler spectrum. Whereas the wideband stochastic model is typically focused on the received power, delay, angle of departure and arrival and Doppler shift, etc. However, the geometry base stochastic models which are similar to the non-model but with a simplified ray tracing. Most of the current MIMO channel modeling uses this model for a wideband system. The difference between wideband and narrowband is the amount of the bandwidth however in narrowband, the delay has no much affects base on the relation $T_m \approx B^{-1}$ and $u(t - \tau_n(t)) \approx u(t)$ but with wideband all of these approximations are not valid. T_m is the multipath delay spread, B is the bandwidth, τ_n is the delay and $u(t)$ is the low pass signal. Moreover, narrowband characterization can be used to model the path loss and coherence time. While wideband used to characterize the channel impulse response to investigate the delay parameters since the narrowband measurements aren't able to resolve the individual multipath components in the time domain [8]. Our focus is the non-geometrical stochastic which is based on real measurement campaigns to collect data to build new models for MmWave bands using artificial intelligence methods.

1.5.1 Objective

During the past decade, wireless channel modeling has gained academic and industrial intentions. Today, there are a couple of challenges to create channel modeling communications and that is due to complexity and inaccuracy. Base on the literature, multiple ways to overcome these issues using MmWave and AI. the usage of higher frequency and AI will assist base stations to interchange information and making decisions. Federal Communications Commission (FCC) already approved using higher frequencies such as 28, 37 and 39 GHz and frequencies between 64-71 GHz for unlicensed for mobile usages. Since there are challenges to implement MmWave frequencies in the communication systems. The current issues with MmWave models are the lack of accurate MmWave channel model, high fading due to vulnerability, mobility and short distance propagation due to low wavelength. During the empirical measurements, a couple of characteristics in the channel modeling need to be investigated such as radio signal strength (RSS), delay spread, Doppler spread, phase shift and bit error rate. Involving these characteristics usually used to improve the performance of communications.

Fading in the wireless communications can be categorized into two parts, large scale and small scale fading where fading refers to the reduction changes of the MPC's power. The large scale fading (LSF) explains the main characteristics of the channel such as path loss and shadow. Furthermore, LSF cases examine the relationship between path loss and the separated distance between the Tx and Rx in different environments such as the suburban. While small scale fading cases go over the relation between the Doppler spread and time coherence. Other characteristics will be examined as well as gradually such frequency bands to assist base stations. These investigations are needed due to the lack of real measurements which will allow developing a reliable wireless communication. The propagation channel of communication systems fundamentals is focusing on channel modeling particularly for a new technology era such as MmWave. The irregularity and complexity of the wireless channel lead to solid ways to achieve accurate models where more trails are always required to accomplish the precise results mainly with new technologies.

Wireless communication systems are non-stationary channels where the medium channel changes during a short time. That change is due to the mobility of one or all transceivers or the environment by itself such as other moving vehicles cause time-variant. In channel modeling, there are six different domains that other parameters can be assigned to it.

- Loss and fading: path loss, shadow fading, and small-scale fading
- Time domain: coherence time and stationary time
- Frequency domain: coherence bandwidth and stationary bandwidth
- Doppler domain: Doppler shift and Doppler spread
- Delay domain: delay spread and mean delay spread and excess delay spread.
- Angular domain: angle of arrival (AoA) and angle of departure (AoD)

The primary concern of wireless channel modeling characteristics is the channel strength variations over time and frequency where those variations can be divided into large scale fading (LSF) or small scale fading (SSF). Regularly, in LSF, the signal loses its strength due to distance in case of the transceiver are far away from each other or due shadowing by environments obstacles. When the signals or multipath gets destructive or constructive, small scale fading occurs and different statistical models can be used to characterize the wireless channel medium such as Rayleigh Model and Nakagami-n (Rice) Model [9]. To investigate the wireless channel, an impulse response is a preform as shown in the following section.

1.5.2 Impulse Response

The channel impulse response (CIR) is a short duration time domain signal which is the impulse response $\delta(t)$ for continuous-time system or $\delta[n]$ for discrete-time systems. Where the system's impulse response $h(t)$ is the output signal of a channel when an impulse is applied to it. CIR can be obtained by the correlation of the received samples with PN sequence in time domain sounding. The reason why CIR is important is predicting the output of the system. CIR can be also obtained in the frequency domain by using the Fourier transform and become channel frequency response (CFR). Both of them can be used to characterize the performance of the system where CIR is helpful when analyzing behavior in the time domain whereas CFR in the frequency domain There are a couple of forms of time-variant channel modeling and the below is a wireless channel model that can be as a linear filter with time-varying impulse response:

$$Y(t,\tau) = x(t) * h(t,\tau) = \int_{-\infty}^{\infty} x(t)h(t,t-\tau)d\tau \tag{1.3}$$

The CIR donated by h(t,τ) which is a function of time (snapshot) and delay τ. Usually, the peak power of the CIR supposed to be close to 20-30 dB above the noise threshold to become a valid CIR whereas the multipath components gain will be set to zero if it's gain only 3dB or lower of the noise threshold. Also, one of the ways to record a time-variant channel is by assuming the estimated largest Doppler frequency value which is needed in estimating the time of the snapshot. The time period between each adjacent snapshot is based on Nyquist rate as $f_{sample} \geq 2v_{max}$ where v_{max} is the Doppler frequency. Moreover, the upper bound of the snapshot period has to be estimated to prevent overlapping with other adjacent snapshots. Where that can be done by knowing the maximum excess delay where the lower bound is estimated by the Nyquist rate.

The traditional method to estimate the fading of a wireless channel is by mimicking the channel effect using statistical methods depending on the environments. For instance, characterizing the amplitude and the phase effect of a wireless channel can be illustrated as follow.

$$h(t) = \sum_{i=0}^{L-1} a_i e^{-j2\pi f_o \tau_i} \tag{1.4}$$

$$= \sum_{i=0}^{L-1} a_i cos(2\pi f_o \tau_i) - j \sum_{i=0}^{L-1} a_i sin(2\pi f_o \tau_i) \tag{1.5}$$

if we assume

$$y = a_i cos(2\pi f_o \tau_i) \tag{1.6}$$

$$x = a_i sin(2\pi f_o \tau_i) \tag{1.7}$$

To obtain the amplitude a and the phase ϕ of an affected propoaged signal through a wireless channel.

$$a = \sqrt{x^2 + y^2} \tag{1.8}$$

$$\phi = Tan^{-1}\frac{y}{x} \tag{1.9}$$

Due to the scatting of the multipath channel, at the receiver side, there will be multiple components of the transmitted signal. Based on Central Limit Theorem (CLT), The random variable a_i and τ_i follow Gaussian Distribution (standard normal distribution) with zero mean and unit variance. Thus, x and y follow the same distribution.

$$pdf(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.10}$$

$$F(x) = \frac{1}{\sqrt{\pi}} \exp^{-x^2} \tag{1.11}$$

$$F(y) = \frac{1}{\sqrt{\pi}} \exp^{-y^2} \tag{1.12}$$

Then, by using the joint distribution

$$F_{x,y}(x, y) = F(y) \cdot F(x) = \frac{1}{\pi} \exp^{-(x^2+y^2)} \tag{1.13}$$

$$F_{x,y}(x, y) = F(y) \cdot F(x) = \frac{1}{\pi} \exp^{-a^2} \tag{1.14}$$

So, by applying a and ϕ in term of joint distribution, we get:

$$F_{x,y}(x, y) \implies F_{a,\phi}(a, \phi) = \frac{1}{\pi} \exp^{-a^2} \cdot J_{x,y} \tag{1.15}$$

where $J_{x,y}$ is the Jacobian factor for changing the variables and after calculating the Jacobian determination, the factor will equal to a. Now, the final caractrization of the fading channel using the joint distribution would be as following.

$$F_{a,\phi}(a, \phi) = \frac{a}{\pi} \exp^{-a^2} \tag{1.16}$$

Now, to obtain the individual distribution of the amplitude and the phase, we need to apply the marginal dissertation from the joint distribution. Let's start with driving the amplitude and then driven the phase would be the same procedure.

$$F_a(ai) = \int_{-\pi}^{\pi} F_{a,\phi}(a,\phi) d\phi \tag{1.17}$$

$$F_a(a) = \int_{-\pi}^{\pi} \frac{a}{\pi} \exp^{-a^2} d\phi \tag{1.18}$$

$$F_a(a) = \frac{a}{\pi} \exp^{-a^2} 2\pi = 2a \exp^{-a^2} \tag{1.19}$$

The above equation is called Rayleigh fading channel. Then, the phase distribution shown as.

$$F_\phi(\phi) = \int_{0}^{+\infty} \frac{a}{\pi} \exp^{-a^2} da \tag{1.20}$$

$$F_\phi(\phi) = \frac{1}{\pi} \tag{1.21}$$

1.5.3 Large-Scale Fading

Large scale fading usually due to the objects that shadow the signal and distance between transmitter and receiver. Path losses and shadowing can cause a large fading to the transmitted signals that degrade the signal strength. However, channel modeling has not been well studied in the 5G. Most of the published path loss models follow the general path loss model in the below equation.

$$PL(\text{in decibels}) = [\text{Free space loss at 1M}] + 10nlog(d) + X\sigma \tag{1.22}$$

where $X\sigma$ is the normally distributed random variable mean and σ standard deviation plus other unknown losses added to this parameter during fitting the model which in our analysis would be investigated more. Additionally, the receive signal strength varies over distance r as $\frac{1}{r^2}$ for free space or $\frac{1}{r^4}$ for ground reflection [10]. Additionally, the relationship between path loss and transmitted power is showing as:

$$PL[dB] = P_t[dBm] - P_r[dBm] + G_t[dB] + G_r[dB] \tag{1.23}$$

where PL is the path loss, P_t is the transmitted power, P_r refers to the received power, G_t and G_r is the transmit and receive antenna gain, respectively

1.5.4 Small-Scale Fading

Small scale fading is due to the constructive and destructive of the Multipath components MPCs at the receiver side which usually leads to fluctuations over small distance or small interval of the signal envelope [6]. Besides, Small-scale fading is used to determine the link-level performance based on bit error rates and average fading. Usually, the amplitude distribution of the received signal is modeled like a Rayleigh distribution for the non-line of side where Ricean distribution is used for the Line of side. Where Rayleigh fading models are the most common statistical model for the random amplitude. Some classes of such fading are summarized as follows:

1.5.4.1 Flat Fading

Flat fading channels are known as amplitude varying channels and narrowband channels (signal BW is narrow compared to channel BW). Usually, flat fading occurs if the signal width pulse or symbol period is larger than the delay spread $T_s >> \sigma_\tau$ or the signal bandwidth is less than the channel bandwidth $B_s << B_c$.

1.5.4.2 Frequency Selective Fading

Frequency selective fading is due to the time spread of the transmitted symbols within the channel [11]. Frequency selective fading is also known as wideband channels and occurs if the signal width pulse or symbol period is less than the delay spread $T_s < \sigma_\tau$ or the signal bandwidth is larger than the channel bandwidth $B_s > B_c$. Frequency selective fading can be solved by two methods. the first approach is for narrowband bandwidth transmissions if frequency selective fading occurs at the transmission frequency, then the entire signal will be lost unless two methods are applied. The first one is by transmitting a wide bandwidth signal or signal spectrum as CDMA where any dips or fading only occurs in a small loss of the signal power instead of a complete loss. The second approach is to separate the transmission into small bandwidth carriers or chops such as in OFDM transmission. In both cases, the loss is small where this information can be recovered.

1.5.4.3 Fast Fading

Fast fading occurs when the channel impulse response changes rapidly within symbol duration or the channel variation is faster than the baseband signal variation. Fast fading occurs if the signal width pulse or symbol period is larger than the coherence time of the channel $T_s > T_c$ or the signal bandwidth is less than the channel bandwidth $B_s << B_c$. Moreover, fast fading occurs when the Doppler spread is high [6].

1.5.4.4 Slow Fading

Slow fading is happening when the channel impulse response changes rapidly within symbol duration or the channel variation is smaller than the baseband signal variation or if the Doppler spread is low. Slow fading appears if the coherence time is large relative to the symbol period delay required by that application or the standard $T_s << T_c$ or the signal bandwidth is larger than the channel bandwidth $B_s >> B_c$.

1.5.5 Distortion

There are two kinds of distortions, time-dispersion (inter-symbol interference) or frequency-dispersion that the wireless channel can get affected by as shown in the below subsections.

1.5.5.1 Time Dispersion

When a signal is propagated through the medium, the signal experiences a distortion affected by reflections, scattered and diffracted by objects where the signal arrived at different times called time dispersion. The different time between the first and last multipath components called delay spread. Time dispersion can cause inter-symbol interference ISI. This dispersion is due to the signal bandwidth larger than the coherence bandwidth while frequency selective fading is due to time spread in the time domain. Usually, the sideband signals mitigate the dispersion than the narrowband [11] [12]. The following paragraph is multipath channel parameters obtained form a power delay profile.

<u>Mean Excess Delay and RMS Delay Spread</u> The mean excess delay is the first moment of the power delay profile. RMS delay spread [6] can be obtained by taking the square root of the second central moment of the PDP. Usually, RMS delay spread values are in order of microseconds for outdoor environments and nanoseconds in the indoor radio channel. Once the RMS delay spread obtain from PDP's, the maximum RMS delay spread, the average RMS delay spread, and the standard deviation RMS delay spread costs can be achieved.

Maximum Excess Delay (X dB) Defined as the time delay value after which the multipath energy falls to X dB below the maximum multipath energy. It is also called excess delay spread.

1.5.5.2 Frequency Dispersion

Frequency dispersion is due to mobility that case time spread in the time domain. While time selective fading is due to frequency spread in the frequency domain. This has to do with the effect of receiver mobility. If the signal changes rapidly enough in comparison with the coherence time $\frac{0.18}{f_m}$, f_m is the Doppler frequency λ, then there is no distortion.

Coherence Bandwidth Coherence bandwidth is a method that is used to characterize the channel in the frequency domain. It is a statistical measure of the range of frequencies over which the channel can be considered flat. Coherence bandwidth also is the difference between frequencies over which two frequency components have a strong potential for amplitude correlation. Both Delay spread and Coherence bandwidth used to describe the time dispersion of the channel in a local area [6].

1.5.6 Building Penetration Loss Modelling

This model was proposed by the 3GPP where it consists of the following parameters which can be used for the future V2I.

$$PL = PL_b + PL_{tw} + PL_{in} + N(0, \sigma) \tag{1.24}$$

Where PL_b is path loss from the urban macro or urban micro models and PL_{tw} is the loss due to building penetration. PL_{in} is losses inside building and σ is the standard deviation.

Furthermore, these parameters can be modified base on the location and versions update of the model.

1.5.7 Shadow Fading (SF) Standard Deviation

Shadow Fading (SF) Standard Deviation can be found using a collection of measured data in the above analysis (PL) where the SF value should be obtained by subtracting the expected PL value from the measured one and then taking the standard deviation

$$SF(d_n, p) = PL(d_n, p) - \overline{PL}(d_n, p) \tag{1.25}$$

$\overline{PL}(d_n, p)$ is the expected path loss and $PL(d_n, p)$ is the measured path loss value. The standard deviation can be shown as:

$$\sigma_{SF} = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} |SF(d_n, p)|^2} \tag{1.26}$$

1.5.8 Doppler Spread

Doppler spread is just the frequency shift of a narrowband channel due to the mobility or the surrounding environments. Doppler spread can be derived as a function of effective mobility through different environments [13]. It's a good indicator of the time selectivity (RMS Doppler spread) and is related to the coherence time. According to [13] doppler spread will be also studied where its effect due to the higher mobility of the vehicle nodes which cause doppler shift (DS) of the received node in the carrier frequency. Mitigating the doppler effect will improve the throughput and the quality of service (QoS) as well. Also, the effect of a doppler frequency shift on the carrier frequencies is happening in different modulations with the same percentage and the SNR is degradation on the OFDM system due to the doppler shift and other reasons such as frequency offset and phase noise.

In my opinion, the DS will not have that effect on the ISI since the spacing between neighbor's frequencies is high on the mmWave. DS is naturally higher at the mmWave than the microwave since the wavelength is shorter which makes it harder to be estimated. DS leads to frequency dispersion and time selective fading where there are two possible fading which are fast fading (small scale fading)and slow fading. DS can be obtained using power angular spectrum plus other parameters such as speeds and angle of both Tx and Rx [14].

Doppler spread gets worse with increasing the speed or the frequency base one equation $f_{DS} = \frac{v}{\lambda} cos\theta$ where v is the speed of the vehicle, λ wavelength and θ is the direction of the vehicle to the target. If $\theta = 0$, the ray is pointed opposite to the vehicle's motion. An example can be provided to investigate these effects. For example, a vehicle is with a speed of 140 Km/hr with a frequency of 1 GHz, the maximum frequency doppler would be 129.6 Hz which is small (If the angle is zero or π) which is the max. Moreover, with the same example but the frequency is 80 GHz, the Doppler spread is 10.37 kHz which is still not quite significant if it's divided by the carrier frequency f_{DS}/f_c is 1.3e-10.

1.6 Path Loss Models

During the past decade, researchers have conducted a lot of designs to propose wireless communication systems toward improving wireless performance. Most of the published path loss model follows the general path loss model follow the relationship between path loss and transmitted power is showing as:

$$PL[dB] = P_t[dBm] - P_r[dBm] + G_t[dB] + G_r[dB] \tag{1.27}$$

Where: - PL refers to path loss. - P_t refers to the transmitted power. - P_r refers to the received power. - G_t and G_r refer to the transmit and receive antenna gain, respectively.

Empirical path loss models will be derived from the data collection and then usually be compared with theoretical models such as free space and two-ray model. There are a couple of requirements that must be accomplished for new channel models such as that channel model must support multi-frequency band up to 100 GHz and has to support large channel bandwidth (2GHz). Moreover, the new channel models supposed to support a range of antennas arrays such as cylindrical, linear, planar, and spherical arrays, with arbitrary polarization. The level of complexity should be limited. The new channel model has to be suitable for channel states such as LOS, NLOS, indoor, outdoor and other environments. In this section, there will be a couple of recent path models candidates as the alpha-beta-gamma (ABG) model, the close-in (CI) model and Close-in with frequency-dependent exponent from previous empirical communications.

1.6.1 Alpha-Beta-Gamma (ABG) Model

ABG model is the current 3GPP 3D model and values may change based on the base station location. Academic and industrial are trying to come up with a single path loss model for each scenario.

$$PL^{ABG}(f,d)[dB] = 10\alpha log_{10} + \beta + 10\gamma log_{10}(\frac{f}{1GHz}) + X_{\sigma}^{ABG} \tag{1.28}$$

ABG model is used to measure how PL gets increased with distance and β is the optimized offset in dB. Also, γ is the PL variation over frequency in GHz and X_{σ}^{ABG} is the fading (SF) in dB. Since there are three parameters, 3GPP mentioned that the ABG PL model always has a lower shadow fading standard deviation than other PL models [15].

1.6.2 Close-in (CI) Model

This model is usually for LOS and NLOS for all UMi, UMa, and InH as well by using close-in reference distance base on Fariis' law [16]. The general version of the close-in model is:

$$PL^{CI}(f,d)[dB] = FSPL(f,1,m) + 10nlog_{10}(\frac{d}{1m}) + X_{\sigma}^{Cl}, \tag{1.29}$$

The above equation is a path loss model where the parameter $FSPL$ is the free space model in dB, α is the path loss exponent (PLE) to show how PL varies with multipath propagation distance and d_o is the reference distance which set to 1 m since there is no shadowing in the first meter and simplifies the equation as well. The FSPL can be obtained using $FSPL(f,1,m)[dB] = 20log_{10}(\frac{4\pi * f * 10^9}{c})$. Where FSPL at 1 m and c is the speed of light. Close-in path loss model is more robust and provides better performance than the ABG path loss model [17].

1.6.3 Close-in with Frequency Dependent Exponent

This model was proposed in the 3GPP meetings where it depends on how high is the frequency bands. Close-in with frequency dependent is an extension of the CI model with frequency dependent [18]. The general model of the CIF model shown as:

$$PL^{CIF}(f,d) = FSPL(f,1,m) + 10n(1 + b(\frac{f - f_o}{f_o})) \\ log_{10}(\frac{d}{1m}) + X_{\sigma}^{CIF} \tag{1.30}$$

where b is a parameter that captures the slop and it would be positive if the both PL and f increase. f_o is the reference frequency [19].

[20] compared the above path loss model for both CmWave and MmWave in different scenarios such as UMa and UMi. All of them showed a good prediction with large data

where the CI model was the most suitable for outdoor cases due to the close-in free space reference. Whereas, CIF has a better performance for the indoor environments due to its small stander deviation values.

1.7 Conclusion

Channel modeling is fundamental to design wireless communication systems. A common practice is to conduct a tremendous amount of channel measurement data and then to derive appropriate channel models using statistical methods. For mobile communications, channel estimation on top of the channel modeling enables high bandwidth physical layer transmission in state-of-the-art mobile communications. This chapter covered an overview of the wireless channel modeling in wireless communications systems. Multiple of popular path loss model in MmWave models were presented. For the coming 5G and diverse Internet of Things, many challenging application scenarios emerge and more efficient methodology for channel modeling and channel estimation are required.

Chapter 2: Artificial Intelligence Techniques

The purpose of this chapter is to cover the theoretical background of artificial intelligence (AI). This section focus on AI methods that can be further involved in predicting the wireless channel parameters based on data driven and elaborate on how to overcome some wireless issues in the new era 5G.

Table 2.1: An Overview of Machine learning Algorithms in Wireless Systems

Category	**Tasks**	**Algorithms**
Supervised learning	Regression, Classification, Bayesian Inference	Neural Networks, Decision Tree. Logistic Regression, K-nearest, SVM, Gaussian Regression
Unsupervised learning	Clustering, Feature Extraction, Dimension Reduction	K-means, Spectrum Clustering, PCA.
Reinforcement learning	Policy learning	Q-learning.

Table 2.1, shows that artificial intelligence can be divided into three main categories that can be applied to the wireless channel modeling in general. These categories are supervised, unsupervised and reinforcement learning. Where some algorithms and tasks are provided in this table.

2.1 Supervised Learning:

For the purpose of learning and predicting, common methodologies to realize ML are training, testing, and validating wherein the training part, the model is trained from the data to predict channel modeling parameters [21]. Supervised learning techniques require input, target labels in a dataset to create a model that is used for estimation. If we have a sample space consist of X_i and output label space y_i where $i \in 1, 2, .., N$ then by using a machine learning algorithm $\mathring{A}$, which is a function that map the input values to the labels to preform future predicting. To measure the quality of the mapping, a loss function is used [22]. The supervised learning can divide into two main subjects which are the regression model, the classification learning. The regression method is considered continuous values whereas the classification is a discrete value. Both of these two subjects are useful in estimating the channel parameters such as the path loss component n and the large-scale random variable ψ_σ. Also, it can be used to estimate the throughput. The error can minimize if the data is sufficient in supervised learning. The following subsections are regression and classification and other algorithms that can be applied to regression and classification techniques.

2.1.1 Regression Algorithms

Regression is a statistical method that is applied to machine learning under the supervise category. The regression models are a learning mechanism based on a dataset from prior measurements or simulations. After finalizing the learning processes, the models can be used for predicting other parameters [23]. Regression uses least squared estimation (LSE) to minimize the square of the error between the observed responses in the dataset and to predict the most accurate model [21] [24]. LSE can be divided into two classes: linear (ordinary) least squares and nonlinear least squares where both of them don't require any prior knowledge. The following sub-subsections are major regression approaches that would be applied later to our analysis.

2.1.1.1 Linear Regression

Linear regression presents the relationship between the dependent and the independent variables and used forecast unknown tasks with a minimum percentage of error based on a given dataset.

So the predicted linear hypothesis function as shown below where θ_0 and θ_1 are the weighted parameters of the prediction $\hat{y}$. From the training dataset, we have independent variables x_i and dependent variables as a label output y_i as an output.

$$h(x) = \theta_0 + \theta_1 x + \epsilon \tag{2.1}$$

The cost function (J) is measured based on the mean squared error (MSE) which is used to measure the loss between the actual and the predicted value to obtain the best fit regression line. The independent variables are supposed to be updated until the model converges and the equation of the cost function can be shown as:

$$\arg\min_j J = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2 \tag{2.2}$$

Above equation is the MSE cost function Where m is the number of training data and $h_\theta(x_i)$ is the predicted $\hat{y}_i$. MSE cost function is used to measure the performance of the training model. MSE used to measure the average amount of the prediction from the actual values. By optimizing the θ_i, the cost function can be reduced to improve the model's prediction.

$$\arg\min_j \sum_{i=1}^{n} J_i^2 \tag{2.3}$$

Optimizing the thetas can be done by using the gradient descent to enhance the model and reduces the cost. By using gradient descent minimization to find the optimum value of the coefficient, the cost value will be reduced.

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(y_i - h_\theta(x_i))^2 \tag{2.4}$$

where the prediction function $h_\theta(x_i)$ can been as:

$$h_\theta(x_i) = \hat{\theta}_0 + \hat{\theta}_i x_i \tag{2.5}$$

By applying the gradient descent to minimize the value, can be done by taking the derivation of the function with respect to each one of the coefficient.

$$\frac{\partial}{\partial\hat{\theta}} J(\hat{\theta}_0, \hat{\theta}_1) = \frac{\partial}{\partial\hat{\theta}_0}\left(\frac{1}{m}\sum_{i=1}^{m}(y_i - h_\theta(x_i))^2\right) \tag{2.6}$$

$$\frac{1}{m}\sum_{i=1}^{m}\frac{\partial}{\partial\hat{\theta}_0}(y_i - h_\theta(x_i))^2 \tag{2.7}$$

Differentiation with respect to θ_0, we get:

$$\frac{2}{m}\sum_{i=1}^{m}(y_i - h_\theta(x_i))\frac{\partial}{\partial\theta_0}(y_i - h_\theta(x_i)) \tag{2.8}$$

$$\frac{-2}{m}\sum_{i=1}^{m}(y_i - h_\theta(x_i)) \tag{2.9}$$

$$\frac{-2}{m}\sum_{i=1}^{m}(y_i - \hat{\theta}_0 - \hat{\theta}_i x_i) \tag{2.10}$$

Now, by making the above equation equal to zero and divide by n to remove the summation, we get after some algebra:

$$\hat{\theta}_0 = \overline{y_i} - \hat{\theta}_i\overline{x_i} \tag{2.11}$$

Where $\overline{y_i}$ is the sample output and $\overline{x_i}$ the sample mean of the input data. Same procedure can be applied to obtain the second coefficient θ_1 and end up as formed as:

$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x_i}^2} \tag{2.12}$$

Applying the gradient descent to update the independent parameters to reduce the J by starting with a random values for these parameters and keep updating until the J converge. The updated coefficient can be compute using the learning rate α as shown below:

$$\hat{\theta}_i = \hat{\theta}_i - \alpha \frac{\partial}{\partial \hat{\theta}_i} \tag{2.13}$$

2.1.1.2 Multiple Regression

Multiple linear regression is a supervised learning and the goal is to infer and predict a parameter by reducing the error using the training data to predict the target by using machine learning method that performs a better estimation $\hat{f}(x_i)$. In our scenario, using dataset that divided into two part, training and testing where the training set consists of 70% and the test is 30% and other as mentioned in [25]. The multiple linear equation can be seen as follow [26]:

$$h_i(x) = \theta_0 + \theta_1 X_1 + ... + \theta_i X_i + \varepsilon \tag{2.14}$$

The estimated of the response variable Y (PL) using X as CSI and minimizing the erroneous follow:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \hat{X}_1 + ... + \hat{\beta}_i X_i \tag{2.15}$$

Then, to maximize the estimation method accuracy, the coefficient of equation 2.2 has to be minimized to obtain a lower difference between the real and the estimation equation. The residual error of the regression estimation can be obtained using the below equations [2].

$$Minimize \sum_{i=1}^{n} e_i^2 \tag{2.16}$$

$$e_i = \arg min \sum_{i=1}^{n} (y_i - \hat{y_i})^2 \tag{2.17}$$

$$(\beta_0, ..., \beta_i) = \arg min_{\beta_0 ... \beta_1} [\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 X_1 - ... - \hat{\beta}_i X_i)^2] \tag{2.18}$$

Then, estimating the of the slope and variance can be shown below [27].

$$\hat{\beta}_1 = (X^T X)^{-1} X^T y \tag{2.19}$$

$$\hat{\sigma^2} = \frac{1}{L-1} (y - X\hat{\alpha})^T (y - X\hat{\alpha}) \tag{2.20}$$

2.1.1.3 Support Vector Machine (SVM)

Support vector machine (SVM) that is derived from learning theory is a method to analyze data for classification analysis. SVM was invented by Hava Siegelmann and Vladimir Vapnik (2001) [28], to identifying datasets using separating hyperplane. Linear regression is suitable for a complex problem and it was extended to SVM and Linear regression is the fundamental of the SVM. SVR is a regression method that can be used toward the wireless channel modeling and it's different from SVM. The output of the algorithm supposes to be continuous values instead of classification which is categorical. The function that will be used to describe this method is called karnel and used to map a lower-dimensional data into a higher dimensional data. Its well known that SVRusese a hyperplane to predict the response instead of using it as a separation line in the SVM to distinguish the classes. SVR is different from the regular linear regression by fitting the error within a certain threshold instead of minimizing the error using least square error. Furthermore, the fitting line in the SVR is base on the maximum data points within certain boundaries.

$$wx + b = 0 \tag{2.21}$$

where ϵ is the error of data point to each boundary that is shown as

$$-\epsilon \leq y - wx - b \leq \epsilon \tag{2.22}$$

2.1.2 Classification

The general goal of classification is to map a categorical label to each data sample in the provided dataset [29]. Where each observation identified to sub-populations (Categories) label. Multiple classification methods are used to classify in terms of binary and multi-binary such as Naive Bayes Classifier, Nearest Neighbors, Support Vector Machines, Decision Trees, Boosted Trees, Random Forest, Neural Networks and others that can be used for both regression and classifications.

2.1.3 Deep Neural Networks (DNNs)

Among many DNN structures, Multilayer Perceptrons (MLP) uses a Feed-forward neural network (FFNNs) and a back-propagation network to compute the loss and adjust the weight [30], which is suitable for deep learning. MLP forms fully connected networks where every single node in a single layer is connected to every node in the following layers. The neurons usually learn how to transform and convert the input data into a corresponding output. The subsequent error is usually obtained by the loss function and optimization methods can be used to minimize the loss such as Adam optimizer. There are multiple loss functions and cross-entropy will be used when the methodology viewed as a classification problem or least square for the regression approach. MLP is also a multivariate multiple nonlinear regression and collection of neurons that serve as a regression by building a decision. Multilayer Perceptrons are typically uncorrelated, and a collection of them make up the network that can be less prone to the notorious overfitting.

The neural network is a data driven approach that learns from the data. Where there are testing data and training data, so the system learns the hidden structure and makes a decision then that decision can be checked with cost/loss/error function. DNNs are the mapping of the input to the output which is a branch of deep learning algorithms. The deep neural network is inspired by the human brain and can be implemented by mapping the input data to the output in terms of the diagram where each vertical group of nodes is called a layer. The first layer usually from the input data and every node is connected to every node in the following layer, the last layer produces the predicted or the response value and other layers between are called hidden layers. Furthermore, the more layer added refers to the depth in the network [31]. Later, the weights are adjusted to obtain the least loss. Once the new data pass to the input layer, the network tries to match it to what it has been learning in previous training.

To accomplish our purpose, Adam optimization algorithm will be adopted which is different from the traditional stochastic gradient descent process to update the weights iterative base in the training data [32] with a learning rate or step size α to adjust weights.

Artificial intelligence (AI), particularly machine learning (ML), is widely studied to enable a system to learn of intelligence, predict and make an assessment instead of the needs of humans [33]. We are in a world that users are engaging with a software or a model that can learn more from the interaction where the neural network is one of the ways that used to learn over time Deep Neural Networks (DNNs) is a set of algorithms and one of the techniques of artificial neural networks (ANNs) [34]. Neural networks used to map the input to the output. In [35], the authors recommend involving the convolution neural network (CNN) to the 3GPP channel model. In [36], the CNN techniques used to predict the speed and direction angle using a sequence of images as an input. By using a dataset that is produced from channel modeling simulation or measurement, the deep neural network is inspired and emulated by the human brain and can be implemented by mapping the input data to the output.

Deep neural methods are an ideal technique that is capable of solving or reduce the complex problems that wireless channel modeling is facing these due to complexity and time consuming that used to collect and build channel models for different environments. Moreover, by applying DL, wireless network applications can reach higher estimation precision and decrease design complexity.

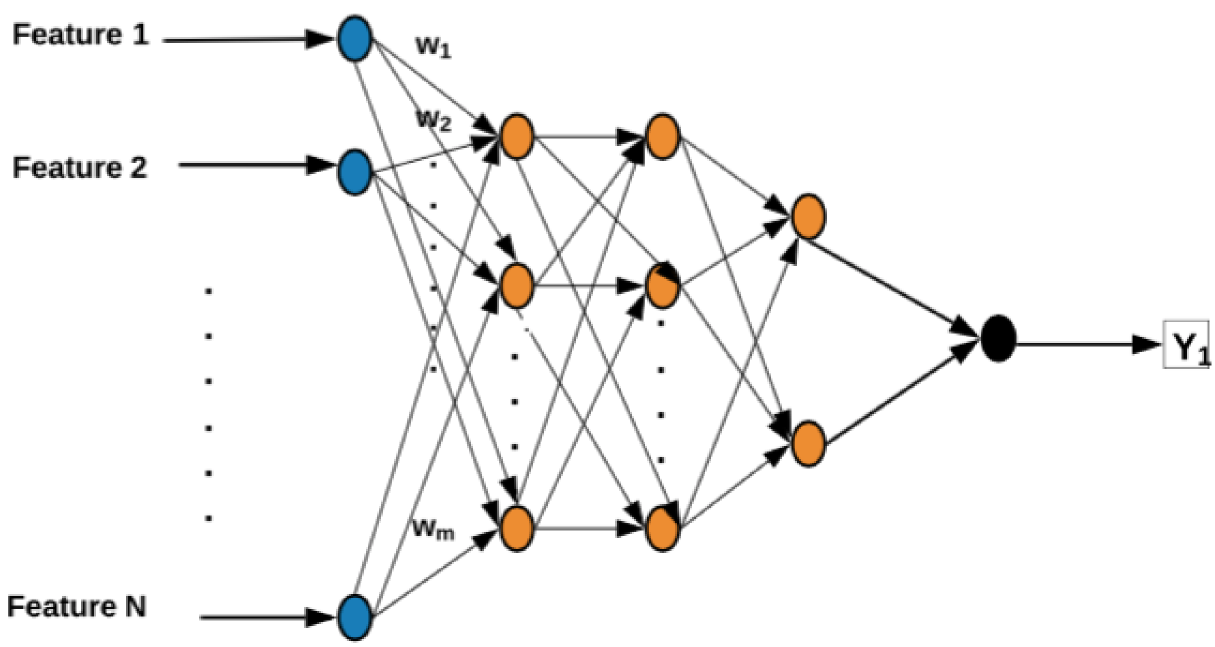

Figure 2.1: ANN Structure

Neural figure 2.1 there are multiple of layers that can be explained as follows:

- Input Layer: The purpose of this layer is to take the data features in terms of numbers or figures, etc and pass it to the hidden layers. There is no computation occurred in this stage.

- Hidden Layers: which are the middle layers that are responsible for processing the data and performing the operations and transfer them to the output layer.

- Output Layer: To generate the decision of the data processing.

Multilayer Perceptrons has the form of some products of Input X and their corresponding weights w plus some error b and apply an activation function $f(x)$ to obtain the output of that layer then feeds it to the input of the following layer. MLP follows the form as shown below.

$$Y_i = f\Big(\sum_i w_i x_i + b\Big) \tag{2.23}$$

Where w is the vector of weights of x vector inputs and b is the error. While the function $f(.)$ is the nonlinear activation function to compute the neuron's output. The activation function is essential to ANN to learn and convert an input signal of a node to an output signal that feeds to the following layer and introduce the non-linearity to learn complex tasks in the neural networks. Furthermore, if an activation function is not involved in the networks, it becomes a linear regression (Polynomial of one degree). There are a variety of activation functions that can be used in the neural networks depending on the task performing as shown:

- **Logistic Sigmoid** mainly used to predict a probability as an output and it lies between zero and one which is preferred in the binary classification.

- **Tanh** activation function is similar to the sigmoid while its output lies between -1 and 1. Tanh function can be in the form of tanh(x) = 2 * sigmoid(2x) - 1 which is a mathematically shifted from the sigmoid function. Tanh is useful with reducing the structure of the network due to some values would be zero due to its range.

- **Hyperbolic Tangent** is a common activation used the backpropagation that evaluated from -2 to 2.

- **Rectified Linear Units (ReLUs)** which is used in the modern neural due to the speedway of processing. With large weight updates, ReLUs result in inactive nodes and become a deadly ReLU since any negative values consider a zero in the output of that node which causes this node to become dead and blocked. The extension of this version is called Leaky ReLU (LReLU) to avoid this disadvantage.

- **Step Function** is also called the threshold function where mostly used in classification methods. The step function passes zero and one to pass the process to the following layer or deny the process. It's helpful for a network that has many layers and neurons to reduce the size, time and complexity of the network.

Activation functions are also responsible for deciding whether a neuron should be active or not base on the summation of the weight value with the input and the bias factor. The raw data of the received and transmitted signals are collected as training data and the model is trained and then the loss L is obtained.

$$\arg\min_{j} L = \frac{1}{N} \sum (\hat{Y}(k) - Y(k))^2 \tag{2.24}$$

To minimize the error between the estimated and the received transmitted signal, the sum of square residual error of sample i is used. Having all neurons running with a fully connected network leads to a complex network structure especially with the growth of the layers and will result in a delay in training the data due to speed limit of computer processing and the large data. Thus, regular forward feed networks have fallen into the disuse due to these limitations and it's the time to apply the backpropagation. Where the weights are tuned and the error is backpropagated through the network couple of times to reduce the cost function.

$$\arg\min_{j} e_i = \sum_{i=1}^{n} [y_i - \hat{y_i}] \tag{2.25}$$

Multilayer Perceptrons Neural Network usually can be used in both classification and regression. When the Multilayer Perceptrons used for the classification, this algorithm works by having binary or multiple classes. However, the regression techniques is usually used for continuous outputs.

2.2 Unsupervised Learning

Unsupervised learning is a branch of machine learning that is mainly used for learning the pattern of the data where it has an input with the unlabelled response. Usually, it is hard to have a large amount of labeled data which causes people to switch to use the unlabelled data that is known as unsupervised learning. Using machine learning techniques such as unsupervised learning to reduce the complexity, improve the measuring performance such as in distributing the wireless cellular tower. Unsupervised learning is useful in the case when the data is not efficient where the hidden variables can be obtained using methods such as Bayesian learning. Clustering is one of the main methods of unsupervised learning wherein it is used in channel modeling to gather multipath components (MPCs) with similar parameters behavior to improve the process of the time-variant channel model precision. These MPCs has similar parameters such as the angle of arrival (EOA), angle of departure (EOD), delay (τ), the azimuth angle of arrival (AOA), the azimuth angle of departure (AOD), and power level will be in in the same cluster [37]. MPCs can be performed in the cluster phenomenon where power delay profiles (PDPs) that have similarities can be presented in one cluster and other MPCs with different clusters. Moreover, unsupervised learning employs the clustering method in the wireless channel modeling where each pattern represents a single clustering such as identifying either the LoS and NLoS.

2.2.0.1 Gaussian Mixture Model (GMM)

The Gaussian mixture model (GMM) is a method of clustering techniques where it's used to implement channel multipath clustering. GMM gives the option to mix the relation of data to clusters. By comparing k-means which is a clustering algorithm to GMM in data distributed, in K-means there is no mix where a single input only belongs to one cluster (Hard clustering) whereas in the GMM an input can belong to different cluster at the same time with a different degree of similarity (Soft clustering). GMM is more flexible then k-means in terms of reducing complexity. Using such simple principal component analysis (PCA) is an example of the big data algorithms to model a channel modeling where the PCA uses an orthogonal transformation to a correlated dataset of an uncorrelated dataset [38]. Moreover, the MPCs are assumed to consist of several Gaussian distributions in the GMM technique that can be applied to channel multipath clustering and the likelihood of this distribution is

$$L(X;\Theta) = \sum_{i=1}^{N} log \sum_{k=1}^{K} (x_i|\mu_k, \Sigma_k)\pi_k, \tag{2.26}$$

The theta parameter Θ is the assembly of parameters such as μ_k, Σ_k and π_k. Where π_k is the prior probability of a class. x_i is the channel multipath, μ_k is the mean and Σ_k is the chi-variance matrix and π_k . Then, parameters of the GMM methods are estimated to solve the likelihood function by the expectation maximization (EM) algorithm.

$$\hat{\Theta}_{ML} = \arg\max_{\Theta} L(X;\Theta), \tag{2.27}$$

2.2.0.2 *K-Means Algorithm*

[22] used the K-means algorithm to identify the clusters and then this algorithm iteratively groups cluster the MPC to minimize the Euclidean distance between the data until it becomes a cluster and then keep iteration to reach the convergence. This algorithm has been used in [8] to cluster the multipath. During the channel modeling, if there are L number of clusters and each cluster consists of $X_1, X_2, ...X_n$, of multipath. By using the centroids $C_1, C_2, ..., C_L$, which is the cluster center in random locations in space and a point in high dimensional vector space. Then, by using the measured data set and for each multipath X_i find the nearest centroids C_i and assign each multipath X_i to every cluster j. Then by using the below equation, Euclidian distance D_E can be obtained to cluster the base of the transmitted signal on their behavior.

$$\arg\min_j D_E(X_i, C_j) \tag{2.28}$$

Where the $D_E = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$ for continuous features and for the categorical variables, Hamming distance is useful as $D_H = \sum_{i=1}^{k}|x_i - y_i|$.

The second step of the algorithm is recomputing each cluster centroid position j=1,2,..., L and that can be done by taking all data inside that cluster and averaging them to obtain the new centroid j. Then, these steps kept iterations to readjust the centroid location until it converged.

2.3 Reinforcement Learning

Another branch of Machine Learning, known as reinforcement learning (RL), is to learn from experience when the prior information is unknown or to interact with dynamic environments. RL is used to learn a system to solve complex problems such as Markov Decision Process (MDP) which usually called model-free based method or model-free methods such as Q-learning that does not learn a system [39]. RL starts with learning, updating and then pursuing the action to maximize a reward. Markov Decision Process (MDP) is the foundation of the RL where the goal of an MDP is to find an optimal policy to maximize the decision [40]. Markov Decision Processes are widely used as a methodology for modeling sequential decision-making in environments with probabilistic dynamics or dynamics of uncertainty.

In the channel modeling of the wireless sensor networks (WSNs) consist of resource-limited devices such as power transmitting. Static decision commands may lead to inefficient energy usage. For example, a node sending data at fixed transmit power without considering the channel conditions will drain its energy faster. Thus, by using MDPs will help to optimize the problem. The MDP model allows a stable design of dissimilar objectives, such as minimizing energy consumption and maximizing sensing coverage. [41] proposed a risk-sensitive reinforcement learning to optimize the beamwidth and the transmit power of cells. The optimal policy can be obtained after having the optimal value function to select the action that maximizes the value in the next state. Reinforcement learning is all about trying to understand the optimal way of making decisions/actions that maximize reward R such as Q-learning. RL can be applied to issues that have been a while that the wireless system dealing with such as reinforce the traffic flow and handover where an action the machine has to do. Applying RL in channel modeling will lead to better performance such as power control and handover [42].

$$\pi^*(x_k) = argmax_{a_k}\{E[r_{k+1}|x_k, a_k] + \beta \sum_{x_k+1} p(x_{k+1} \mid x_k, a_k)V^*(x_{k+1}) \quad (2.29)$$

A useful algorithm that helps to find the optimal value function is called the value iteration algorithm.

$$max_{x \in S}|V^{l+1}(x) - V^{l}(x)| < \delta \tag{2.30}$$

Where l is an iteration counter. The above equation can be used to compute optimal policy using value iteration which is the most effective and widely used technique to resolve infinite time horizon discounted MDP. This method has many advantages such as quick convergences especially if the state space is too large and easy to implement. One of the Algorithms that represent RL is using Iteration which usually repeats the optimization until the optimal state value function V(x) converges. The converges occurs once the maximum value difference between two iterations is less than a specific threshold. For the policy iteration: we keep repeating improving until it converges.

Machine learning techniques have been a major method for the problems that deal with routing such as in adaptive routing, shortest path, and multicasting routing. One of the ML techniques to solve these issues. Q-learning is one of the reinforcements learning that is used to control the routing follow [43]. Usually, Q-learning is a model-free method that doesn't learn a system and can handle the routing by using stochastic transitions and rewards. Q-learning calculates the policy that maximizes the expected value to maximize the overall reward. Q-function is used to solve the MDP problem in cases when the information of the environment isn't needed. [44] mentioned that Q function estimates the expected reward when taking an action in a state such as pursuing a handover and power control problems. RL can be valuable in blockage and selecting beam issues where there will be a reward if the beam of the signal is selected successful and if there is a blockage, a negative reward is incurred and reelecting the link base on the training mode is required. To the best of our knowledge, there are a rear works that have been done on applying reinforcement learning on wireless communication channel modeling [40]. Deep reinforcement learning (DRL) as mentioned in [1] can be used for cognitive radio network (CRN) control to sense the free channel.

2.4 Massive Wireless Channel Modeling Data Analytics

The channel modeling dataset involves a massive amount of data. Effective machine learning and feature extraction can alleviate the thrust from the size of data sets, and the complexity with comparable functionality can be substantially reduced. Authors of [42] explore complexity reduction techniques to improve signal efficiency. The first technique is using distributed optimization algorithms, for instance, alternating direction method of multipliers (ADMM) and primal/dual decomposition. Both of these methods are used to decouple the statistical learning issues into small sub-issues. The second technique is used for incomplete data and to perform that, online or active learning is needed.

Other types of classification reduction are dimension reduction which is used to reduce a high dimensional space without losing any useful information. Methods such that are liner projection and nonlinear projections methods. An example of the linear projection methods is principal component analysis (PCA) and for the nonlinear methods are isometric mapping (ISOMAP) and locally linear embedding. [45] went over an interdisciplinary study of using massive channel modeling data and used a cluster-nuclei based channel model and matching the propagation signals to machine learning algorithms using measurement data.

Data Mining is the process of discovering large dataset that is used to learn rules and relationships that help to build a model to predict automatically from the updated data. Data mining can be obtained from different many resources where this dissertation focuses on wireless channel modeling which is becoming a Data Mining. Data is useful for scientists and during this decade, wireless communication devices have had grown by a factor of one thousand [46] [12]. Since the measurement of channel modeling will produce big data with a large of Gigabits especially with cellular systems. MIMO is one of the major things that switch the channel modeling to be data mining. [47] recommended using the benefits of data mine to model wireless channel modeling.

Usually, the parameters of a channel model are generated by measurements or simulations of the propagated signal through a channel where that signal gets disturbed by fading, blocking and distance which leads to MPCs. The highest MPCs is the strongest link and from there, channel parameters can be obtained to create a dataset. During the previous wireless generation where SISO with lower bandwidth, the amount of data is not sufficient enough to be used for accurate prediction. Whereas with the usage of MIMO and higher bandwidth, data can be used to estimate instead of taking so much measurement and leads to support the accuracy of wireless communication modeling. Thus, channel modeling can be considered as a sort of data mining.

The data is needed for establishing a channel model where the features and structures are obtained for the empirical data. Generally, wireless communication including cellular and others can be data mining since it's geographically distributed, producing data continuously with updated versions of the same parameters [48]. Data mining methods used in Machine learning as other algorithms to build models that predict our needs Authors of [49] believe that the combination of data mining and machine learning methods and huge dataset, channel modeling parameters can be extracted and controlled. [50] has confirmed that a deep model reduces the channel modeling complexity by allowing DNN algorithms to learn complex functions. Moreover, some techniques can be applied to the data analytics unit such as confirmatory data analysis (CDA) which uses statistical techniques to find out whether hypotheses of the data set are true or false. Also, an exploratory data analysis (EDA) technique can be applied to the data analytics to obtain the relationships and patterns in the given data [51].

Chapter 3: Overview of the ML Toward Wireless Channel Modeling

Channel modeling is fundamental to design wireless communication systems. A common practice is to conduct a tremendous amount of channel measurement data and then to derive appropriate channel models using statistical methods. For highly mobile communications, channel estimation on top of the channel modeling enables high bandwidth physical layer transmission in state-of-the-art mobile communications. For the coming 5G and diverse Internet of Things (IoT), many challenging application scenarios emerge and more efficient methodologies are needed for channel estimation. In the meantime, machine learning has been initiated successfully demonstrated efficiently handling big data. In this chapter, applying machine learning to assist channel modeling and channel estimation will be introduced in this chapter and results with evidence will be in the following chapters.

Artificial intelligence (AI) methods have shown a great performance in multiple domains such as image recognition and with this motivation, we decided to apply AI to solve complex issues in the wireless channel modeling. Machine learning has been successfully demonstrated efficient handling of big data. In this chapter, applying machine learning to assist channel modeling and channel estimation will be introduced. Moreover, an overview of artificial intelligence (AI) and its applications towards wireless channel modeling will be introduced as well. Additionally, previous works from the literature exploring AI in the field of wireless channel modeling.

Machine learning applied to wireless communication channel modeling is expected to improve the derivations of channel models and contribute to more effective channel estima-

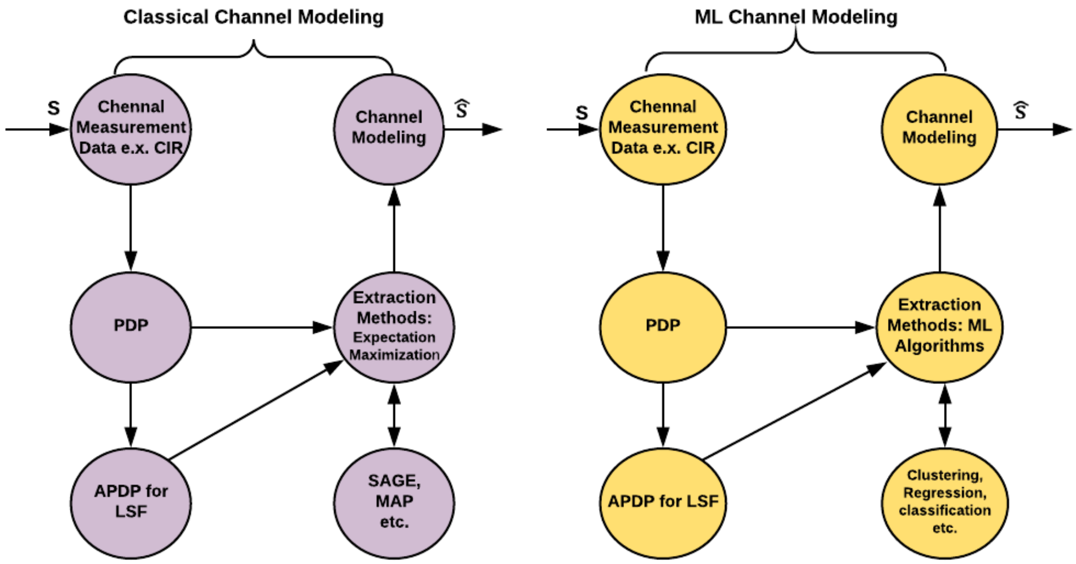

Figure 3.1: A General Visualization Explaining the Difference Between Traditional and New Wireless Channel Modeling

tion. Consequently, machine learning for wireless channel modeling is not only best fitting a model with appropriate parameters but also keeping the model updated to obtain more precise operating of wireless communication systems.

Figure 3.1, demonstrate a general compression of the classical wireless communication channel modeling and the ML wireless channel modeling. Both of them start with have CIR that is obtained from the sounders and then power delay profile (PDP) can be determined by taking the absolute square of the CIR to determine the small scale parameters. If the target is to estimate large scale fading (LSF), average the power delay profile (APDP) is calculated [52]. Then, using the classical estimation method such as SAGE to extract the channel parameters. On the other hand, involving and applying machine learning to extract the CSI and then applying estimation techniques using a method such as regression, classification, and clustering, etc. to estimate the channel parameters that will lead to better performance and lower complexity. Channel estimation is a mathematical predicting of the natural propagation of the signal that helps the receiver to approximate the affected signals. [1] suggests using deep learning to identify and classify the modulation nodes, improving the interference alignment and locate the optimum routing path. [53] addressed that DNNs are capable to learn the channel model characteristics of the wireless channel that get affected by fading. In [35], the authors recommend involving the convolution neural network (CNN) to the 3GPP channel model. In [36], the CNN techniques used to predict the speed and direction angle using a sequence of images as an input. [24] used multiple machine learning algorithms such as KNN and Random Forests to estimate the path loss model from a dataset and showed results that prove the accuracy with small estimation errors.

$$h(t,\tau) = \sum_{i=0}^{L-1} a_i(t)\delta(\tau - \tau_i(t)) \tag{3.1}$$

The time-varying impulse response of the channel is shown in the above equation and after applying the APDP to it using $APDP(\tau) = \frac{1}{N}\sum_{i=0}^{N}|h(\tau)|^2$. Then, the output of that equation is fed to the structure of the DNN. To predict the Y_i values from the training data $X_1, X_2, ..., X_N$

3.1 ML Usage in the Wireless Communication Channel Modeling

Artificial intelligence (AI), particularly machine learning (ML), is widely studied to enable a system to learn of intelligence, predict and make an assessment instead of the needs of humans. Switching the traditional channel modeling to machine learning channel modeling still in its early stage. One of the main issues in current communication is to accurately predict the channel parameters, whereas using machine learning techniques could enhance the prediction and reduce the complexity. ML can be used to predict and estimate the wireless channel parameters and examine large and small-scale fading including parameters such as path loss, delay path loss exponent, carrier phase shifts, Doppler spread and random variable that explains the large scale fading. [24] used multiple machine learning algorithms such as KNN and Random Forests to estimate the path loss model from a dataset and showed results that prove the accuracy with small estimation errors.

Estimation of these parameters is required at the receiver. In ML, there must be a different way than the previous methods such as minimum mean square error (MMSE) estimation or maximum a posterior probability (MAP). The new ML methods estimate the channel parameters using different methods base on their categories. [54] used deep learning for estimating the carrier frequency offset (CFO) and [28] showed how to predict the transmitted signals using deep learning techniques such as neural networks other than the classical method where existing receivers estimate the parameters and then recover the data using estimation. Whereas with deep learning, the channel state information (CSI) is estimated and recovering the transmitted signals directly. [55] used neural network methods such as learned denoising-based approximate message passing (LDAMP) network to estimate and learn CSI and solve the limited number of frequency chains in cellular systems from training data. [56] used a model-based method using Cramer-Rao lower bound (CRLB) to predict the CSI using the deep neural networks. During the data processing, the noise threshold has to be identified where any PDP close to the noise floor will be identified during the data processing. Specifying the noise threshold to extract the CIR is an essential thing. The traditional ways to overcome the noise from the CIR and signal higher than the threshold is considered acceptable and choosing the noise threshold is sensitive since removing all noise means getting rid of part of the signal and reducing the noise level leads to having some noise in the signal itself. Machine learning methods are capable of choosing an accurate threshold. [57] presented several novel ideas of applying deep learning to the end-to-end communication systems such as autoencoder and redesigning those communications in a single process. Where the autoencoder helps to decrease the block error rate (BLER) by 1-7 time in scenarios that has Rayleigh fading. Moreover, the Same author presented in [58] other channel modeling challenges that reach the complexity level such as blockage, atmospheric effects, handover, beam direction, MIMO and it is the time for machine learning to get evolved. [59] proposed a unique way to train deep learning of channel modeling without any assumptions which improve the system performance. Authors used a back-propagating stochastic approximation method to overcome the corrupted signals due to hardware, fading, and improper schemes, etc. The cross-entropy loss function used to measure the performance of the system as shown below. Regression is one of the main methods that is used in machine learning where the regression models are a learning mechanism based on a dataset from prior measurements or simulations. After finishing the learning processes, the models can be used for predicting other parameters. Authors of [23] applied three regression models support vector, linear and DNNs regressions then, made a compression to control the high-speed channel modeling errors. Where all of these techniques help to reduce the wireless channel modeling complexity and need to be investigated more. [60] used the regression algorithms to estimate the RSS over delay bins.

$$L(\mathbf{s}_{1:N}, \hat{\mathbf{s}}_{1:N}) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{s_i \in s_n} s_i \log(\hat{s}_i) \tag{3.2}$$

where s is the input symbol and $\hat{s}$ is the estimated output.

3.2 Applications of Machine Learning in Channel modeling

The propagation channel of communication systems fundamentals is focusing on channel modeling particularly for a new technology era such as machine learning. Channel modeling is so wide either in cellular or wireless sensor networks channel model systems. The irregularity and complexity of the wireless channel lead to solid ways to achieve accurate models where more trails are always required to accomplish the precise results mainly with new technologies. Where the primary focus of this chapter is to go over Machine learning in different channel modeling applications.

Machine learning is needed to examine the channel modeling and focusing on large scale and small-scale fading where fading refers to the time changes of the MPC's power. The large-scale fading (LSF) explains the main characteristics of the channel such as path loss, shadow, angular spread, and delay spread. Moreover, LSF cases will be examined the relationship between the path loss and the separated distance between the Tx and Rx in different environments such as the suburban. Where the small-scale fading case will go over mainly the relation between the Doppler spread and time coherence. These studies are needed due to the lack of real measurements which will allow developing reliable communication. Generally, large-scale parameters are chosen base on mobility, shadowing, delay spread and angular spread. while the small-scale parameters are measured based on power, Doppler spread, an angle of arrival and departure of the MPCs.

3.2.1 MmWave Channel Model in High-speed Applications

Investigations on the wireless channel modeling would allow other applications to interchange data to make the communication more secure by using a robustness detection scheme to avoid jamming attacks [61]. Getting machine learning involved in measurement from simulation or campaign will provide us with unbiased results. Machine learning algorithms investigate the feature of wireless channels deeply in MmWave. Moreover, these measured data will lead us to build a new statistical (Empirical) channel model to design an optimized communication with high-speed applications. By using a dataset from prior measurements and designs, ML methods would be applied to reduce the errors. Machine learning is known to improve performance and reduce complexity. [23] applied ML techniques for high-speed channel modelings such as support vector and deep neural network (DNN) regressions. Authors of [23] applied three regression models support vector, linear and DNNs regressions then, made a compression to control the high-speed channel modeling errors. Where all of these techniques help to reduce the wireless channel modeling complexity and need to be investigated more. [60] used the regression algorithms to estimate the RSS over delay bins. In the following section, regression will be applied to predict path loss.

[1] mentioned that deep learning is an ideal technique that is capable of solving or reduce complex problems. Furthermore, by applying DL, wireless network applications can reach higher estimation precision and decrease design complexity. Authors of [60] used the neural network to optimize the estimated parameters with two hidden layers. The parameters that were estimated by DNNs are received signal strength (RSS) and delay over angle of departure (AoD) bins. This method could lead the effort to increase the accuracy of the link budget due to channel fading, interference and mobility.

[53] purposed a method to train the receiver by computing the loss function and using the DNN. The input of the DNN is the received signal y and the received signal pilot y_p (for the time-variant) that leads estimate the channel without estimating the channel by itself. The transmitter and channel can be trained and estimated as shown in Hao *et al.* work. Lee at. el in [50] purposed DNNs to improve based automatic modulation classification (AMC) technique. Authors had revealed that using DNNs is powerful for channel modeling and different than the traditional channel modeling. [62] used DNN to estimate the channel of MIMO systems using the direction of arrival (DOA). Authors trained the DNN using data of different channel scenarios that confirm the better performance than the conventional methods. Deep Neural Networks (DNNs) is a hot topic of artificial neural networks (ANNs) [34].

Figure 3.2 shows a feedforward DNNs that can be applied to channel modeling to estimate the channel parameters such as delay, azimuth, and fading. Moreover, it can be used to estimate the link budget, for instance, path loss parameters. The number of neurons in each layer can be shown in a vector $L = (L_{in}, L_1, ..., L_h, ..., L_n, L_{out})$ where L_{in}, L_{h_i}, L_{out}

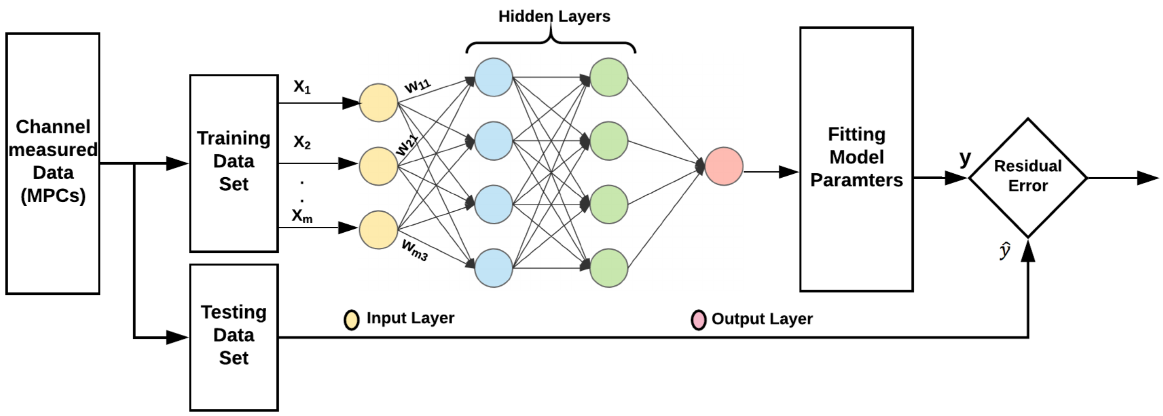

Figure 3.2: Applying Deep Neural Networks in Wireless Channel Modeling

are the number of nodes in input, nth hidden. Let (x,y) the measurement data of the wireless channel such CIR is used in the train part using the backpropagation algorithm or Newton algorithm [?] and the DNN algorithms to estimate the weight parameters to build the model channel. Then, the test data used to examine the model. lastly, the residual error is measured to obtain the accuracy of applying the DNN algorithm to the wireless channel modeling. Equation 3.3 shows the output of the DNN where w_i is the weight of each node which is needed to be estimated and n is the number of nodes. The number of the output layer is usually base on the target and deciding the number of hidden layers and number of neurons in each hidden layer is still an open research topic where few or more neurons lead to underfitting and overfitting. Authors of [28] used the DNNs algorithm to predicts the transmitted signal and then modeling the estimated data to recover the transmitted signals. Besides, Hao Ye at el.

$$Y_i = f\Big(\sum_i w_i x_i + b\Big) \tag{3.3}$$

where b is obtained by the training data and one way to determine the bias is using root mean squared error (RMSE) and by using an iteration of the neural networks, the b value can be reduced [21]. Then, after the model is trained, the loss of L is obtained.

$$\arg\min_{j} L = \frac{1}{N} \sum (\hat{Y}(k) - Y(k))^2 \tag{3.4}$$

To minimize the error between the estimated and the received transmitted signal, the sum of square residual error of sample i is used.

$$\arg\min_{j} e_i = \sum_{i=1}^{n} [y_i - \hat{y}_i] \tag{3.5}$$

The dataset of the wireless channel is supposed to divide into two fragments as in figure 3.2. The fragments are training, and testing which help to produce a perfect model. Other research use training, validation and testing [63]. The selected model will be trained and validated on the dataset then will be tested using the unseen data as showing in figure 3.2. Usually, the train part takes 60% of the data set and 25% for the validation and 15% for tasting the model [64] [65]. The training process can be repeated several times and then taking the average as [56] repeated the training process ten times. Usually, the dataset is splatted into two subsets which are training and testing the model. Training the dataset is used to train the algorithm for the DNNs model and then later the tested data use to evaluate the final model [66]. [1] recommends using deep learning for the incomplete or erroneous data to be reconstructed as if the human brain can recognize letters, numbers, faces, etc.

During the past decade, vehicle-to-vehicle communication has gained academical and industrial intentions where the vision is to have communicated vehicles for safety and other purposes. Today, there are a couple of challenges to create wireless communication and that is due to latency which will slow the interchange of data from sensors or cameras between connected vehicles and complexity. Base on the literature, the only way to overcome these issues is by using Machine learning especially with future technologies that support high data rates [67]. The usages of higher frequency will assist IoT to interchange information to create what is known as a connected IoT due to the amount of bandwidth and high data rate. By using machine learning algorithms, Efficiency and lower complexity can be achieved than the current cellular communication systems.

During the mobile communication, the Tx, Rx and the environment moving which leads to high Doppler shifts. The channel modeling in cellular is a time and frequency variant. There are a couple of cellular channel modeling such as Two Ring Model [68] [69] and Regular-Shaped GBSMs (RS) [70] where both of them are 2D GBSMs. A third cellular channel model was proposed by [71] is a 3D two-cylinder based and later was extended by the same author into a non-stationary 3D model [72]. Too many cellular channel models in the literature but since it is a moving environment, Time-Delay Line (TDL) is a non-stationary correlated scattering and modelled based on real measurement in different places [73]. 3GPP [**?**] has release outdoor-to-indoor car penetration loss and that can be modeled as

$$PL = PL_b + N(\mu, \sigma_P^2) \tag{3.6}$$

From the above equation, PL_b is the outdoor PL and that model can be applied for frequencies from 0.6-60 GHz.

Penetration Loss Modeling was proposed by the 3GPP, which consists of the following parameters which can be possibly used for the future V2I.

$$PL = PL_b + PL_{tw} + PL_{in} + N(0, \sigma) \tag{3.7}$$

where PL_b is path loss from the urban macro or urban micro models and PL_{tw} is the loss due to building penetration. PL_{in} is losses inside building and σ is the standard deviation. Furthermore, these parameters can be modified base on the location and versions update of the model [**?**].

Since we have the three main challenges to implement ML in cellular communication systems, i.e., the lack of ML vehicular channel models, the ML algorithms, and usage of deterministic or stochastic such as the empirical models. The current issue with ML cellular systems is the lack of an accurate ML channel model. Couple characteristics in the channel modeling need to be investigated with ML algorithms such as radio signal strength (RSS), delay spread, Doppler spread, phase shift and bit error rate. These characteristics usually used to improve the performance of the communications and ML will reduce the complexity. Cellular communication is a non-stationary channel where the medium channel changes during a short time. That change is due to the mobility of one or all transceivers or the environment by itself such as other moving vehicles cause time-variant. These actions also lead to delay spread and that is noticeable due to very high delay resolution [52].

By applying machine learning to the cellular systems to the wireless propagation channels to measure the parameters such as large scale and small scale fading such as path loss, shadow, delay parameter. Moreover, how to predict these parameters form a row data. One of the ways to predict these parameters is using Bayesian models such as the hidden Markov models that can use the current state data to predict the other parameters. Other methods may be used for estimating the channel model such as neural networks based on the dataset.

3.2.2 MmWave Channel Propagation Model in Cellular Systems

Similar channel model to the previous sections, channel modeling of cellular systems can be achieved using Deterministic and Stochastic Channel Models. The geometrical stochastic models such as the empirical models can be used to identify the large-small scale model. In this section, path loss models and how to apply machine learning techniques will be exploited. Moreover, there will be a couple of recent large-scale models to obtain the path models that consider the shadowing effects to be one of the parameters. PL describes how the mean received power varies with the distance between the transmitter and receiver where it's due to attenuation due to shadow fading between (SF) Tx and Rx. PL is a measure of the channel range and quality in the link budget. By matching the channel modeling formulas to the statistical analyses such as the path loss.

MmWave emerges as a new technology opportunity for 5G new radio. Most of empirical path loss models derive from the data collection and then compared with theoretical models such as free space and two-ray model.

$$PL(d) = PL(d_0) - 10n * log_{10}(\frac{d}{d_0}) + X_\sigma(d); d \geq d_0 \tag{3.8}$$

where $PL(d_0)$ is the path loss at reference distance d_0, n is the PL exponent which is an environment's dependent that is determined by field measurements, and X_σ is the large-scale fading about the mean power which follows normal distribution with zero mean and variance σ^2 as X_σ follows $N(0, \sigma^2)$. $PL(d_0) + [10nlog]_{10}(d/d_0)$ can be used as the mean of the path loss $\mu(d)$ and can be estimated and the large-scale fading X_σ can be also estimated using empirical measurements or regular estimation using a distribution. To obtain the large-scale fading, the effect of the small-scale fading is supposed to be removed from the data set which usually done by averaging the set over time sample or other methods.

ML still relies on relatively old-fashioned techniques such as logistic regression, SVM, decision trees, K-NN, naive Bayes, Bayesian modeling, ensembles, random forests, signal processing, filtering, graph theory, gaming theory, and many others to extract the channel modeling information. Involving machine learning techniques to estimate the path loss parameters still rely on traditional techniques such as regression and Bayesian theory. The PL candidates are the alpha-beta-gamma (ABG) model, the close-in (CI) model and Close-in with frequency-dependent exponent [74].

3.2.3 ML Channel Estimation

Due to the increasing complexity due to a large number of operations and large measurements, researchers tend to perform machine learning to enhances channel modeling. There are multiple machine learning estimator methods used to predict the channel parameters such as regression that use Least Square Estimator, the Minimum Mean Square Estimator [23]. These methods will obtain the channel state information (CSI) from the detected signals. Authors of [60] used regression to estimate the received signal strength over AoD bins. [21] have applied several of machine learning methods to estimate the wireless network using data measurements. The authors use Gaussian regression and random forests to predict where reducing the operational cost and enhancing the performance have been confirmed in this chapter. [75] used machine learning algorithms such as regression to predict the receive signal strength using beam training. This approach could raise the awareness of autonomous vehicle accidents using their locations using machine learning and ray tracing. [76] used linear and logistic regression to solve the switching band and band assignment problem then compared the results with different structures of neural networks. DNN can be used also to estimate the path loss as shown in [77].

ML can estimate the large-scale fading using supervised learning such as the Regression model. By using data to predict and infer the main parameters of the PL such as n and the large-scale fading σ. σ is from the error term in the general path loss equation and since the error is assumed to be i.i.d. where its mean is zero and variance is σ.

3.2.3.1 PL Estimation Using Regression

Regression is a machine learning under the supervise category [23]. Regression use LSE to minimize the square of the error between the observed responses in the dataset and to predict the most accurate model [21] [24]. LSE can be divided into two classes: linear (ordinary) least squares and nonlinear least squares where both of them don't require any prior knowledge.

Path loss (PL) considered the main goal for the link budget and by using a likelihood expression for a normal distribution to estimate the PL and the large-scale parameters. Where the data set is assumed to be Gaussian distributed due to a large amount of data that converge to normality and since large-scale fading is already assumed to be Gaussian. If we assume having N sample data point where $(X_1, Y-1), (X_2, Y_2), ..., (X_n, Y_n)$. By using the general form of multiple Regression to estimate the PL variables is:

$$y = f(x) + \epsilon \tag{3.9}$$

where y is the dependent variable (PL) and sometimes it called the output response, x is the independent variable and the ϵ is the theoretical error. However, since we have more than one sample the function of x would shown as:

$$f(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \tag{3.10}$$

The estimated of the response variable y and minimum goal of erroneous follow:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + ... + \hat{\beta}_p X_p \tag{3.11}$$

$$\arg\min_{j} \sum_{i=1}^{n} e_i^2 \tag{3.12}$$

Then, estimating the mean and variance can be shown below.

$$\hat{\alpha} = (X^T X)^{-1} X^T y \tag{3.13}$$

Using the above estimate, we can infer the variance σ^2;

$$\hat{\sigma^2} = \frac{1}{L-1}(y - X\hat{\alpha})^T(y - X\hat{\alpha}) \tag{3.14}$$

Machine learning estimation would be involved to predict the channel modeling parameters that would help to reduce the signal error $e(n)$. Also, Bayesian or ANNs methods can be used to predict the output parameters of the channel. Moreover, the minimum mean square estimator is a channel estimator which predict the modeling parameters that maximize the likelihood.

3.2.3.2 Support Vector Machine (SVM) in Wireless Channel

The support vector machine (SVM) that is derived from learning theory is a method to analyze data for classification analysis or regressions. SVM is used to identifying datasets using separating hyperplane. Linear regression is suitable for a complex problem and it was extended to SVM. Linear regression is the fundamental of the SVM. [78] showed how SVM method can help to predict path loss which made it easier and more reliable than the classical prediction. Zhao at el. used a method to estimate path loss by training PL data and then built a model that is capable of predicting the unmeasured data. Moreover, SVM can be used to classify by finding the hyperplane that differentiates the Los and NLoS.

3.2.4 K-Means Algorithm

K-means Algorithm is an unsupervised technique that can be implemented for wireless channel modeling to perform multiple of the process such as enhancing the CSI parameters using pattern learning to reduce the complexity based on data driven. As showing in [79], after collecting the measurement data and estimation the channel parameters, several studies have noticed these parameters such as the delay (τ), azimuth angle of arrival (AOA), azimuth angle of departure (AOD), elevation angle of arrival (EOA), elevation angle of departure (EOD) usually have the same cluster. [80] used a clustering algorithm to sort the MPCs with similar behavior such as delay and angular characteristics. Additionally, authors of [81] used K-power means clustering algorithm to automatically cluster the MPCs based on multipath component distance (MCD). MCD is used to enhance the clustering performance and [82] uses other distance measures and compares them. [38] is also proposed K-means-based clustering algorithm to classify the delay of MPCs and keep iteration the MPCs until the distance of the slimier data is minimized. The number of an optimum cluster can be calculated using the Kim-Park method [83] or other use numbers of a cluster as a random variable as in COST 259 models [84]. The total MCD between the ith and the jth multipath is shown as

$$MCD_{ij} = \sqrt{(||MCD_{Rx,ij}||)^2 + (||MCD_{Tx,ij}||)^2 + (||MCD_{\tau,ij}||)^2} \tag{3.15}$$

$$MCD_{\tau,ij} = \frac{|\tau_i - \tau_j|}{\Delta\tau_{max}} . \frac{\tau_{std}}{\Delta\tau_{max}} \tag{3.16}$$

K-power means (KPM) is a clustering algorithm that is used to assign a clustered index to each MPCs which is beneficial to minimize the MPCs base on their power.

$$\arg\min_j KPM = \sum_{l=1}^{L} P_l * MCD(x_l, C_j) \tag{3.17}$$

L is the number of MPCs, P_l is power of each MPC as $l = 1, ..., L$ and parameter vector x_l represent delay and other angular parameters. Then, iteration is used until it converged to its own clustering.

Zhang *et al.* in [49] proposed a channel modeling method to exploits the feature and structure obtained from the channel impulse response using an unsupervised method called principal component analysis (PCA). The CIR data obtained from measurement that was conducted in the teaching building of BUPT, China. [85] and [86] are one of the researchers that proved the ML techniques can improve the wireless channel modeling. The authors proposed a tracking algorithm based on maximization estimation and K-power means algorithm for the time-variant wireless channel and showed a good performance by clustering and estimating channel parameters using the nonlinear least square regression method. [87] purposed estimate the MPCs by space-alternating generalized expectation-maximization algorithm and used clustering methods such as MCD tracking algorithm and the K-power means algorithm.

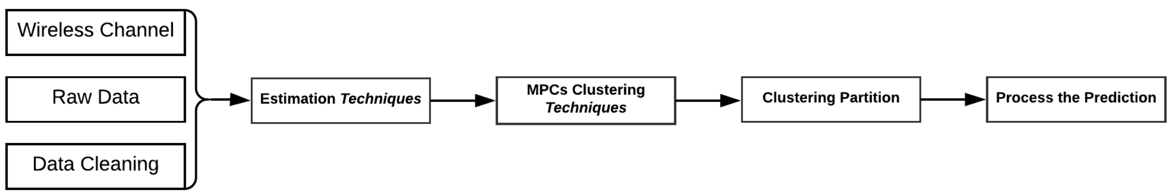

Figure 3.3: Flow Chart of Channel Cluster Modeling Processes

Figure 3.3 illustrates the processes of machine learning techniques usages in wireless channel modeling. The block diagram explains how MPCs are estimated and then clustering using the clustering algorithm such as K-power means. Data later will be fed to the clustering partitions to identify the observations into k clusters then the output shows the channel parameters. cluster validation is also recommended during the clustering processes by using methods such as the Calinski-Harabasz index and the Davies-Bouldin criterion to generate significant results [88].

In, telecommunication propagation the MPCs are cluster into the group based on LoS and NLoS and the strength of the received signal energy. Classical clustering such as k-means and spectrum clustering for the above causes. Li at.el on [31], compared two unsupervised methods, K-means, and GMM to model using clustering techniques in the area of wireless channel multipath. [56] applied Cramer-Rao lower bound (CRLB) to switch the CSI inference problem to extract the channel modeling parameters.

3.3 Conclusion

This chapter presents an overview of the theoretical framework toward machine learning problems for future wireless communication channel modeling. Even with the high demand to emerge machine learning in the field of modeling, there are still some challenges and block thoughts that need to be solved. This article presented the 5G in the channel modeling domain and summarized the current progress of applying ML to channel modeling. The main difference between the current LTE and 5G is the usage of data and how to involve machine learning to it. This exposition went over some useful methods in machine learning in the field of wireless communication such as regression analysis which was used for data fitting to predict the channel modeling parameters. Moreover, how classification identifies the input data to be mapped to the output which can be used to predict the traffic flow. Reinforcement learning is all about trying to understand the optimal way of making decisions/actions so that we maximize reward R such as Q-learning. Applying RL in channel modeling will lead to better performance such as autonomous vehicles.

Furthermore, other channel modeling scenarios that reach the complexity level such as fading, atmospheric effects, handover, beam direction, MIMO, suggest the time for machine learning to get evolved. Machine learning techniques show the viability proceeding with the estimation of the channel parameters and extracting the channel information as showing in estimating the path loss parameters using regression and clustering techniques.

Additionally, this chapter presented several open research topics that required more effort. So far, there are so many investments and time from academia and industry that have been invested in machine learning of the wireless channel modeling and we still have a long way to overcome the challenges and achieve the precise modeling. Involve machine learning to the wireless channel modeling is still in the preliminary phase and such that work will lead and guide scholars to achieve that goal. The scientific explanation of using machine learning in wireless communication channel modeling has been shown with pieces of evidence and reasons through this journey, while mobility introduces further technological opportunities. As a conclusion, applying effective predictive ML methods can reduce the complexity and increase the accuracy than the regular channel modeling. Involving ML techniques to the wireless channel modeling becomes more efficient as the number of measurements is reduced as good generalization abilities.

Chapter 4: Predicting the Path Loss of Wireless Channel Models Using Machine Learning Techniques in MmWave Urban Communications

4.1 Introduction

The classic wireless communication channel modeling is performed using Deterministic and Stochastic channel methodologies. Machine learning (ML) emerges to revolutionize system design for 5G and beyond. ML techniques such as supervise leaning methods will be used to predict the wireless channel path loss of a variate of environments base on a certain dataset as a model-free approach. The propagation signal of communication systems fundamentals is focusing on channel modeling particularly for new frequency bands such as MmWave. Machine learning can facilitate rapid channel modeling for 5G and beyond wireless communication systems due to the availability of partially relevant channel measurement data and model. When irregularity of the wireless channels leads to a complex methodology to achieve accurate models, appropriate machine learning methodology explores to reduce the complexity and increase the accuracy. In this paper, we demonstrate alternative procedures beyond traditional channel modeling to enhance the path loss models using machine learning techniques, to alleviate the dilemma of channel complexity and time-consuming process that the measurements take. This demonstrated regression uses the measurement data of a certain scenario to successfully assist the prediction of the path loss model of a different operating environment.

In this chapter, artificial-intelligence-learning-based prediction methods for CSI such path loss in 5G mmWave channels are proposed. Multiple AI algorithms are applied such, like random forest, and K-nearest- neighbors (KNN), are engaged to make predictions. "The principle of machine learning is to map the inputs to the outputs through flexible inside structures. The features selected as the inputs are often decided by the outputs. In general, different features have different importance. The features with low importance may reduce the generalization performance, over-fit, or increase computational complexity.

State-of-the-art channel modeling is the process of predictably incorporating wireless channel parameters into a channel model using the minimal number of measurements. Radio propagation models can be traditionally obtained via Deterministic and Stochastic Channel Models and applying a regular statistical method to build a model. These traditional methods are becoming more complex and time-consuming by using the new measurements while employing new technologies/frequency-bands and the increase of data traffic [46] [12]. In this paper, we are introducing a new procedure to overcome this dilemma of channel modeling by machine learning (ML) that emerges to revolutionize systems design for 5G and beyond. ML techniques, particularly supervise leaning methods, will be used to predict the wireless channel path loss, a key component of channel modeling. As mmWave frequency bands are widely introduced to 5G and thus require tremendous of new measurements, the irregularity of the wireless channels leads to a complex methodology in order to achieve accurate models. Machine learning algorithms, therefore, aim to reduce such complexity and increase accuracy while reduces the number of measurements. From the computational aspect, channel modeling can be considered as a sort of data mining and machine learning techniques considered as a valid solution to predict models instead of empirical and deterministic methods [89] [47]. [54] applied machine learning techniques to predict the carrier frequency offset (CFO) and presented better results toward applying machine learning to the wireless channel modeling. The prediction of the channel model emerges as a critical mission in modern AI-assisted communication systems. Furthermore, ML assists the extraction of useful information from the vast amount of channel measurement data in the wireless communication system [90]. Investigation of machine learning methods is suitable and capable to derive a channel modeling in better shape than the traditional ways as other researchers have accomplished and recommended [78].

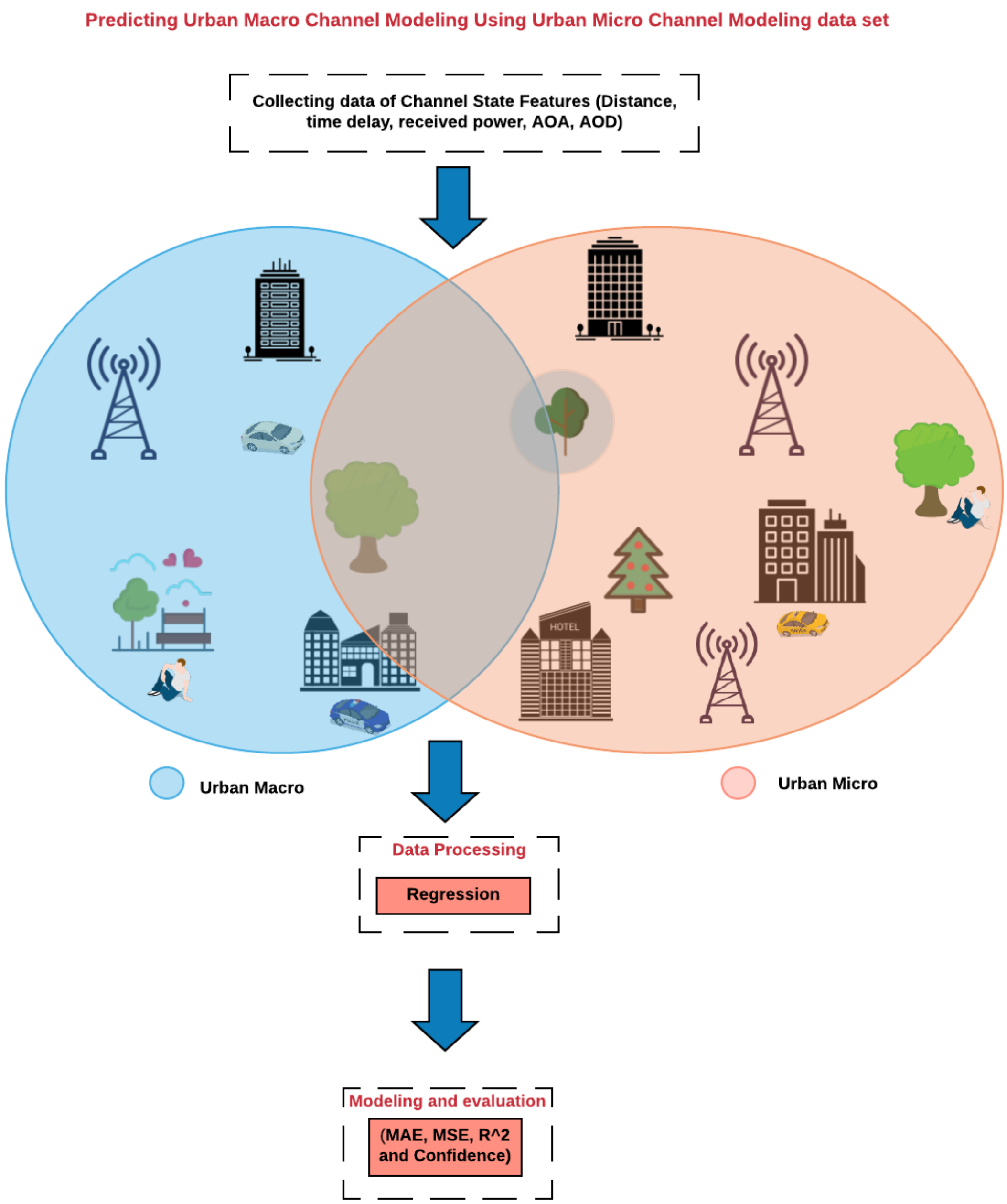

Figure 4.1: System Model of Urban Micro and Macro Channel Modeling Prediction

Figure 4.1 illustrates the proposed idea of this journey, which can be seen as instead of applying a specific measurement campaign to obtain a certain wireless channel model. ML can be used to predict a model of the variant of environments to predict a wireless channel model base on reliable data from a different environment. The traditional way is accomplished by conducting tremendous of measurements in a particular environment and obtaining its model using regular statistical techniques. The general mathematical model can be shown as:

$$y(t) = x(t) * h(\tau, t) = \int_{-\infty}^{\infty} h(\tau, t)x(t - \tau)d\tau \tag{4.1}$$

Where y(t) is the received signal, u(t) is the transmitted signal, * is the convolution sign and $h(\tau, t)$ is the delay spread function with respect to delay and time [91]. The characteristics of the wideband channel such as the power delay profile PDP, RMS delay spread, and other channel parameters are derived from the channel impulse response $h(\tau, t)$. The received passband signal is shown below:

$$Y(d, t) = \sum_{i=1}^{L-1} \alpha_i s(t - \tau_i) + n(t) \tag{4.2}$$

Large scale fading (LSF) usually due to the object that shadow the signal and explains the main characteristics of the channel such as path loss, shadow, angular spread, etc. Moreover, LSF cases will be examined the relationship between the path loss and the separated distance between the Tx and Rx in different environments such as the suburban.

$$APDP(\tau) = \frac{1}{N} \sum_{i=1}^{L} |h(t - \tau_i)|^2 \tag{4.3}$$

In this manuscript, common solutions to model large scale path loss are introduced in II. Followed by section III demonstrate how machine learning will be involved in wireless channel modeling. Then, generating a dataset of this work explained in section IV. Then models validation and results in V. Lastly, a conclusion is shown in section VI.

4.2 Conventional Model of Large Scale Path Loss

The common ways to predict the path loss in a channel model vary base on different characteristics, such as environments, types of antenna and frequency scales. Inferring the path loss of a different communication environment using existing path loss models or data from different environments has not been well investigated in the literature, which serves our target and novelty with the aid of ML. There are non-ML types of path loss models that are used to predict the signal loss of the propagated link via a wireless channel. These path loss models are briefly explained in the following subsections.

4.2.1 Close-in (CI) Model

This model is usually for LOS and NLOS for all urban micro (UMi), UMa, and InH as well by using close-in reference distance base on Farris' law. The general version of the close-in model is [16]:

$$PL^{CI}(f,d)[dB] = FSPL(f,1,m) + 10nlog_{10}(\frac{d}{d_o}) + X_{\sigma}^{Cl}, \tag{4.4}$$

The parameter FSPL is just the free space model in dB, n is the path loss exponent (PLE) to show how PL varies with multipath propagation distance and d_o is the reference distance which set to 1 m since there is rarely shadowing in the first meter and simplifies the equation as well. The above PL models should apply to our usage in the work and have the form of a linear regression model where other path loss models such as the following models can be applied to multiple linear regression due to the many channel features.

4.2.2 Close-in with Frequency Dependent Exponent

This model was proposed in the 3GPP meetings where it depends on how high is the frequency. Close-in with frequency dependent is an extension of the CI model with frequency dependent [18]. The general model of the CIF model shown as [19].

$$PL^{CIF}(f,d) = FSPL(f,1,m) + 10n(1 + b(\frac{f - f_o}{f_o})) \\ log_{10}(\frac{d}{1m}) + X_{\sigma}^{CIF} \quad (4.5)$$

here b is a parameter that captures the slop or the dependence of the path loss of the weighted average of the reference frequencies f_0 and it would be positive if both PL and f increase.

[20] compared the above path loss model for both μWave and MmWave in different scenarios such as UMa and UMi. All of them showed a good prediction with large data where the CI model was the most suitable for outdoor cases due to the close-in free space reference. Whereas, CIF has a better performance for the indoor environments due to its small stander deviation values. The path loss exponent in CI/CIF models shows loss with distance for urban macro then urban micro and that seems applicable due to the obstructions blocking the signal from than lower base stations while the urban macro is commonly higher than the micro communication.

4.2.3 Floating-Intercept (FI) Model

FI model is also called the alpha-beta (AB) path loss model. This PL model can be combined with log-normal shadowing as shown:

$$PL^{FI}(f,d)[dB] = \alpha + 10\beta log_{10}(d) + X_{\sigma}^{FL} \quad (4.6)$$

The values of α and β can be obtained using the least square fitting as a slope and floating intercept respectively. Also, the shadow fading is represented by X_{σ}^{FL} fallowing a Gaussian random variable with zero mean and standard deviation of σ.

4.2.4 Alpha-Beta-Gamma (ABG) Model

ABG model is the current 3GPP 3D model and values may change based on the base station location.

$$PL^{ABG}(f,d)[dB] = 10\alpha log_{10} + \beta + 10\gamma log_{10}(\frac{f}{1GHz}) + X_{\sigma}^{ABG} \tag{4.7}$$

ABG model is used to measure how PL get increased with distance and α is the slope of PL with log distance. β is the optimized floating offset in dB, γ is the PL variation dependence over a frequency in GHz and X_{σ}^{ABG} is the fading (SF) in dB. Since there are three parameters, 3GPP mentioned that ABG PL model always has a lower shadow fading standard deviation than other PL models [15].

4.3 Modeling the Path Loss Using Machine Learning Techniques

Estimating a path loss can be solved by machine learning techniques to overcome challenging issues such as complexity and time consuming due to the required tremendous measurements. The classic wireless communication channel modeling is performed using Deterministic and Stochastic channel methodologies. In the following, ML techniques such as supervise leaning methods will be used to predict the wireless channel path loss of a variate of environments base on a certain dataset, with application scenarios like mmWave bands communication channels. In this paper, we demonstrate applying machine learning to develop alternative procedures to enhance the path loss models using machine learning techniques. Furthermore, the investigation of machine learning methods is suitable and capable to derive a channel modeling in better shape than the traditional ways as other researchers have accomplished and recommended [78]. This section will presents how to apply ML methods to estimate the channel parameters using different regression methods. Regression is a basic supervised learning technique [23]. Regressions use least square error (LSE) to minimize the square of the error between the observed responses in the dataset and to predict the most accurate model [21] [24]. Regression is one of the main methods that is used in machine learning where the regression models learn the mechanism based on a dataset from prior measurements or simulations. After the learning processes are completed, the model coefficients can be obtained. Authors of [23] applied support vector and DNNs regressions then made a compression to control the high-speed channel modeling errors. Where all of these techniques help to reduce the wireless channel modeling complexity and need to be investigated more. The machine learning techniques that will be used in this work are linear and multiple linear regression algorithms. Multiple linear regression techniques take the advantages of other channel modeling features to enhance the path loss prediction comparing to regular linear regression and that can be seen in the future result section. Furthermore, this journey will test how the wireless channel features affect the path loss prediction. The reason for using regression techniques instead of other machine learning methods is due to the desire of prediction continuous values.

$$Y_i = f_i(\boldsymbol{X}) + \epsilon_i \tag{4.8}$$

Y is the dependent response which is in our case the path loss, X is the independent variable in form of $\boldsymbol{X} = [x_1, x_2, \ldots, x_p]$ which is the channel state information (CSI) features such as distance, time delay, received power, azimuth AoD, elevation AoD, azimuth AoA, RMS Delay Spread, and frequency (GHz). In order to make a prediction of Y on new data, we need to estimate $\hat{f}(X)$. Thus, the estimate coefficients have to be accurate as possible to enhance the accuracy. Path loss models suppose to be applicable and have the form of regression algorithms such as linear regression for Floating-Intercept (FI) model. While other path loss models such as the Alpha-Beta-Gamma (ABG) model can be applied to multiple linear regression due to the other channel features in the previous section . By considering linear regression where distance is the only channel feature used to estimate the path loss model as show below and estimating the parameters can be seen in equation 14.

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i X_i \tag{4.9}$$

$$E(\beta_0, \beta_i) = \sum_{i=1}^{p} [y_i - (\beta_0 + \beta_i X_i)]^2 \tag{4.10}$$

By applying this approach to estimate the the coefficient parameters and then characterizing the theoretical loss L is obtained.

$$\arg\min_{j} L = \frac{1}{N} \sum (\hat{Y}(k) - Y(k))^2 \tag{4.11}$$

4.3.1 Multiple Linear Regression

Multiple linear regression techniques take the advantages of other channel modeling features to enhance the path loss prediction comparing to regular linear regression. Furthermore, this journey will test how the wireless channel features affect the path loss prediction. Machine learning techniques would be adopted to estimate the channel modeling parameters that would reduce the estimation error $e(n)$. An example of the ML techniques is the multiple linear regression method that can be used to predict the modeling parameters of the channel following the ABG model that was introduced by the 3GPP. Multiple linear regression is a supervised learning and the goal is to infer and predict a function by reducing the error using the training data to predict the target by using machine learning method that perform a better estimation $\hat{f}$. In our scenario, using dataset that divided into two part, training and testing where the training set consists of 70 % and the test is 30% and other as mentioned in [56]. The multiple linear equation can be seen as follow [26]:

$$f_i(x) = \beta_0 + \beta_1 X_1 + ... + \beta_i X_i + \varepsilon \tag{4.12}$$

The estimated of the response variable Y (PL) using X as CSI and minimizing the erroneous follow:

$$\hat{y_i} = \hat{\beta_0} + \hat{\beta_1}\hat{X_1} + ... + \hat{\beta_i}X_i \tag{4.13}$$

Then, to maximize the estimation method, the coefficient of equation 16 has to be minimized to obtain a lower difference between the real and the estimation equation. The residual error of the regression estimation can be obtained using the below equations [2].

$$Minimize \sum_{i=1}^{n} e_i^2 \tag{4.14}$$

$$e_i^2 = \arg min \sum_{i=1}^{n} (y_i - \hat{y_i})^2 \tag{4.15}$$

$$(\beta_0, ..., \beta_i) = \arg min_{\beta_0 ... \beta_1} [\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\beta_0} - \hat{\beta_1} X_1 - ... - \hat{\beta_i} X_i)^2] \tag{4.16}$$

Then, estimating the of the slope and variance can be shown below [27].

$$\hat{\beta_1} = (X^T X)^{-1} X^T y \tag{4.17}$$

Using the above estimate, we can infer the variance σ^2 in order to find the shadow fading parameter of the Close-in and ABG models ()equations 7, 10 and 11).

$$\hat{\sigma^2} = \frac{1}{L-1}(y - X\hat{\alpha})^T (y - X\hat{\alpha}) \tag{4.18}$$

4.4 Data

Investigation the channel modeling would allow other applications to interchange data to make the communication between them more precisely. Getting machine learning involved in measurement from simulation or campaign with ML algorithms will provide homogeneous works and unbiased results. Machine learning algorithms investigate the feature of wireless channels deeply in MmWave. Machine learning is used to improve performance and reduce complexity. By using a measurement dataset, ML methods would be applied to enhance the accuracy or to interpret/extend non-measured scenarios. Channel modeling parameters are generated by measurement campaigns or simulations and the propagated signal through a channel that gets disturbed by fading which leads to MPCs.The highest MPCs are the strongest link and from there, channel parameters can be obtained to create a dataset. We modified an open-source Matlab simulation that was provided by New York University throughout their wireless lab [92] [93] to meet our specifications. Then, that simulation was used to obtain a sufficient amount of data to enhance the accuracy of the models. Then, we purposed methods of using multiple machine learning techniques and the generated data and then do the interpretation of performance comparison between the algorithms to check the path loss model as shown in the results section. The regression techniques in consideration include linear regression methods. Python was used to preform data analysis.

Table I., exhibits that the channel measurement parameters of the data raw that was used for this paper. Regression is considered the main method to investigate the relationship between the channel features [33]. With the glory of having a large amount of data, the behavior of the wireless channel modeling becoming more interesting and obvious to obviate the complexity. The step following cleaning the data is applying the machine learning scheme to stars the learning processes. Then, a model can be used for predicting the path loss and

Table 4.1: Channel Measurement Parameters

Parameters	**Values**
Distance (m)	1-40
Frequency (GHz)	28
Bandwidth (MHz)	800
TXPower (dBm)	300
Scenario	UMi
Polarization	Co-Pol
TxArrayType	ULA
RxArrayType	ULA
Antenna	SISO
Tx/Rx antenna Azimuth and Elevation (red)	10

evaluating the model will be accomplished by Mean Absolute Error, Mean Squared Error, Root Mean Squared Error and R-square as shown in the result section.

4.5 Results

Checking how significant the data is performed by using residual plots, where plot, as shown in figure 4.2, can demonstrate how the data distributed among the horizontal line and it appears reasonably random. Thus, it confirms the data used for regression is unbiased. Residuals method is used to forecast errors that can be obtained by subtracting the forecast from the expected values.

Below Figure shows the prediction of path loss using the linear regression method and evaluating this model is shown in table III. While table II explores the coefficient parameters of the Linear Regression (LR), Multiple Linear Regression (MLR1) and Multiple Linear Regression (MLR2). The second model (MLR1) implemented with only three wireless channel feature while the third model used eight features.

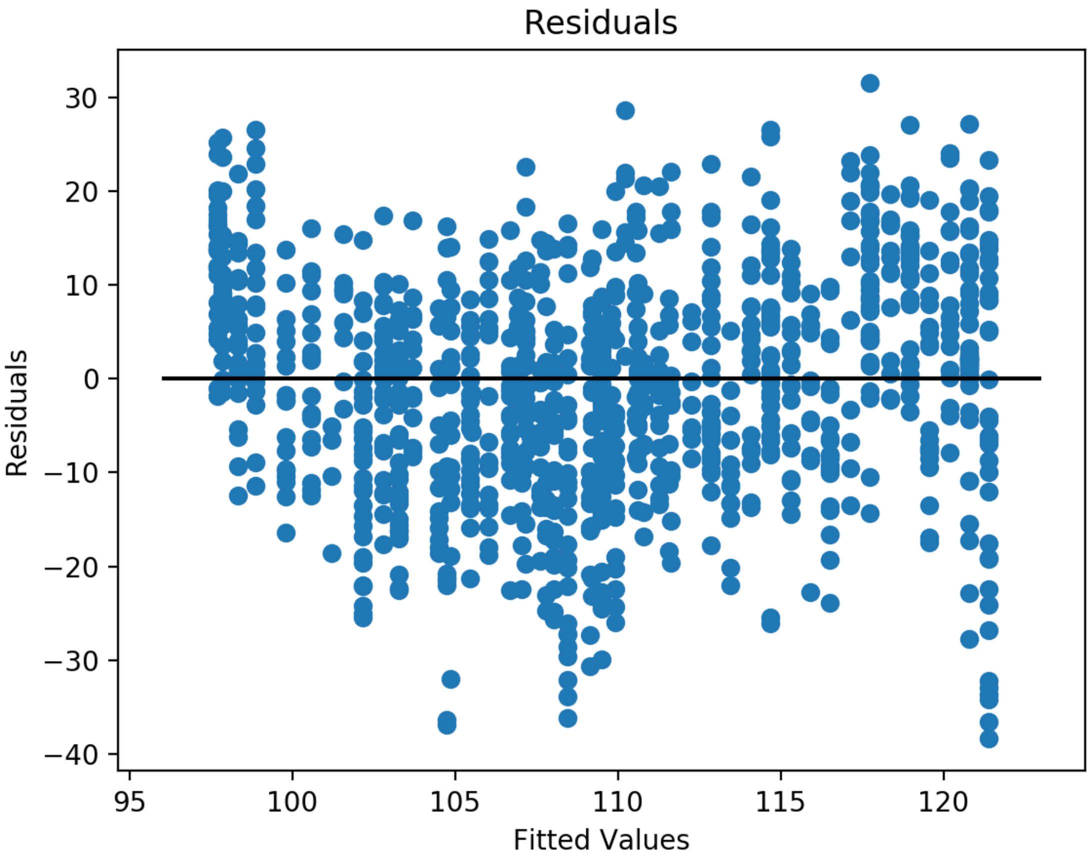

Figure 4.2: Residuals Plot

While figures 4.3 and 4.4 demonstrate two model urban micro (Uma) and urban micro (Umi) of wireless communication. The date belongs to umi that was used to generate a regression line, while the regression line is the uma model that was obtained from uma scenario and then applied to the umi.

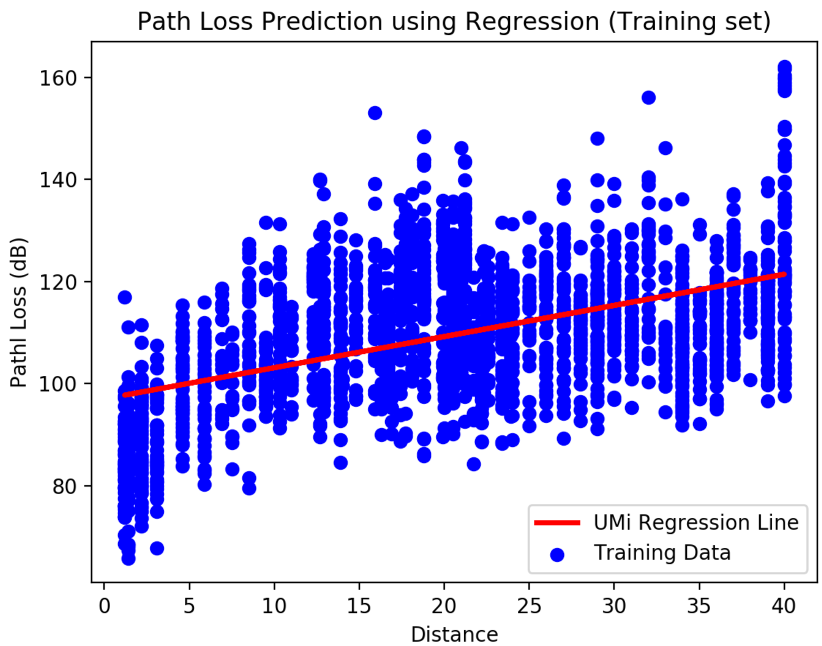

Figure 4.3: Training Path Loss Prediction Using Linear Regression Algorithm

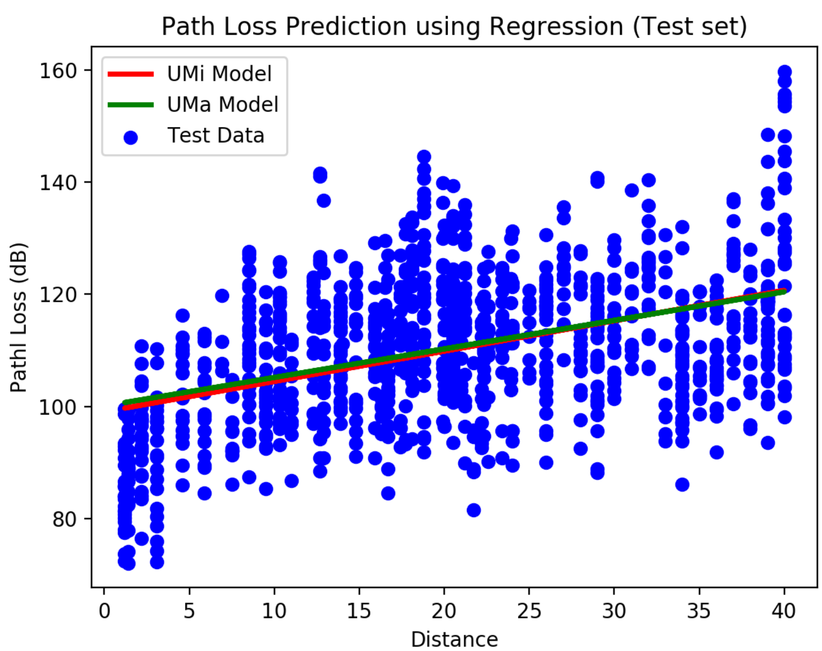

Figure 4.4: Predicting Path Loss Using linear Regression Algorithm

With the usage of different wireless channel features, table II illustrate an adequate work by comparing the results of applying the model that was obtained from micro/urban environments with table I specifications and applied it in the macro/urban communication. Thus, the wireless channel measurements can be reduced by applying a model from a single environment to others by applying machine learning techniques that can learn the logic.

Table 4.2: Communication Scenarios Comparison

Environment Scenario	**UMi**	**UMa**
MAE	8.92	6.66
MSE	126.60	74.32
RMSE	11.25	8.62
R Square	0.21	0.533
Confidence	0.21	0.533

Table 4.3: Channel Measurement Parameters for UMi Communication

Test	**LR**	**MLR**	**MLR**
T-R Separation Distance (m)	0.56	0.46	0.48
Time Delay (ns)	-	-0.08	-0.09
Received Power (dBm)	-	-0.69	-0.69
RMS Delay Spread (ns)	-	-	0.29
Elevation AoD (degree)	-	-	-0.10
Azimuth AoD (degree)	-	-	-0.002
Azimuth AoA (degree)	-	-	-0.004
Elevation AoA (degree)	-	-	-0.001

As can be followed by equation 4 and 5, the estimated path loss model shown as for linear

Table 4.4: Linear Regression Model

Environment	α	$L_0[dB]$	$X_\sigma[dB]$
Outdoor Micro Urban	97	.51	13.6

regression and a single feature loss $L_0[dB]$ as the separated distance.

$$\hat{PL} = \alpha + 10logL_0[dB](d) + X_\sigma[dB] \tag{4.19}$$

While for the model for multiple regression that consists of multiple wireless channel features loss $L_N[dB]$ as equation 7 proved shown as:

$$\hat{PL} = \alpha + L_0[dB] + L_1[dB] + ... + L_9[dB] + X_\sigma[dB] \tag{4.20}$$

Both parameters of above two equations can be obtained from tables III and IV. Moreover, using the statistical parameters Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and R square (R^2) values level to achieve the significant of the predicted or the used model. RMSE is the square root of MSE and used to check the accuracy of the wireless channel propagation predication where it measures the differences between the predicted and observed model where zero value indicates the fit is optimum [6]. Furthermore, these parameters can be used to validate the significance and check the accuracy of the proposed models. Table III illustrates the analysis of this journey, where there is three model that can predict the path loss of an outdoor Micro environment using a 28 GHz. The features that are used in the second model are T-R Separation Distance (m), Time Delay (ns) and 'Received Power (dBm). While the third model's features are T-R Separation Distance (m), Time Delay (ns) and 'Received Power (dBm), RMS Delay Spread (ns), Elevation AoD (degree), Azimuth AoD (degree), Azimuth AoA (degree) and Elevation AoA (degree). Then these models are evaluated using Mean Absolute Error, Mean Squared Error, Root Mean Squared Error and R-square. From table III, multiple linear regression with eight features performs the best among other models which leads to increasing the feature enhance the prediction but until the model reaches the steady-state. R square particularly presents how the models improved with increasing the number of channel variables which provide an acceptable prediction result.

Table 4.5: Micro Urban Channel Measurement Parameters

Test	Linear Regression	Multiple Linear Regression (3 Feature)	Multiple Linear Regression (7 Feature)	Feature Selection with 4 Features
MAE	8.92	6.66	5.10	5.52
MSE	126.60	74.32	44.51	52.3
RMSE	11.25	8.62	6.67	7.23
R Square	0.21	0.533	0.73	0.72

4.6 An Update Results Using Different ML Algorithms

Current and previous wireless generation technologies have some technical concerns that have not been solved or partially solve for a specific generation. In this manuscript, alternative procedures are applied to overcome these issues using Artificial intelligence (AI). Machine learning techniques have the capability to interpolate and extrapolate the channel state's values of a wireless communications application based on a given dataset to overcome wireless channel modeling communications matters. Issue such as the complexity of empirical wireless channel modeling is overcoming using the method of machine learning (ML) techniques. Methods such as regression techniques have been implemented to enhance the communications to meets the 5G requirements. Methods of machine learning and deep learning have been compared to ensure optimum accuracy is reached. Throughout this manuscript, the robustness of AI methods will be illustrated to revolutionize the system design of 5G - new radio (5G-NR) and beyond that led to reducing the complexity, and enhance the reliability.

Machine Learning methods are used to develop alternative procedures to enhance the wireless communications of different applications in the 5G era to reach the 3GPP requirements. In this manuscript, two wireless communications issues will be investigated to ensure the AI techniques capability to overcome them. The issues are the complexity of modeling wireless channel modeling. The classical approaches either by applying empirical or deterministic approaches. The deterministic methods are based on theories such as ray tracing, Maxwell equations and others [?]. While the stochastic is based on empirical measurement models such as Okumura–Hata [?] and COST 231 [?]. The path loss models are an estimate based on varieties of factors such as the transmitted signal power, distance, frequency, antenna high, etc. For every environment, there must be a unique model that estimates the attenuated strength of the signal. Channel modeling is the process of incorporating wireless channel parameters into a model that has the capability of forecasting and making a prediction. The propagated electromagnetic waves usually face the surrounding environments that cause the signal to be reflected, diffracted or propagated through the medium that leads to multipath components and selecting the optimum signal is the second issue in this work [46] [12]. Investigating and overcoming these concerns are based on applying AI techniques that shows great results as shown in selection slowromancapii@. Millimeter-wave provides an alternative frequency band of wide bandwidth to recognize pillar technologies. These requirements were introduced by International Telecommunication Union (ITU) to provide ultra-reliable low-latency communication (URLLC), machine-type communication (mMTC) and enhanced mobile broadband (eMBB) to the 5G applications [?]. The wireless channel modeling is unstable and these effectiveness signals degregate the wireless communications and cause fading, interference and other wireless issues. The wireless channel modeling is the middleman between the transmitter and the receiver in the wireless communication systems. Knowing and learning the propagation environments channel requires either a deterministic or stochastic method such as the ray tracing or the empirical method which are complex and inaccurate [?]. Moreover, both these methods are time consuming and every scenario has to be measured and there must be a way to compromise them to meet the 5G-NR requirements. It's time to adopt and apply machine learning techniques toward wireless communication issues.

Reducing the complexity and other cons in the classical wireless channel modeling. ML techniques have been used to investigate this issue where the regression method is used to predict the path loss model instead of the classical approaches. Details of the regression method will be shown next section

4.7 Wireless Channel Modeling

A fundamental design of the wireless communication system is the radio frequency channel modeling. The state of the art of channel modeling is defined as a fundamental part of the physical layer in communication systems that can be represented by mathematical parameters to model the channel. Wireless channel is the transmission medium for mobile communications which is the process of involving wireless channel parameters into a model that has the capability of forecasting and making a prediction. [?] have proposed the Extreme Learning Machine (ELM) algorithm to predict the path loss model for lower microwave frequency of an outdoor scenario to modify the base station deployment. The propagated signals face an issue with the surrounding environment that causes the signals to be either destructive or constructive during the propagation. During the past decade, deterministic and empirical methods were performed to measure the amount of degradation of the transmitted signal and that is still considered as a drawback especially with the requirement of the 5G- new radio.

The first issue that would be investigated in this manuscript is the complexity of conducting measurement campaigns in every single scenario is a time consuming and inaccurate. The classical procedure to come up with a wireless channel model is to establish a new measurement campaign for every environment such as urban, rural, suburban and others. Instead of measuring every single field to obtain a path loss model, a regression can be used to estimate a path loss model from a specific field. An example, building a new wireless channel modeling for urban environments requires performing a measurement campaign to collect data to build the model. While with machine learning regression techniques, that urban model can be estimated if we have any previous data from a different environment. Meaning, having a single environment dataset can be used to create a new model of different environment. This method will reduce complexity and reduce the time and the number of money [**?**].

4.8 Updated Results

In this section, both regression techniques will be used to investigate wireless channel modeling. Where the regression approach will be applied to assure the capability to predict a certain wireless environment based on a dataset from a different environment. Three methods of regression techniques will be used to overcome the complexity of wireless channel modeling. These methods are linear, multiple linear and Support Vector Machines Regression.

4.8.1 Linear and Multiple Linear Regression

The classic wireless communication channel modeling is performed using Deterministic and Stochastic channel methodologies. Nowadays, machine learning (ML) emerges to revolutionize system design for 5G and beyond. ML techniques such as supervise leaning methods will be used to predict the wireless channel path loss of a variate of environments base on a certain dataset. The propagation signal of communication systems fundamentals is focusing on channel modeling particularly for new frequency bands such as mmWave. Machine learning algorithms can facilitate rapid channel modeling for 5G and beyond systems such as cellular communications systems due to the emerging of wireless big data. When irregularity of the wireless channels leads to a complex methodology to achieve accurate models, machine learning algorithms explore to reduce the complexity and increase the accuracy that reduces the number of measurements. In this paper, we demonstrate applying machine learning to develop alternative procedures to enhance the path loss models using machine learning techniques. Due to channel complexity and time consuming the measurements take, this journey develops alternative procedures to reduce the channel measurements that are used to predict the path loss model. This can be done by using the measurement of a certain measurement scenario to predict and assist the prediction of path loss model of a different environment.

Estimating the path loss model such as the floating model as (1) can be taken to the form of regression where linear regression has only one channel feature which is distance. While multiple linear regression consists of many parameters that represent different wireless channel features such as delay, receive signal strength, distance and others. t

Floating-Intercept (FI) Model is a path loss model and can be estimated using the regression method as shown:

$$PL^{FI}(f,d)[dB] = \alpha + 10\beta log_{10}(d) + X_{\sigma}^{FL} \tag{4.21}$$

$$\hat{Y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i X_i \tag{4.22}$$

$$E(\beta_0, \beta_i) = \sum_{i=1}^{p} [y_i - (\beta_0 + \beta_i X_i)]^2 \tag{4.23}$$

By applying this approach to estimate the the coefficient parameters and then characterizing the theoretical loss L is obtained.

$$\arg\min_{j} L = \frac{1}{N} \sum (\hat{Y}(k) - Y(k))^2 \tag{4.24}$$

4.8.2 Support Vector Machines Regression

Support Vector Machines (SVM) is another method of supervised ML algorithms that can be used in either regression or classification [94]. In this section, SVM used to fit the data in terms of regression based on a theoretical foundation based on the Vapnik-Chervonenkis theory to minimize the error. The basic idea of SVM is to map a dataset from a finite dimensional space to a high dimensional space in a form of linear and nonlinear shapes. The extension of that technique is support vector regression that used in regression techniques in continuous cases.

$$f(x, w) = \sum_{j=1}^{m} wx + b \tag{4.25}$$

SVR is a regression method that can be used toward the wireless channel modeling and it's different from SVM. The output of the algorithm supposes to be continuous values instead of Classification which is categorical. The function that will be used to describe these methods is called karnel and used to map a lower dimensional data into a higher dimensional data. Its well known that SVR use a hyperplane to predict the response instead of using it as a separate line in the SVM to distinguish the classes. An advantage of applying SVR is the absence of the local minima to and used instead of the margins [**?**].

SVR is different from the regular linear regression by fitting the error within a certain threshold instead of minimizing the error using least square error. Furthermore, the fitting line in the SVR is base on the maximum data points within certain boundaries.

$$wx + b = 0 \tag{4.26}$$

While the boundaries are shown as

$$-\epsilon \leq y - wx - b \leq \epsilon \tag{4.27}$$

where ϵ is the error of data point to each boundary.

$$e_i^2 = \arg min \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \tag{4.28}$$

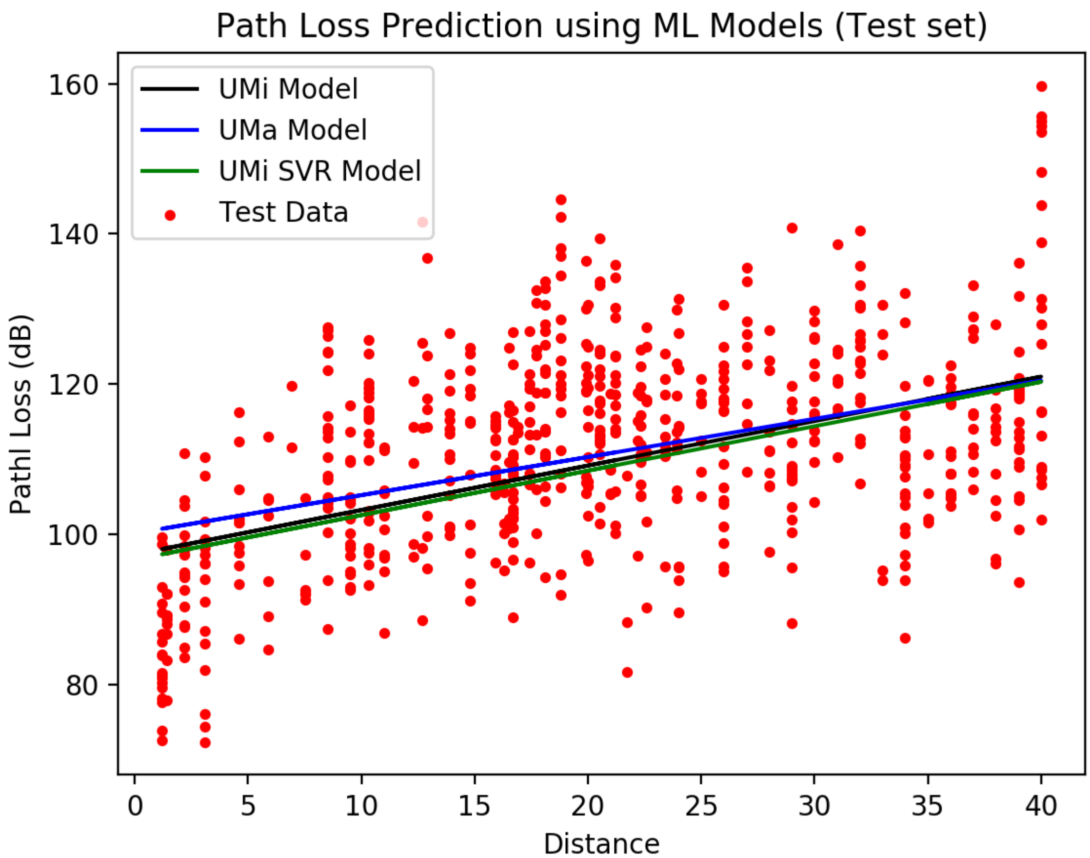

Figure 4.5: ML Regression Fitting Line Techniques

Figure 4.5 shows three models, the red dots are simulated data from MUi environment and the black model is urban micro-cell (UMi) model. The blue line is an urban macro model which can be seen how close the model with an R-square of 0.21 for UMi and 0.53 for UMa which consider a great results. While the green regression line is linear SVR.

Conveniently, this procedure helps with reducing the time consuming to collect several measurements in terms of modeling the wireless channel modeling, enhances accuracy and reliability. Solving these issues can be applied to other applications such as IoT and other networking utilization. Determining and solving this issue meet the 5G-NR pillars technologies of enhancing mobile broadband (eMBB) and ultra-reliable and low-latency communication (uRLLC) and massive machine type communications (mMTC). The loss of the SVM can be obtained using the below formula can be shown as following.

$$\mathcal{L}_\epsilon(y, f(x,w)) = \begin{cases} 0 & \text{if } |y - f(x,w)| \leq \epsilon, \\ y(x) - f(x,w) - \epsilon & \text{o.w..} \end{cases} \tag{4.29}$$

Outliers can cause problems with certain types of models. For example, linear regression models are less robust to outliers than decision tree models. Generally, we should not remove outliers until we have a legitimate reason to remove them. Sometimes, removing them improves performance, sometimes not. So, one must have a good reason to remove the outlier, such as suspicious measurements that are unlikely to be the part of real data.

4.9 Conclusion

Inaccuracy, complexity, and number of measurements of the wireless communications have been not solvable issues for past decay. This chapter presented a new ML procedure to overcome this issue with the assist of machine learning techniques. The traditional solutions base on top wireless communication organizations have not sufficiently overcome these issue and with the new era of big data, it's the time to resolve them base on machine learning algorithms. A new approach to applying supervises learning to model the wireless channel. We have used regression techniques to defeat channel modeling issues. Using the data of a certain communication environment, we were able to predict the model of the new communication scenario. Thus, the required number of measurements and the complexity have been reduced.

Multiple AI algorithms have been covered to overcome wireless communications drawbacks in a variety of communication applications such as the wireless system in terms of 5G and beyond. The capability and potential of Machine learning (ML) techniques that have been applied in this journey to meet the 5G-NR. The purpose of this letter elaborates on the ability of machine learning to solve other communications issues. Regression techniques have been used as well to overcome the complexity and time consuming of collected measurement of every single field to build communication models. Other machine learning techniques already used such as SVM, logistic regression to overwhelm other wireless communication problems.

Chapter 5: Predicting Optimum Propagated Link for 5G New Radio via Artificial Neural Networks

5.1 Introduction

Millimeter-wave supplies an alternative frequency band of wide bandwidth to better realize pillar technologies of enhanced mobile broadband (eMBB) and ultra-reliable and low-latency communication (uRLLC) for 5G - new radio (5G-NR). When using the mmWave frequency band, relay stations to assist the coverage of base stations in the radio access network (RAN) emerge as an attractive technique. However, relay selection to result in the strongest link becomes the critical technology to facilitate RAN using mmWave. An alternative approach toward relay selection is to take advantage of existing operating data and apply appropriate artificial neural networks (ANN) and deep learning algorithms to alleviate severe fading in the mmWave band. In this paper, we apply classification techniques using ANN with multilayer perception to predict the path loss of multiple transmitted links and base on a certain loss level and thus execute effective relay selection, which also recommends the handover to an appropriate path. ANN with multilayer perception is compared with other ML algorithms to demonstrate effectiveness for relay selection in 5G-NR.

Relay selection [95] to form cooperative communication has become a critical technology with the 5G - new radio (5G-NR) and future mobile communications. Relay selection in multi-hop communication was shown as an adorable technique for mobile communication over mmWave frequency bands [96,97] The sensitivity of mmWave signal to fading remains a fundamental challenge in communication systems, especially for 5G era. Authors of [98] have proposed a novel adaptive multi-state selection utilizing different of mmWave frequencies. The fifth-communications generation goals are prioritized base on three pillars which are enhanced mobile broadband (eMBB), ultra-reliable, low latency communications (URLLC), and massive machine type communications (mMTC). In this work, we are trying to meet the first two goals base on relay selection with a new mechanism that increases communication strength and improves reliability. Moreover, this work may enhance communication between massive devices to meet the mMTC. During the communication propagation, the transmitted signal can be affected by the surrounding environments resulting in signal diffraction, scattering, and reflection as showing in Figure 5.1. Having a line-of-sight (LOS) transmission does not mean obtaining a proper transmission but relying on other propagation links may have better performance and coverage. Figure 5.1 demonstrates three propagated signals from (BS) where the first link PL_{LOS} is the LOS signal, $PL_{Bs,r1}$ is the second transmitted signal which handover to another station and the third link is affected due to obstacles OB then penetrate through to the destination mobile station MS. To allow the destination point to select the best link base on the propagation signal strength to meet the 5G -new radio (NR) requirement (the ultra-reliable and low latency communications) in terms of reliability with 99.999% [99]. In this paper, we demonstrate a new machine learning methodology that acculturates the link selection from base stations or the user's equipment side.

$$C_i(x) = \begin{cases} 1 & PL < 120dBm \\ 0 & PL \geq 120dBm \end{cases} \tag{5.1}$$

Where $C_i(x)$ is the link selection classes which depends on the path loss of the link that can be calculated using models such as Floating-Intercept (FI) model.

$$PL^{FI}(f,d)[dB] = \alpha + 10\beta log_{10}(d) + X_{\sigma}^{FL} \tag{5.2}$$

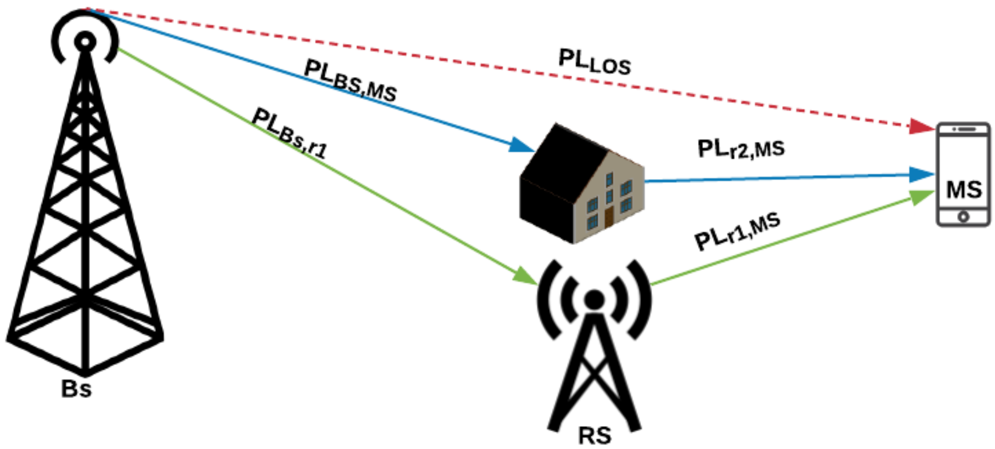

Figure 5.1: System Model of Wireless Signal Propagation Selection

$$BL = \arg\min_{L_s}(PL_n) \tag{5.3}$$

Where BL is the best link selection and PL_n is the path loss of the propagated links. While n is the number of transmitted links between the base station and destination.

To meet the 5G pillars technologies particularity the eMBB and uRLLC, we propose a new mechanism to overcome the affected propagation link using artificial neural networks. Using relay link selection, we note that the classification of machine learning is suitable for description and prediction of the optimum link/path. Having a reliable mechanism to meet the uRLLC with trustworthy communications and eMBB to enhance the coverage and improve the communications. The supervised classification algorithms can be performed to predict categorical class labels using the training dataset of the outdoor urban environment and can be implemented by the artificial neural networks (ANN) with deep learning. Furthermore, Multilayer Perceptrons (MLP) methods in ANN, together with different models, are evaluated to predict and select the strongest propagation link. MLP, therefore, serves the classification in ANN to identify and characterize the new link candidates using the path loss parameter or the receive signal strength. This work also will influence the massive machine type communications (mMTC). The classification technique that was selected is a binary where there are only two classes which are strong links and weak links. Once we obtain propagation results, the new mechanism of this work divides the signal losses into categories, these levels are classes. In our case, we have binary classes, where each path loss signal strength is either considered sufficient or insufficient (no fading). The base station makes a decision based on that categorical. Thus, identifying the optimum link using classification algorithms to meet the reliability and coverage of the 5G NR. The base station learns how to predict the weakest path loss and select the minimum loss path. The base station selects the propagated signals base on a certain energy strength (threshold) and once that link energy reaches this threshold the base station will switch to another link. In this study, the threshold is base on the path loss which is equal to -120 dBm and below this threshold is considered a poor propagation. Thus, eMBB and uRLLC will be achieved.

[1] suggests using deep learning to identify and classify the modulation nodes, improving the interference alignment and locate the optimum routing path. Furthermore, applying prediction techniques using methods such as classification and clustering, etc. to estimate the channel path loss models that lead to better performance and precision. Multilayer Perceptrons is a ML classification technique which is a neural network. The data can be classified based on maximum probability in Multilayer Perceptrons techniques to predict the path loss that can be expressed as:

$$\hat{C} = \arg\max_{C_i} \prod_{i=1}^{n} P(C_i/X) \tag{5.4}$$

$\hat{C}$ is the prediction path loss class and $P(X_i/C)$ is the conditional probability of dataset feature given the class. [50] published an article showing how machine learning techniques such as Deep Neural Network (DNN) reduce complexity and increase performance. In this manuscript, Multilayer Perceptrons Neural Network is introduced in slowromancapii@. Followed by section slowromancapiii@ that shows the dataset. Then, model validation and results in slowromancapiv@. Lastly, a conclusion is shown in section slowromancapv@.

5.2 Multilayer Perceptrons Neural Network

Among many DNN structures, Multilayer Perceptrons (MLP) uses a Feed-forward neural networks (FFNNs) and a back-propagation network to compute the loss and adjust the weight [30], which is suitable for deep learning. MLP forms a fully connected networks where every single node in a single layer is connected to every node in the following layers. The subsequent error is usually obtained by the loss function and optimization methods can be use to minimize the loss such as Adam optimizer. There are multiple of loss functions and cross entropy will be used when relay selection can be initially viewed as a binary classification problem. MLP is actually a multivariate multiple nonlinear regression and collection of neurons that serves as a classification by building decision decision. Multilayer Perceptrons are usually uncorrelated, and a collection of them make up the network that can be less prone to the notorious overfitting. MLP is mathematically mapping in the form of:

$$\Re^n \longrightarrow \Re^m : (y_1, y_2, ..., y_n) \tag{5.5}$$

$$y_n = g_s\left(w_0 + \sum_{i=1}^{n} w_i y_i\right) \tag{5.6}$$

$$y_2 = g_{out}\left(w_{k0}^{(2)} + \sum_{j=1}^{M} w_{k0}^{(2)} \gamma\left(w_{j0}^{(1)} + \sum_{i=1}^{n} w_{ji}^{(1)} y_i\right)\right) \tag{5.7}$$

The above structure can proceed with only two layer, where $y_0 = 1$ as the output of the first layer. g_s is the activation function and here we are using the step function as can be shown as:

$$g(\cdot) : R \rightarrow R \tag{5.8}$$

$$g_s(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{5.9}$$

To accomplish our purpose, Adam optimization algorithm will be adopted which is different from the traditional stochastic gradient descent process to update the weights iterative base in the training data [32] with a learning rate or step size α. Artificial intelligence (AI), particularly machine learning (ML), is widely studied to enable a system to learn of intelligence, predict and make an assessment instead of the needs of humans [33]. Switching the traditional link selection such as the Adaptive selection scheme [100] to machine learning link selection still in its early stage. One of the main issues in current communications is the accuracy of handover, whereas using machine learning techniques could enhance the prediction and reduce the complexity. The new ML methods predict the best link using different mechanisms base on path loss and receive signal strength. [28] showed how to predict the transmitted signals using deep learning techniques [55] used neural networks methods such as learned decisions-based approximate message passing (LDAMP) network to estimate and learn channel state information (CSI) than solving the limited number of frequency chains in cellular systems from training data. [25] used a model-based method using Cramer-Rao lower bound (CRLB) to predict the channel state parameters in the deeper neural network. Moreover, other authors presented in [58] and [2] some communications challenges that reach the complexity level such as atmospheric effects, handover, beam direction, MIMO and it is the time for machine learning to get evolved. Machine learning uses training and testing for letting the machines learn and keep predicting. In the training part, learning from the data while the testing method, a trained model is used for predicting such as the ray selection.

Supervised learning techniques require input, target and training data to create a model that is used for predicting. If a sample space consists of X_i and output label space y_i where i $1, 2, .., N$ then by using a machine learning algorithms $\mathring{A}$, which is a function that map the input values to the labels that helps for future predicting. To measure the quality of the mapping, a loss function is used and see [22] for more details. The classification algorithms that will be used in this journey is Multilayer Perceptrons Neural Network and can be described in the following.

Multilayer Perceptrons Neural Network usually can be used for both classification and regression. When the Multilayer Perceptrons used for the classification, this algorithm works by having binary or multiple classes. However, the regression techniques is usually used for continuous outputs while our goal here is to classify the link strength to binary classes to predict the optimum link propagation. Multilayer Perceptrons follow the form as shown below.

$$y = \Phi(\sum_{i=1}^{n} w_i X_i + b) \tag{5.10}$$

where w is the vector of weights of x vector inputs and b is the error. Φ is the nonlinear activation function. In this work, we proposed six models of Multilayer Perceptrons with different specification as shown below:

- Model 1: One Hidden Layer of 10 Neurons
- Model 2: Two Hidden Layers of 50 and 10 Neurons
- Model 3: Three Hidden Layers of 10, 50 and 10 Neurons
- Model 4: Four Hidden Layers of 10, 50, 50 and 10 Neurons
- Model 5: Five Hidden Layers of 10, 50, 100, 50 and 10 Neurons
- Model 6: Eight Hidden Layers of 10, 50, 100, 100, 50 and 10 Neurons.

- Model 7: Logistic Regression Model
- Model 8: Dummy Classifier Model
- Model 9: Support Vector Machine

Multilayer Perceptrons neural network will be employed to predict the optimum propagated link in our relay selection. Then, compare with other machine learning techniques base on precision, recall, F1 Score, accuracy and support. Results will be explored using simulated data showing the accuracy of applying deep learning techniques and how this algorithm performs well in relay selection. By considering more from wireless communications, the result and compassion of these machine learning models to predict and selection of the best relay link will be shown later in Section IV.

Since both prediction of link performance and classification to select an appropriate link, MLP neural networks appear to fit our purpose due to the capability of predicting the link with low path loss, which allows a reliable handover to meet the need of eMBB and uRLLC. While other ANN structures such as convolution neural networks are for images where there exist 2D or 3D inputs and the RNN is for sequential models like time series, machine translation, language generation. Further investigations to check the fitness of the MLP models compared to other machine learning models will be conducted later in Section IV. Thus, we proved that deep learning technique (MLP) is a capable technique to overcome this wireless communications fading issue using link selection base on MLP technique and performed better than other machine learning techniques.

5.3 Dataset

The dataset of this investigation was generated after some modification using open-source Matlab simulation by New York University [92] [93]. The dataset of the wireless channel is composed of two fragments. The selected model will be trained and validated on the dataset then will be tested using the unseen data. In this work, the train part took 75% of the data set and 25% for tasting the model in our training/testing scenario and others can be found in [65]. Classes data which is zeros and ones are specified base on the path loss strength and other channel states information (CSI) are used to predict the path loss. The measurements are specified based on distance from 1 m to 40 m. That simulation is suitable for frequencies in a range of 500 MHz to 100 GHz, bandwidth up to 800 MHz with different scenarios and environments. As a summary, simulation parameters are listed in Table slowromancapi@, which exhibits that the channel measurement parameters of that data raw that was used for this paper.

Table 5.1: Channel Measurement Parameters

Parameters	**Values**
Distance (m)	1-40
Frequency (GHz)	28
Bandwidth (MHz)	800
TXPower (dBm)	30
Scenario	UMi
Polarization	Co-Pol
TxArrayType	ULA
RxArrayType	ULA
Antena	SISO
Tx/Rx antenna Azimuth and Elevation (red)	10

The dataset used in this work is consisting of channel properties of a communications link such that the information helps the base station to execute supervised classification based on a dataset from prior measurements or simulations.

5.4 Models Validation and Results

To accomplish a broad exploration, Multilayer Perceptrons Neural Network, Logistic Regression, Dummy Classifier, and Support Vector Machine are used to perform the classification techniques. Evaluating the performance of these classification algorithms by confusion matrix which counts the outcomes of the prediction models compared to the training dataset [101]. Moreover, the precision usually shows how often a model makes a positive prediction and recall shows how the model is confident in predicting all positive targets. Accuracy, Precision, Recall, and F1 Score metrics were used to evaluate the machine learning classification algorithms (classifiers). The accuracy is measured by counting the number of true predictions to the total number of predictions. Which is the number of correctly predicted selected links over the total number of links, which tells how the classifier is able not to misclassify a positive path loss (a sample). Precision is the number of true positives (T_p) over the number of true and false positives (F_p). Recall stands for the number of true positives over the number of true positives and false negatives (FN). While F1 Score measures the harmonic mean for both precision and recall, we obtain the following mathematical expressions:

$$\text{Average Precision} = \frac{1}{n}\sum_{i=1}^{N}\frac{Tp}{Tp+Fp} \tag{5.11}$$

$$\text{Total Recal} = \sum_{i=1}^{N}\frac{Tp}{Tp+FN} \tag{5.12}$$

$$F_1\text{Score} = 2 \times \frac{precision \times recall}{precision + recall} \tag{5.13}$$

The interpretations of performance measures that were used to check the process of a model via precision, recall, f1 Score, accuracy and support. From the two above tables interpretation of performance measures of the Multilayer Perceptrons Neural Network algorithms, we can conclude that all of these techniques did a decent job and model 5 has the highest accuracy which consists of five Hidden Layers of 10, 50, 100, 50 and 10 Neurons. Model 5 gained in the best performance in Precision, Recall, and F1 Score among other models.

Table 5.2: Interpretation of Performance Measures

ANN Models	Precision	Recall	F1 Score
Model 1	0.39	0.61	0.47
Model 2	0.88	0.87	0.87
Model 3	0.86	0.86	0.86
Model 4	0.93	0.91	0.92
Model 5	0.98	0.98	0.98
Model 6	0.88	0.87	0.88
Logistic Regression	0.86	0.86	0.86
Dummy Classifier	0.56	0.57	0.57
SVM	0.92	0.93	0.93

Table 5.3: Accuracy Compression of all Models

Models	Accuracy	ROC AUC Score
Model 1	0.623	0.484
Model 2	0.868	0.877
Model 3	0.857	0.842
Model 4	0.925	0.932
Model 5	0.982	0.981
Model 6	0.882	0.866
Logistic Regression	0.882	0.866
Dummy Classifier	0.857	0.848
SVM	0.934	0.973

Thus, it has the best performance in classifying the ray selection, followed by Model 4 and the worst one is model 1 among the ANN model and the dummy classifier compared to all models. The reason for that is some of the features depend on each other such as distance and received power. Electing the number of hidden layers and the number of neurons is still an open research topic where few or more neurons lead to underfitting and overfitting. An assumption from our trail and error trails, we noticed that the number of hidden layers should lower than the number of input by 30%. Model 6, began degrading once the number of hidden layers has reached 70% of the number of the inputs as can be seen in figure 5.4.

Figure 5.2 presents the Receiver Operating Characteristic (ROC) [102] visualizes a classifier's performance that illustrates the wellness of the classification models used in this paper.

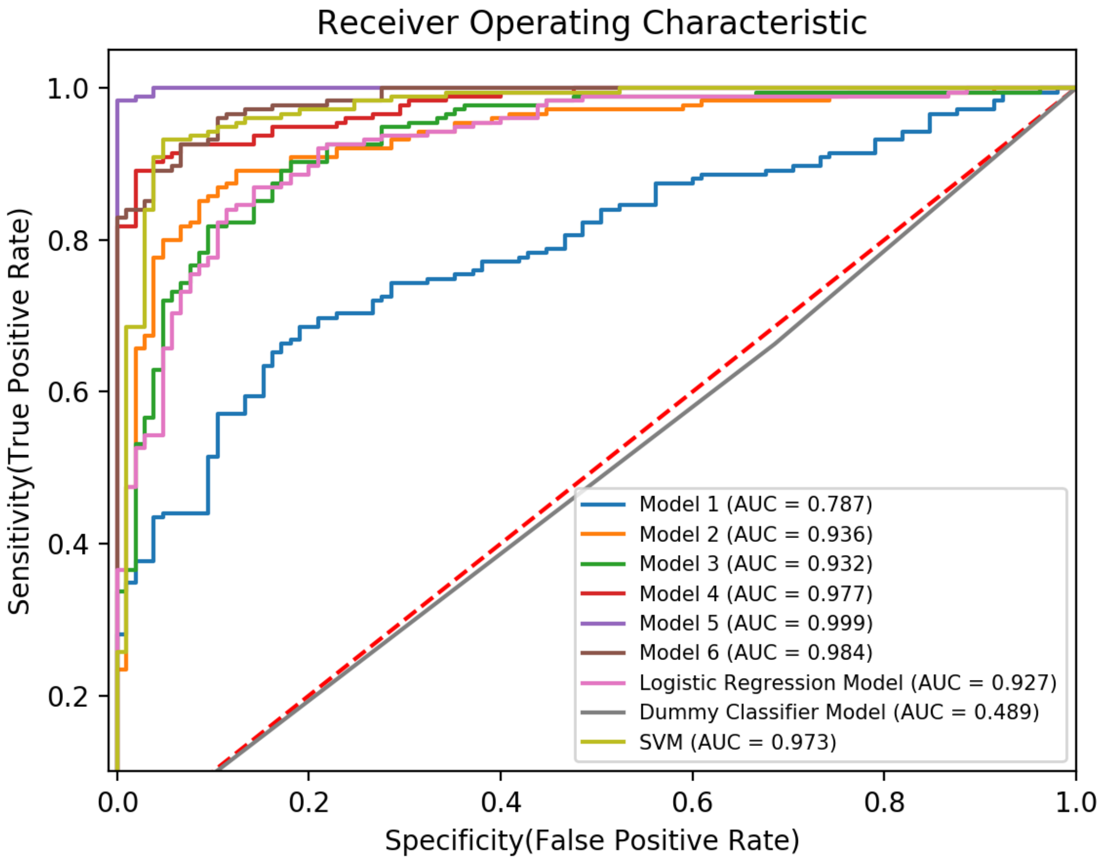

Fig 5.2: ROC Curves of Classification Techniques

The false-positive rate is plotted against the true positive rate and it's very obvious that the model 5 is almost closer to the optimum with 99%. Other models have been compared to our models such support vector machine (SVM) that was presented in [94] that could not perform well in overcoming the relay link selection. The usage of the sigmoid kernel while using the linear kernel poly kernel with degree 4 would improve the performance. Underfitting and overfitting were examined base on checking the accuracy and the ROC score of the training and the testing data which were showing close to each other and performing well. Figure 5.3 illustrates the relay selection precision recall curve where model 5 is performing the best among other models. The precision recall curve shows the relationship between the true positive rate and the positive predictive value for a variety of models. While figure

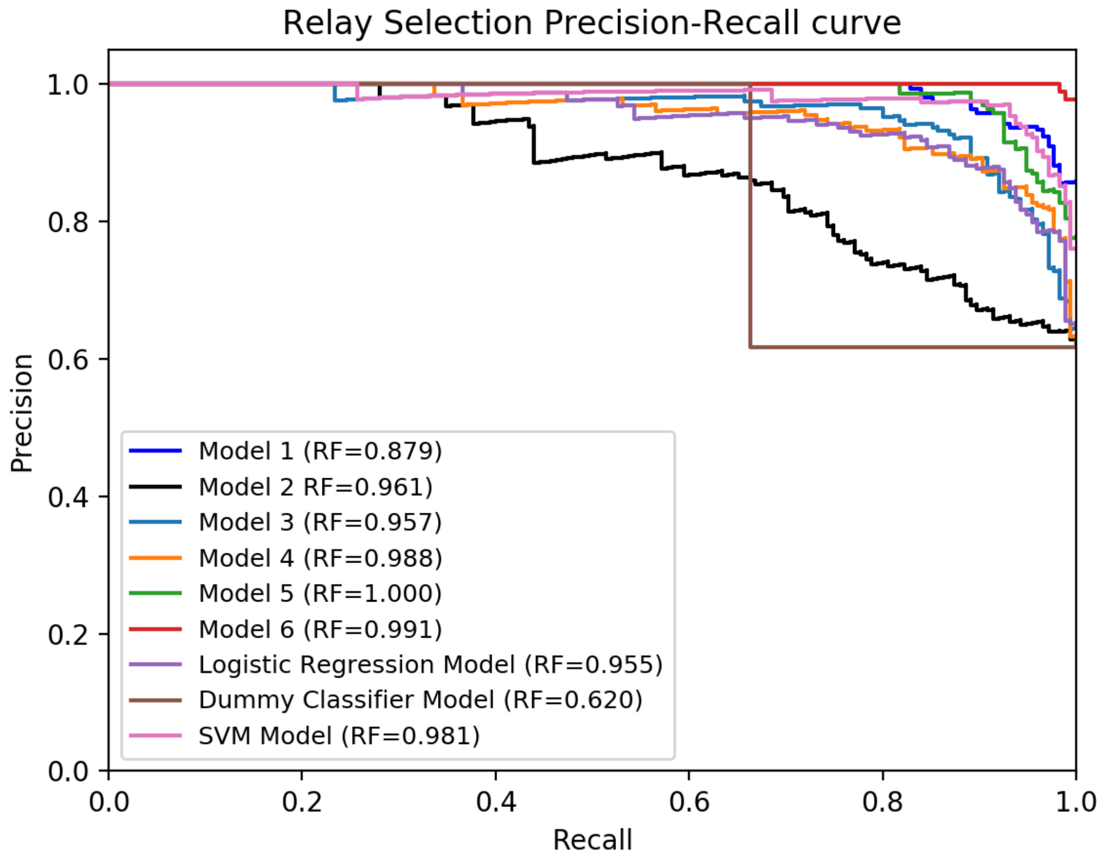

Fig 5.3: Precision and Recall Curve of Classification Techniques

5.4 shows the history loss verse neural iterations when the training data will not improve the performance of the model by at least tolerance (usually assigned 1e-4) or having a constant loss for a multiple of iteration. Moreover, figure 5.4, we notice the losses of models decreased nicely and smoothly except model one due to the adjusted learning rate of this model which was 1e-5 while others are 0.05 and increase that rate will affect the accuracy of the model. Moreover, by looking at model 6, we notice at iteration number 185, it starts increasing which is a sign to stop the model to avoid issues such as overfitting and decreasing the efficiency of the model. Figure 5.4 again confirm the best performance goes to model 5 among other NN models. Simply we conclude this journey from the result section that ANN performs better in term of selecting the optimum link comparing to other machine

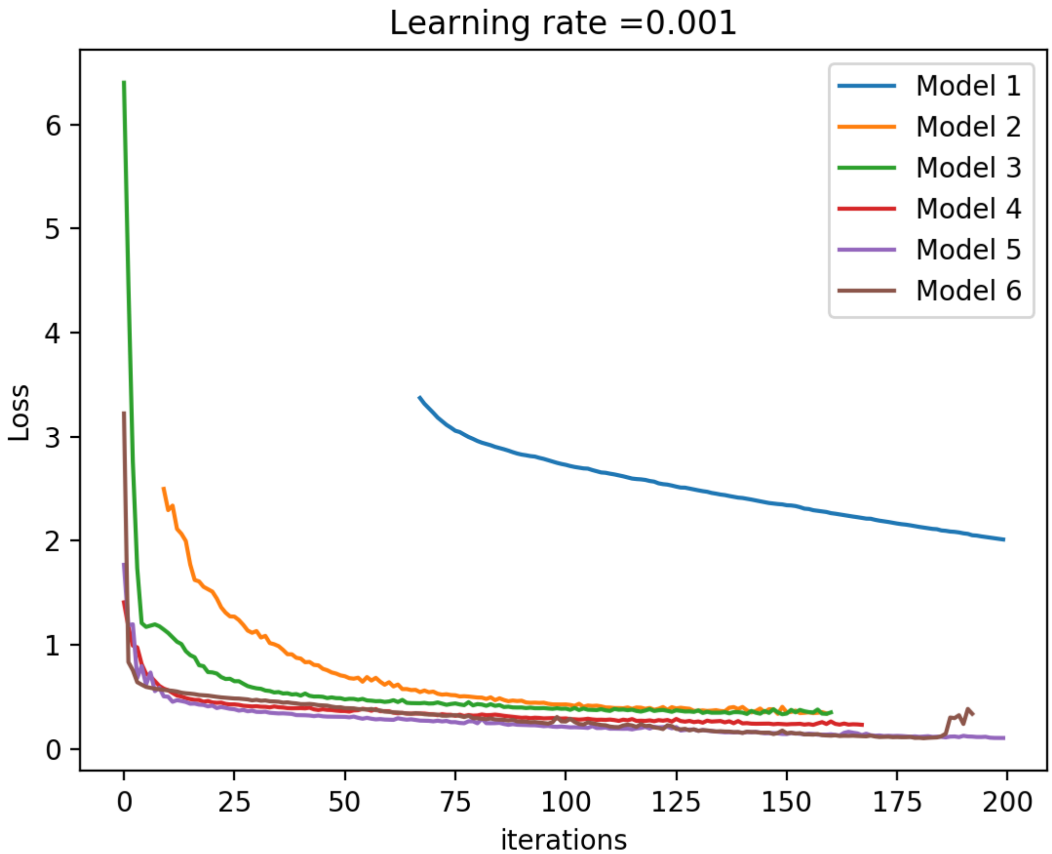

Fig 5.4: Loss VS Iteration of ANN Classification Techniques

learning techniques where the accuracy of selecting the optimum link is 99% which meet third 5G -new radio (NR) requirement (the ultra-reliable and low latency communications). The future works to develop are extraordinarily rich, and powerfully using machine learning algorithms toward solving wireless communications issues to enhance the communications and reduce the complexities for future wireless generations. This approach can be further extended to solve other wireless communications issues such as near-far problem base on more than a binary classification. Base on a certain path loss, power transmission control can be manipulated to achieve the efficiency of a full communication.

5.5 Conclusion

Wireless communications with the new era of 5G and beyond require overcome classical issues in order to meet the 5G pillars that include eMBB, uRLLC and mMTC. One of these critical issues is relay selection to handover with reliability to the strongest link to meet the eMBB and uRLLC. This can be solved by applying machine learning techniques such as Multilayer Perceptrons Neural Network, Logistic regression, Dummy classifier and support vector machine to develop alternative techniques to predict the signal path loss strength that's usually get affected by the wireless channel. Multilayer Perceptrons is a classification technique used in this journey and compared the result of each model by using the interpretation of performance measures such as accuracy, precision, recall, and F1-score and compare with previous studies to end up with a better performance. Other techniques influenced a perfect prediction that confirms the usage of machine learning towards wireless communications.

Chapter 6: Investigation and Prediction of the Wireless Channel Modeling for High Frequency Bands Using Artificial Intelligence Methods

6.1 Abstract

The exploitation of higher MmWave is promising for wireless communication systems. Involving AI and its subcategories of deep learning to beyond 5G (B5G) is to learn from data and make a prediction or a decision other than relying on the classical procedures to enhance the wireless design are auspicious for future wireless systems. The new wireless generation should be proactive and predictive to avoid the previous drawbacks in the exiting wireless generations to meet the 5G target services pillars. One of the aspects of Ultra-Reliable Low Latency Communications (URLLC) is moving the data processing tasks to the cellular base stations. Investigating the wireless channel modeling is an essential task in the wireless communication systems and fuse it with Artificial Intelligence (AI) is our main goal. With the rapid usage of wireless communications devices, base stations are required to execute and make a decision to ensure communication reliability. In this manuscript, an efficient new methodology to predict channel state information based on data-driven to assist the base station to predict the frequency bands and the path loss. Systems that consume different bands such as in telecommunications with uplink and downlink transmission and other IoT devices need an urgent response between devices to alter bands and other channel state information to maintain the requirements of the new radio (NR). Thus, machine learning techniques are needed to learn and assist a base station to fluctuate between different bands based on a data-driven. Then, to testify the proposed idea, we compare the analysis with

other deep learning methods. Furthermore, to validate the proposed models, we applied these techniques to a different case study to ensure the success of the proposed works. To improve the accuracy of supervised learning, we modified the technique by combining an unsupervised algorithm. Eventually, the superiority of AI towards wireless communication has shown great accuracy of 90.24%.

6.2 Introduction

Artificial intelligence deployments in the wireless communication systems rapidly gaining attractiveness nowadays with the usage of data-driven. The justification of that is the usage of the availability of the data solutions to complex and high computational machines for wireless communication problems [103]. Furthermore, The new emerged of the Quantum Computing (QC) technologies have a strong potential to be applied to the future wireless generations with the deployment of AI methods [104]. The potential investigations of AI usages in the field of wireless communications have been increased especially channel modeling. Initiation studies of applying ML to the general field of wireless networks can be found in [105]. Since the AI learning approaches are still considered a black box which it's hard to analyze, [106] recommended studying the theoretical analysis of machine learning to enhance efficiency. Implementing ML techniques to estimate the received signals are complex due to the high computational requirements and [107] proposed an online fully complex extreme learning machine (C-ELM)-based channel estimation to avoid these limitations using a deep neural network. The expectations of deploying ML techniques in the B5G will be a major part of designing the networks in terms of more autonomous, dynamic and self organize. [108] used deep learning methods such as multiple layers perceptrons (MLP) algorithm to predict the wireless channel parameters for dedicated short-range communication and were able to outperform formal empirical models. While others used deep learning methods to perform channel estimation and modulation [109]. Furthermore, other ML algorithms have been used in [110] to assist BS to predict the acceptable received signal in lower MmWave based on classification methodologies. [111] used ML to classify and identify the wireless channel LOS and NLOS. Applying AI techniques towards wireless channel modeling seems successful usage. However, these methods performance relies on data [3] and to avoid the complexity of the wireless channel characterization, a crucial artificial intelligence methodologies have to be involved.

Some facts of the wireless channel modeling such as the higher MmWave bands, the high signal's energy degradation that leads to changing the behavior of the signal with higher MmWave frequencies, the propagated signal becomes more sensitive during the communication that leads to increase the path loss and that will be proved in the following sections. Moreover, penetration of the propagated signal through objects is less in the higher MmWave frequencies [8]. Thus, investigating the wireless channel modeling in higher frequencies is crucial for B5G. Our main goal here is to assist the base stations to predict the channel state information and that includes the evolution to predict the frequency bands and the path loss using some ANN and random forest techniques. The wireless communications scenario will be an urban maco environment and the frequency band will be 1-100 GHz to investigate the bands and another dataset to validate our methodology for varieties of MmWave frequencies. The construction and development of the wireless channel modeling with the usage of the AI is becoming impotent for channel wireless features such path loss which is the reduction of power strength during signal propagation through a channel. The Federal Communications Commission (FCC) has already assigned the MmWave band in the United States as 24, 28, 37, 39, 47 GHz for the 5G and selected bands above frequency 95 GHz for experimental usage to ensure the United States stays at the forefront of wireless innovation. Furthermore, the FCC encouraging industrial and academia to start investigating the higher MmWave frequencies [112]. Researchers have already started performing case studies towards higher bands such as 28 GHz in [113]. Thus, we formulate this work to meet this request by investigating the higher MmWave bands. Additionally, our main goal is to focus on the wireless channel modeling to investigate the path loss and bands in high MmWave frequencies for B5G other than the traditional procedure to pave the road towards future generations as shown in figure 1. Concerning classical works, [114] presented a new measurement campaign to measure high MmWave using classical empirical measurements. The classical procedure can be obtained using empirical or deterministic models. The empirical models are based on applying statistical methods to data that have been collected via the measurement filed. The downside with this approach is, the time consuming that every environment has to be measured to obtain an inaccurate path loss model. While the deterministic models are based on geometrical theories of electromagnetic propagation such as ray-tracing which is accurate but more details of the surrounding environment that require a three dimensional (3D) which is complex [115]. [**?**] proposed a new methodology to predict the path loss of a wireless channel based on a dataset from the different environments where ML algorithms applied to learn the wireless medium pattern and predict the loss with perfect accuracy. This manuscript proposes artificial intelligence techniques to assist the base station to predict frequency bands and path loss. Moreover, we demonstrate and evaluate a new investigation of the high MmWave bands for B5G. Investigating the new bands will be based on two categories, frequency bands, and path loss where artificial intelligence methodology will be applied to predict these factors. Then, this investigation will be compared to the lower MmWave bands. First and second order optimizations methods will be also involved to minimize the cost of the predictions in the base stations. Additionally, during the investigation of the higher MmWave bands, other properties of the wireless channel modeling will be included such as the relation between path loss and delay in lower and higher frequency bands.

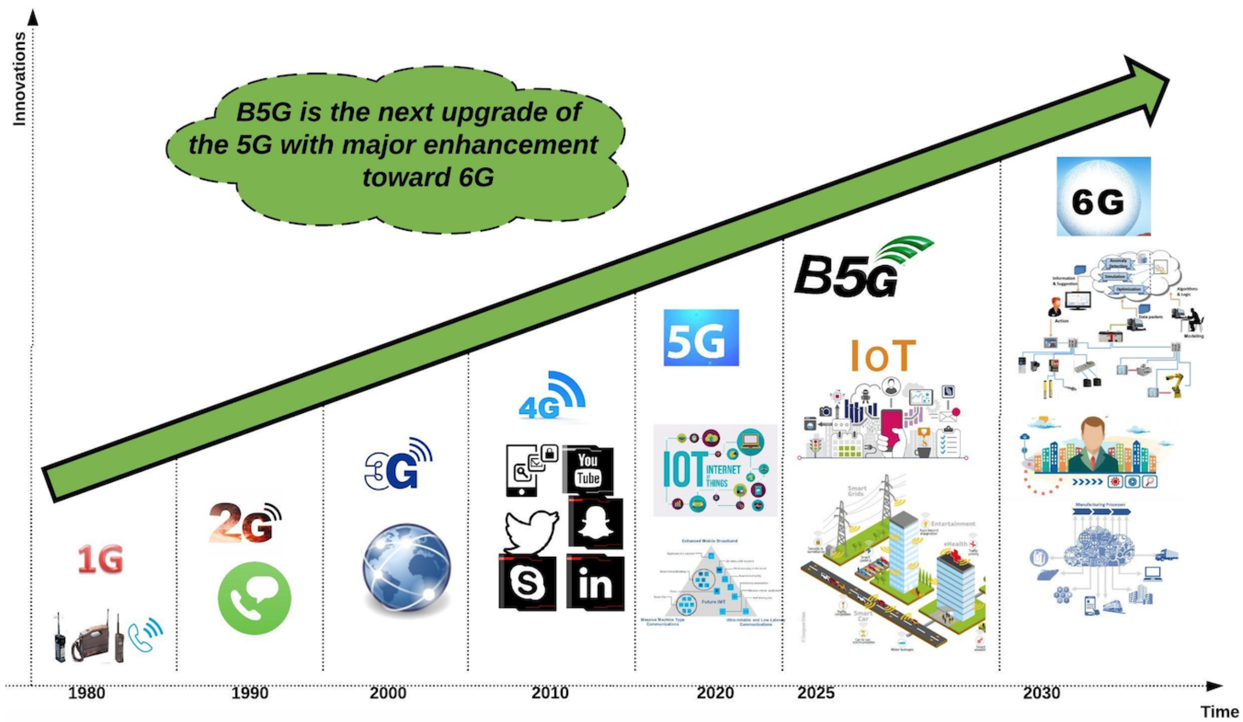

Figure 6.1: Evolution of Cellular Network Generations

The dataset was used to train the neural network's method to predict the CSI features in the higher bands and how these features are preformed compared to the lower bands. Deep learning neural networks methods and other ML techniques such as random forest are used to assist base station to predict the frequency bands, path loss and how higher MmWave band affects the wireless channel modeling. The deep learning technique will be Multiplelayers Perceptrons (MLP) since it works better than other artificial neural networks algorithm with non-stationary data. Forward, backward problem propagation and error correction methods will be used during the testing phase. Moreover, optimization algorithms that deal with optimizing the neural networks will elaborate in the following sections. To enhance the accuracy of the prediction models, the unsupervised learning technique is involved in supervised learning to obtain the CSI features patterns and denominational reduction.

This work is a promising for future wireless systems since each standard has its own regulations. One of these regulations in wide area networks (WAN) such as Cellular networks or Wireless Local Area Network (WLAN), is assigning the frequency bands. For instance, assisting base station to predict the LTE and 5G frequency bands or the 802.11g systems utilizes 2.4 GHz and the 802.11a system uses 5 GHz. The newest Wi-Fi system is called 802.11n and it uses both 2.4 GHz and 5 GHz. The proposed scheme can be further implemented towards and Wireless Sensor Networks (WSN) or even further to the Personal Area Network (PAN) systems. Moreover, other applications of this scheme are telecommunication companies where each cellular company has its own bands. Since the new devices are SIM-less, new devices are capable of predicting the frequency bands that belong to that specific company. Thus, work is a promising for future regulations and standards in future wireless generations. Therefore, by implementing the proposed work, future wireless devices should not be specified for a specific frequency band.

The components of this journey are arranged as follows. Section II system model formulation, while section III presents data-driven and prepossessing. Section IV elaborates on the supervised learning methods for B5G. Results and discussion are presented in section V while the conclusion is in section V.

6.3 System Model Formulation

This manuscript proposes a new scheme to assist the base station to predict a wide range of frequency bands. In a practical scenario, however, the receiver has no access to the actual bands, and suppose to predict it. The proposed system is a single-input single-output (SISO) wireless communication system with a macro urban environment. The system has a radio frequency bandwidth is 800 MHz with a transmitter power of 30 dBm and transmit and receive array type are URA. Due to the weakness of higher MmWave bands, other effects related to weather have been considered such barometric pressure 1013.25 mbar, humidity

50% , the temperature is 20 degrees Celsius and rain rate was assumed to be zero [116]. Other system parameters used in the simulation can be summarized in Table Table I.

Table 6.1: Channel Measurement Parameters

Parameters	**Values**
Distance (m)	1-300
Frequency (GHz)	95
Bandwidth (MHz)	800
TXPower (dBm)	300
Scenario	UMa
Polarization	Co-Pol
TxArrayType	ULA
RxArrayType	ULA
Antenna	SISO
Tx/Rx antenna Azimuth and Elevation (red)	10

Table I, exhibits that the channel measurement parameters of the data raw that was used for this paper. While Python was used to perform the data analysis.

Once the data is generated, by investigating the higher MmWave and that can be seen with driving the Friis theorem. This can be seen using the free space path loss theorem which confirms the increase of the wireless bands, based on the prove in the following subsection and other details can be seen in section V-A.

6.3.1 Theorem

Investigating the effects of higher MmWave can be driven with the usage of the Friis Theorem [117]. By assuming an environment with only a transmitter T_x and a receiver T_r and no obstacles between them to create a free space. With a distance d and a transmitted power of P_t and omnidirectional antennas at both T_x and T_r. Therefore, the received power P_r is calculated using

$$P_r(dBm) = \frac{P_t}{4\pi d^2} G_t \tag{6.1}$$

Where G_t is the gain of the transmit antenna. Assuming the antenna at the receiver side has an effective aperture A_{ER}. Then equation 6.1 becomes as shown below.

$$P_r(dBm) = \frac{P_t}{4\pi d^2} G_t A_{ER} \tag{6.2}$$

The effective aperture of the antenna can be found from:

$$A_{ER} = \frac{\lambda^2}{4\pi} G_r \tag{6.3}$$

$$P_r(dBm) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2} \tag{6.4}$$

Where $\lambda = \frac{c}{f_c}$, λ is the wavelength in meters, c is the speed of light and f is the frequency in GHz [118]. Moreover, Friis theory can lead to the same conclusion that path loss is proportional to the frequency bands.6.4 can be seen as

$$P_r(dBm) = \frac{P_t G_t G_r c^2}{(4\pi f d)^2} \tag{6.5}$$

From equation 6.5, we can infer the frequency bands as shown below.

$$f(MHz) \frac{c^2}{4\pi d} \sqrt{\frac{P_t G_t G_r}{P_r}} \tag{6.6}$$

The above theorem proves driven results in affecting the path loss of the propagated signals that results in increasing the path loss as shown in the below equation.

$$PL(dBm) = 10log_{10}(\frac{16\pi^2 d^n}{\lambda^2}) \tag{6.7}$$

$$P_r(d) = P_t \frac{G_t G_r \lambda}{(2 * \pi)^2 d^n L} \tag{6.8}$$

By understanding the behavior of the path loss with respect to distance and other CSI features with continuous dependent variables leads to the use of the proper procedure of the machine learning categories. Therefore, it's cleared in our case that the machine learning strategy will be a supervised regression problem based on a data driven. One technique to measure the energy loss is by using the path loss model which measures the reduction of the energy during the propagation. By using path loss models, energy loss can be governed. The complexity and the accuracy of the path loss models can be varying with many factors such as environments, interference levels, energy, distance and so on. The wireless bands have a high impact on the amount of loss and coverage following the equation of $d = \frac{\lambda}{c}$ that is different from the lower bands. The channel state information (CSI) features that are used in our system are frequency bands (GHz), T-R separation distance (m), received power (dBm), phase (rad), azimuth AoD (degree), elevation AoD (degree), azimuth AoA (degree), elevation AoA (degree), path Loss (dB) and RMS delay spread (ns). While frequency band (GHz) and path Loss (dB) will be used as a target where other CSI will be used to feed the neural networks to assist the base station (Bs) to predict the bands or the path loss.

In the macro environment, user equipment (UEs) is assumed to be non-stationary and placed uniformly distributed and measurement parameters can be seen in table 6.1. In our model, we applying deep learning neural networks techniques such as Multilayer Perceptrons (MLP) as supervised learning to predict the dependent variable. Having a dataset that consists of the dependent variable y_i and channel state information x_i, features of the high MmWave bands can be predicted. In this manuscript, we will focus on predicting the path loss and the frequency bands based on previous data $p(y/x)$. The data are splatted into training $D = (x_i, y_i)_{i=1}^{N}$ and testing $T = (x_i, y_i)_{i=1}^{N}$ where $D \cap T = \phi$.

Since our investigation is dealing with the side effects of the higher wireless bands and we're trying to obtain the relation between a response variable y and other CSI variable. Thus, our focus will be on applying AI methodologies to formulate the work. AI algorithms learn from the data and make predictions based on mapping the input features to output variables through data learning patterns. The machine learning algorithms learn the pattern and assist base stations to make a decision. Artificial Neural Networks (ANN) are widely used nowadays to learn the complex pattern of the wireless channel to avoid complex and unreliable mathematical formulations. Since the presented data is nonlinear and multivariable characteristics, ANN could be involved to predict the frequency bands and path loss to assist the base stations as alternative model structures for received power. The ANN's structure consists of at least three main layers, input (i), hidden (j_n) and output (k) layers and for simplification, we assume the system only consists of only one hidden layer. Each layer composes of one or more number of neuron N_{li}, where l to represent which layer and $i \in \mathbf{N} = (1, 2, \cdots, k)$ for identify the specific neurons which considered the main component and the processing unit. N_k neurons in layer j_n feeds from N_{k-1} neurons in the previous layer by a weight vector. Input of CSI $X = (x_1, x_2, \cdots, x_n)$ are feed to networks and then multiply by the weight vector $W = (w_1, w_2, \cdots, w_n)$ and with addition of bias variable b_i then summed at the hidden layer. Then, activation function $f(\cdot)$ is used for every node in the hidden layers to produce an output where more details of the activation functions will be elaborated later in this manuscript. The summation of the hidden and output layer of the neural networks denoted by s_j and s_k respectively and can be formulated as follows.

$$s_j = \sum_{i=0}^{k} w_{ij} \cdot x_i + b_i \tag{6.9}$$

$$s_k = \sum_{i=0}^{k} w_{jk} \cdot x_k + b_i \tag{6.10}$$

$$y = f(\sum_{i=1}^{n} w_i x_i + b_i) \tag{6.11}$$

Figure 6.2 is a simplified systems model, where the training data samples $D = (x_i, y_i)_{i=1}^{N}$ are feed to the system. forward, backward and optimization techniques are implemented in this system and more details will in section ??. Deep learning algorithms will be compared with machine learning methods and then optimization techniques will be applied to reduce the loss of the prediction.

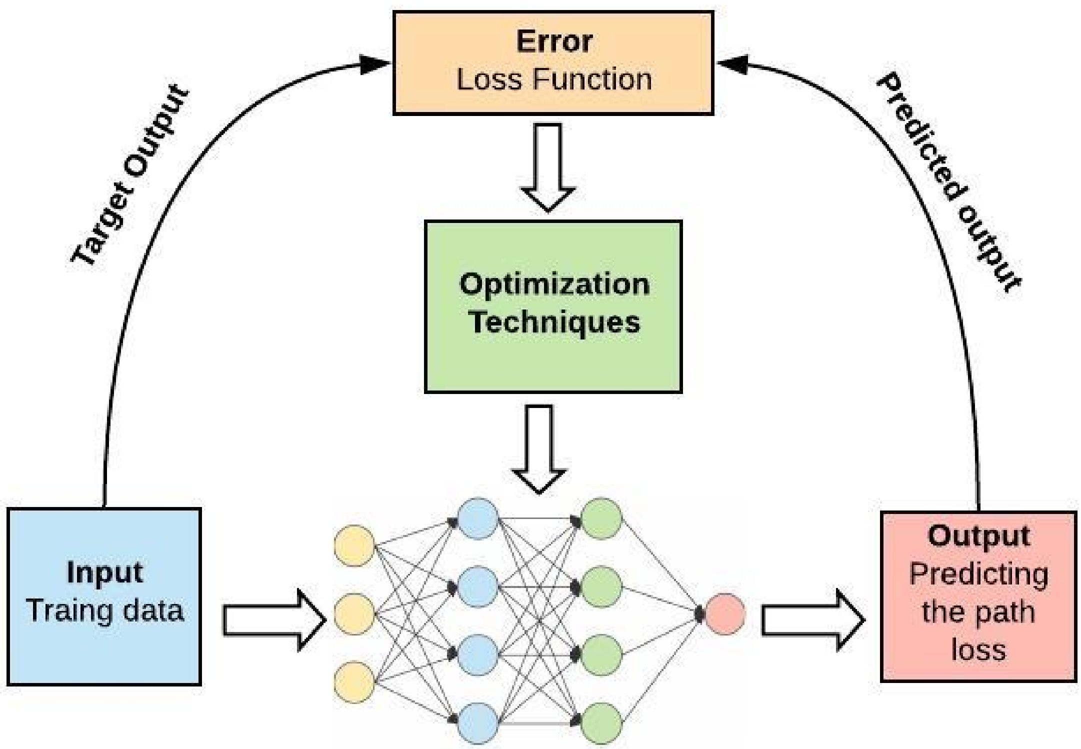

Fig 6.2: Visualization of the System Model Using ANN Technique

6.4 Data Driven Modeling and Prepossessing

Data-driven is considered as a new approach for next-generation generation models [119]. Data in the proposed cases were produced using the NYUSIM simulator which was based on an extensive measurement campaign for varieties of MmWave bands and different scenarios environments [120]. Prior to analyzing the data, data cleansing is considered the backbone of data analytics. Data cleansing is the process of identifying, removing and recovering the nonrelevant data. After that data has been produced, several data prepossessing have to be performed such as the data cleansing. Data cleansing is one of the main parts of machine learning that perform a significant part to enhance the accuracy of the model. Data cleansing is just the process of removing erroneous or unwanted observations from the data then replacing them with samples based on several concepts such as the averaging. Furthermore, managing unwanted outliers and then verifying the data. Removing unwanted observations including detecting redundant or irrelevant samples from the dataset [121]. The dataset of CSI features has been divided into training, testing and validating with percentages of 70%, 15% , and 15% respectively.

6.4.1 Channel Impulse Response

The wireless channel can be modeled deferentially depending on the communication environments. Linear time-variant (LTV) channel model is a popular wireless channel model $h(t)$ and in theory, if the received signal $r(t)$ with an input signal $s(t)$

$$r(t) = x(t) * h(t) + n(t) \tag{6.12}$$

Where the $n(t)$ is noise that affect propagated signal and the convolution is given by integration with a channel delay τ as shown below.

$$r(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau + n(\tau) \tag{6.13}$$

$$h(t) = \sum_{i=0}^{N-1} a_i e^{j\theta_i} \delta(t - \tau_i) \tag{6.14}$$

$$APDP(\tau) = \frac{1}{N} \sum_{i=0}^{N} |h_i(\tau)|^2 \tag{6.15}$$

Equation 6.14 is the channel impulse response of the wireless channel where N is the multipath components (MPCs) and a_i is the random amplitude of MPC which can be characterized statistically using Rayleigh or Rician distribution based on the communication environment. Then, the power delay profile can be obtained by taking the squared absolute value of the impulse response. While equation 6.15 is the power delay profile (PDP) that involves removing the effects of the small scale fading by averaging the power delay profile. PDP usually gives the intensity of the received signal of a multipath channel as a function of time delay. [52].

$$\begin{aligned} PDP &= |h(t)|^2 && (6.16) \\ &= \sum_{i=0}^{N-1} (a_k e^{j\theta_i})^2 \delta(t - \tau_k) && (6.17) \\ & && (6.18) \end{aligned}$$

6.5 Supervised Learning Methods for B5G

In the following subsections, multiple machine learning and deep learning techniques will be used to investigate the wireless channel of the high MmWave band for beyond 5G. Furthermore, applying these techniques to the base station to predict the frequency bands and the path loss. The Artificial Intelligence techniques establish the mapping relationship between the dependent and independent variables. Where in our case, the dependent variable is the path loss and bands while other channel state information is the independent variable.

6.5.1 Random Forest

Random forests is a machine learning method that is based on ensemble methods and can be applied to regression and classification tasks. Random forests use Bootstrap Aggregation to divide the data into multiple bags and each bag pass and trains the data to the decision tree. Then, multiple decision trees are combined to determine the output instead of relying on an individual decision. Thus, for predicting the frequency bands to assist the base station, the final prediction value is obtained by averaging the whole predictions from all single tress as shown in equation 6.19 [3]. To enhance the random forest regression prediction, Principal Component Analysis (PCA) technique which is unsupervised learning was implemented to find more patterns to enhance the prediction.

$$\hat{y_b} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_t(x) \tag{6.19}$$

Another way to investigate the signal attenuation is through checking the path loss for different MmWave bands from 1 GHz through 100 GHz. As can be seen in the below figure that in specific bands, the path loss is considered and performed better than other bands.

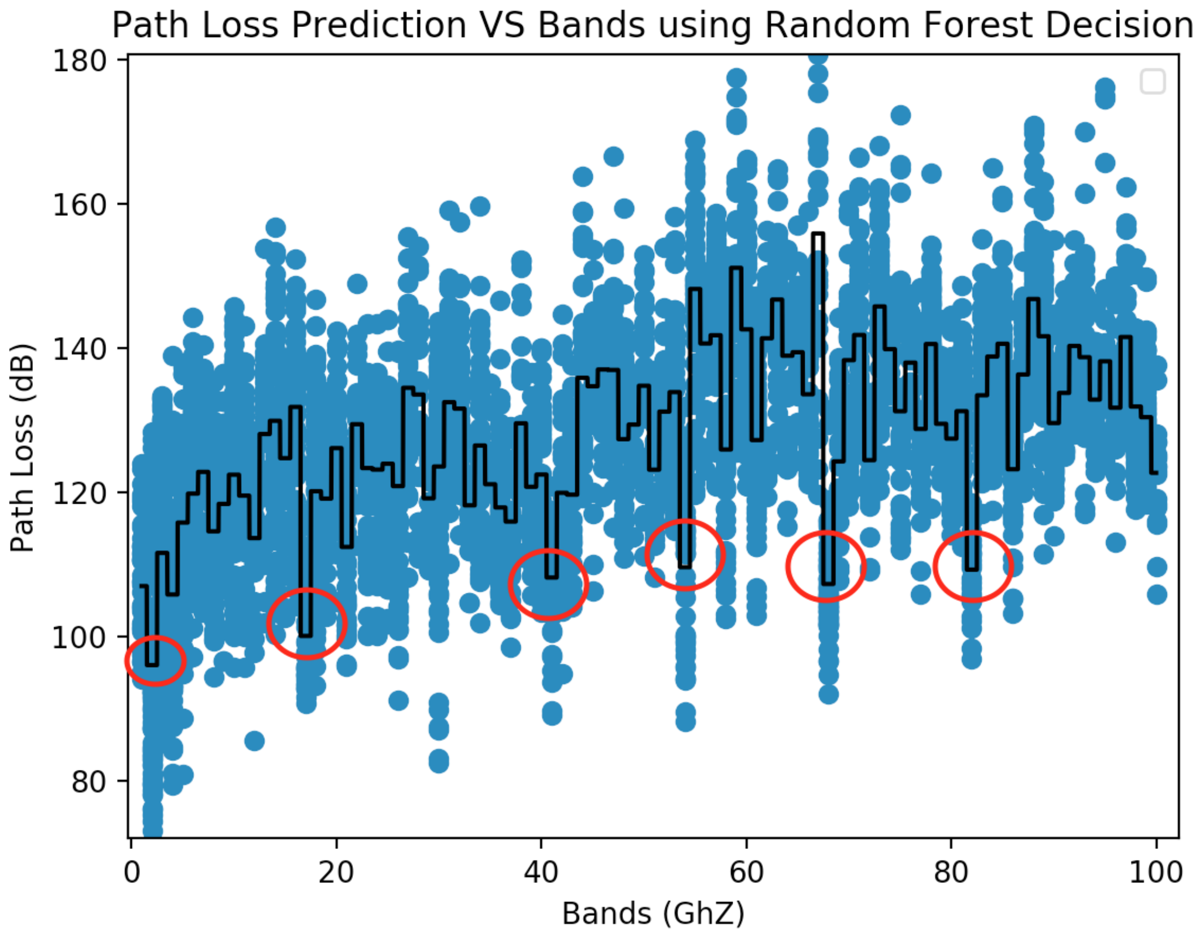

Fig 6.3: Visualization of the High Loss with Different MmWave Frequencies

Based on figure 6.3, path loss (dBm) VS Bands (GHz) visualization explaining the low loss through different frequency bands. We can identify the perfect path loss in band such as 2.4, 17, 41, 54, 68 and 82 GHz. The number of estimators in this kind of ensemble method is the number of trees to be used in the forest. In our case, we used 100 estimators where each tree or estimator makes a prediction and with other tree's prediction can be averaged to obtain the final prediction.

6.5.2 Multiplelayers Perceptrons

Multiplelayers Perceptrons (MLP) method is a branch of artificial neural networks (ANN) for the purpose of computing and prediction. MLP combines neurons in different layers to solve complex problems. MLP will be used as a regression technique to investigate the high MmWave where Feed-forward neural networks and a back-propagation computation will be used to assist the prediction in the wireless base station. This section describes the propose neural networks technique to investigate wireless channel modeling in terms of path loss and bands. The output of the neural network in equation 6.10 will be fed back to the network to compute the backpropagation to reduce the loss using equation the following equation.

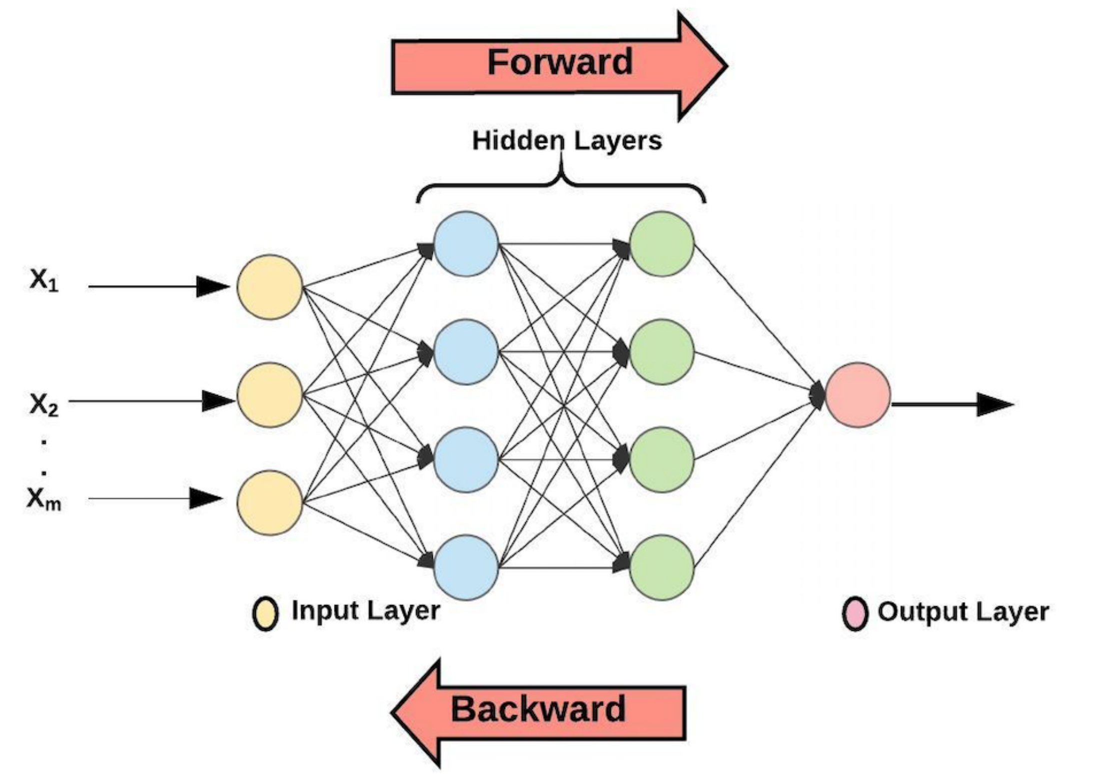

Figure 6.4: An Explaining the MLP Forward and Backward Algorithms

$$\mathcal{L}_\epsilon(y, f(x,w)) = \frac{1}{2}\sum_{i=0}^{k}(y_{k_i} - \hat{y_{k_i}})^2 \tag{6.20}$$

Where $\mathcal{L}$ is the loss, y is the actual dependent variable and $\hat{y_k}$ is predicted value. The predicted value of the neural networks can be seen as follows.

$$\hat{y_k} = \sum_{i=0}^{k}(w_{jk} \cdot S\sum_{i=0}^{k}(w_{ij}x_i + b_j) + b_k)) \tag{6.21}$$

Then, optimization methods are used to reduce the error to reach the minima of loss. For simplification, the bias variable will set or initialize from now to zero. Since the loss have been obtained in equation 6.20, back propagation feeds the parameters values back to the network to reduce the error then after some iterations, the parameter values are ready to be used to form the path loss model of 5G advance. Back propagation calculations shall be started from the output layer toward the first hidden layer and by taking the derivative with some derivative properties such as the chain rule as shown below.

$$\begin{aligned}
\frac{\partial\mathcal{L}(w)}{\partial(w_{jk})} &= \frac{\partial\frac{1}{2}\sum_{i=0}^{k}(y-\hat{y}^2)}{\partial w_{jk}} \\
&= (y-\hat{y})\frac{\partial(y-\hat{y})}{\partial w_{jk}} \\
&= (y-\hat{y})\frac{\partial\hat{y}}{\partial w_{jk}} \\
&= (y-\hat{y})\frac{\partial\hat{y}}{\partial x_k}\cdot\frac{\partial x_k}{\partial w_{jk}} \\
&= (y-\hat{y})\cdot A_{x_k}\cdot(1-A_{x_k})\cdot x_j \\
&= Error\cdot A_{x_k}\cdot(1-A_{x_k})\cdot x_j
\end{aligned} \tag{6.22}$$

Where A_{x_k} is the derivation of the activation function and x_k,x_k,x_k are unit value in input, hidden and output layer respectively.

Then the value of the derivation in the above equations is called the backpropagated error from the output layer which passes the output value to the hidden layer. This procedure has to come through every layer in the networks until it reaches the input layer as shown in the below derivation to obtain the final backpropagation value. Then the procedure of updating the weight and bias parameters begins and proceeds a new forward propagation to the networks.

$$
\begin{aligned}
\frac{\partial \mathcal{L}(w)}{\partial (w_{ij})} &= \frac{\partial \frac{1}{2} \sum_{i=0}^{k} (y - \hat{y}^2)}{\partial w_{ij}} \\
&= \sum_{i=0}^{k} (y - \hat{y}) \frac{\partial \hat{y}}{\partial w_{ij}} \\
&= \sum_{i=0}^{k} (y - \hat{y}) \frac{\partial \hat{y}}{\partial x_k} \cdot \frac{\partial x_k}{\partial x_j} \cdot \frac{\partial x_j}{\partial x_i} \cdot \frac{\partial x_i}{\partial w_{ij}} \\
&= \sum_{i=0}^{k} (y - \hat{y}) \alpha(x_k)(1 - \alpha(x_k) w_{jk} \alpha(x_j)(1 - \alpha(xj)) x_i
\end{aligned}
\tag{6.23}
$$

During the MLP regression procedure, optimization algorithms have been used to optimize the square loss. Gradients are usually used to find the minimum loss using a vector that reaches to the optimum loss. We started with Broyden–Fletcher–Goldfarb–Shanno algorithm (LBFGS) optimizer which is a family member of quasi newton methods. Then, started comparing the results with stochastic gradient descent (SGD). Moreover, another stochastic gradient-based optimizer has been used called "adam" or some times called adaptive moment estimation. Adam is an extension of stochastic gradient descent to update neural network weights iterative based in training dataset [122]. The reason most of this work focused on adam optimizer is our data is non-stationary and it does not conserve a specific learning rate for every single weight update [123].

6.5.3 Neural Network Optimization

Optimization algorithms are often used in neural networks such as Gradient Descent (GD) which is an iterative technique to obtain the values of the parameters of a function to reduce the cost function. Optimization algorithms are methods used to reduce the error function $E(x)$ that depends on the internal learning variables that are considered undependable to compute the response variable y. Thus, the main purpose of the optimization is to estimate the error gradient of the current state, then updates the weights and the bias of the model using the backpropagation of error algorithm. Other stochastic methods can be used to update the neural networks but there must be a trade-off between the precision of the updated parameters and the running time to update. Optimization methods are used to minimize the loss function $L(w_i, b_i)$ by fining the optimum wight w_i and b_i bias variables.

Optimization can be categorized into two major codes based on first or second-order optimization to obtain the information of the slope as a vector to identify the minima direction. The first-order optimization algorithms are used to minimize or maximize the loss function using a derivative. An example of the first-order methods is Gradient Descent (GD) where the derivative is used to identify whether the loss function is increasing or decreasing toward the minima. While second-order optimization algorithms are considered as optimizing the optimizer based on the second-order derivative to check how the gradient varies which usually involves a high computational cost. To avoid this dilemma, techniques such as LBFGS is applied with limited memory usage. Other SGD methods such as Adaptive Moment Estimation (Adam), Momentum, and Root Mean Square Propagation (RMSProp) are used as second-order optimization methods. Gradient Descent is considered as the main optimizer to optimize an intelligent system. GD is used to reduce the loss by updating the model's parameters to reach the converge limit. Once the gradient vector $\nabla J(W)$ is calculated, the new parameter can be obtained by multiplying a learning rate η by the gradient parameter and then subtracting them from the current weight parameter. Where the subtracting value is used to move toward the opposite direction since the gradient is considered as a vector that points to the direction of the increase of the loss function. Where the newly updated parameters are feedback to the network to reduce the cost. Where the cost can be calculated using the mean squared error. The new update parameter W_{t_1}

$$W_{t_1} = W_t - \eta \cdot \nabla J(W) \tag{6.24}$$

$$W_{t+1} = W_t - \eta \cdot J(W) \tag{6.25}$$

$$W_{t+1} = W_t - \eta \cdot \frac{\partial J(W)}{\partial w_t}$$

Where W_{t+1} is a new weight that we assume will feed to the networks to reduce the loss and W_t is the previous weight parameter. Usually, a random weight is generated and if the new weight $W_t + \partial W$. Initialization of the parameters can be in the form of ones, zeros, normal distribution, truncated normal distribution, etc. A different version of gradient descent recommends using where the difference between these versions is based on the number of data samples that feed to the network for each iteration.

- Gradient descent (GD) which based on applying the gradient algorithm to every single observation in the training set.
- Stochastic gradient descent (SGD) is the opposite of the SG method. SGD introduces a random sample of the data on its iteration. The con of this method is the slowness.
- Batch gradient descent is based on feeding all data to the network at the same time. The limitation of this method is the high risk while the positive is the processing speed.
- Mini-batch gradient descent: is based on feeding the networks with N random of a group of samples to overcome the cons on the SGD such as the acceleration processes.

During our compression between the optimization methods, we notice that ADAM preformed better with large amount of data while LBFGS is works for small data since it converge faster and may preform with alike results to other stochastic optimizes.

6.5.4 The Neural Network Structure

The activation function can have a general form where W is the weight vector and b is the bias variable. Neural networks that don't use an activation function will end up with a linear output. Where non-linear functions are preferred to learn more complicated data patter that linear functions aren't able to do.

$$h = f(W^T X + b) \tag{6.26}$$

The activation function is used in the forward and backward propagation process in the network to compute the summation of the error to show the loss of the models. The number of iterations is set to 200 and the number of hidden layers and neurons in the proposed models can be shown in Table 6.2.

Table 6.2: MLP Parameters

MLP Models	**Hidden layers**	**Learning Rate**
Model 1	3	e^{-1}
Model 2	3	e^{-2}
Model 3	5	e^{-3}
Model 4	6	e^{-4}
Model 5	6	e^{-10}
Model 6	7	e^{-20}

6.6 Results

In this section, the results of investigating the wireless channel modeling in the high MmWave bands is presented. We started with constructing a logarithmic path loss model using different channel state information. Furthermore, we were able to assist base stations to predict the bands and path loss using AI techniques. To elaborate the results more, we used multiple evaluation matrices to evaluate the accuracy of the AI models such as the average root mean square error (RMSE) defined as:

$$\bar{\mu} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\bar{\mu_j}} \tag{6.27}$$

$$\bar{\mu_j} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - h(\mathbf{x}_i, \mathbf{w}))^2} \tag{6.28}$$

Where μ_j in the above equation can be obtained for the test set, given by and k is the number of data samples in the test set. Additionally, to evaluate the proposed models, regression evaluation metrics are preform. Mean squared error (MSE) is used to average the sum squared difference between the predicted label $\hat{y}$ and the actual label y. Mean absolute error (MAE) is measured by taking the average absolute difference between the predicted label and the actual label. Where low values are targeted to reach perfect prediction without errors in both MSE and MAE. R-Square (R^2) which is used to evaluate the models and how close it fits the regression line by measuring the sum square error to the total square error [124]. Other statistical metrics can obtain from the below properties

$$MSE\ \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{6.29}$$

Where y_i and $\hat{y}_i$ are the actual and predicted values for multiple of purpose in this manuscript such as predicting the frequency bands and the path loss of a high MmWave in maco urban environment.

6.6.1 Bands

With the new wireless generations, the wireless bands are becoming more widely and it's essential to allow the base station to predict the bands. In this section, the base stations can predict the bands based on machine learning algorithms. With the usage of wireless channel information, the bands can be predicted. Machine learning methods such as algorithms and deep learning techniques are implemented to infer the frequency bands. Figure 5, shows that as the bands increase, the path loss increase as well.

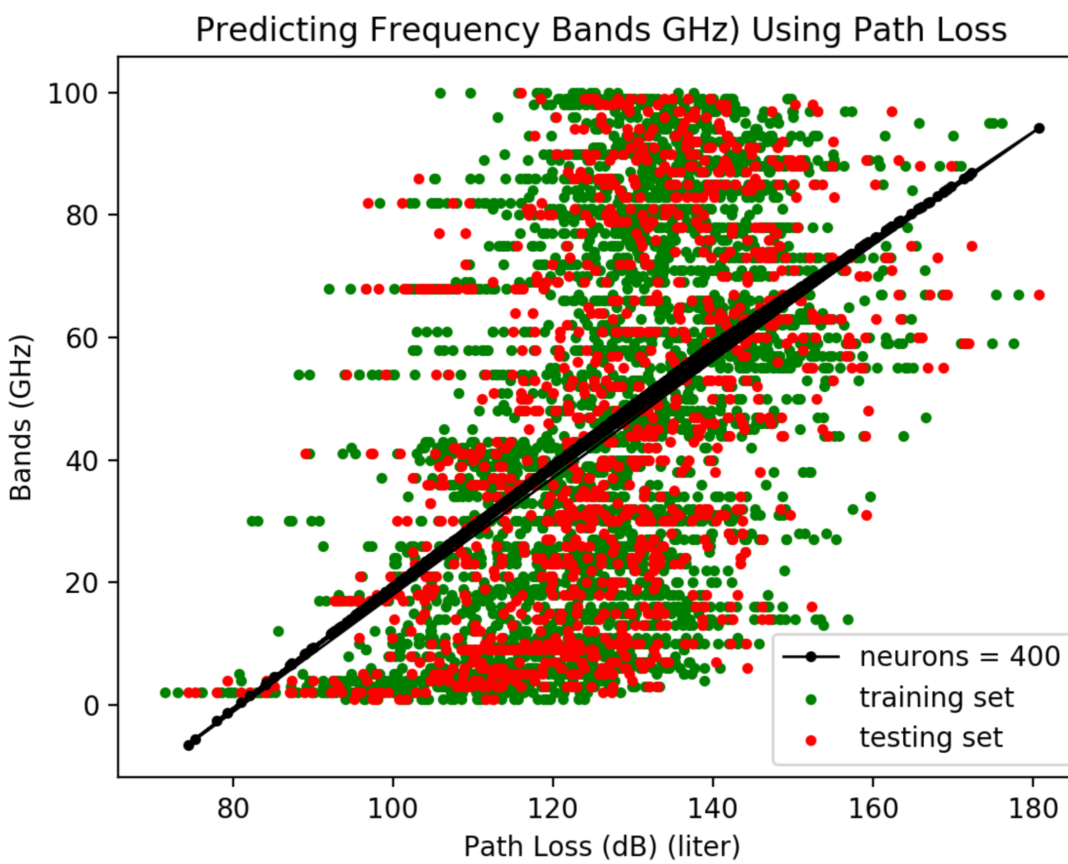

Figure 6.5: The Relation of MmWave Bands with Path Loss

Figure 7.2 shows the prediction of frequency bands with using only a CSI feature where the purpose of this figure is to show the proportion relation between path loss and frequency to meet equation 6.7.

Table 6.3: Frequency Bands Prediction Evaluation Metrics

AI Models	R-Square	MAE	MSE	RMSE
Model 1	65.6	0.73	12.2.9	17.3
Model 2	63.3	13.3	322.2	17.9
Model 3	65.0	12.9	306.8	17.5
Model 4	67.1	12.2	287.7	17.0
Model 5	65.0	12.7	307.0	17.5
Model 6	66.2	12.6	296.4	17.2
Model 7	73.0	11.7	238.9	15.45
Random Forests Regress (RFR)	83.7	10.5	142.5	4.99
RFR and PCA	90.23	9.8	132.6	3.89

Table 6.3 shows the evaluations and the capability of base stations to predict the wireless frequency bands that allocated from 1 to 100 GHz. Where the highest prediction was the modified schemes of the combination of RFR and PCA techniques as marginal of supervised and unsupervised learning. That combination of random forest regression and principal component analysis (PCA) results in a higher prediction than using the random forest by its self due to the capability of PCA to find the pattern of the CSI features. Moreover, to illustrate how the loss function of predicting the bands can be seen in the below figure. Where Sven models have been implemented with different learning rate to predict the frequency bands of a wireless systems. Moreover, number of iteration reached to 300 which indicate the lowest loss error obtained with this procedure.

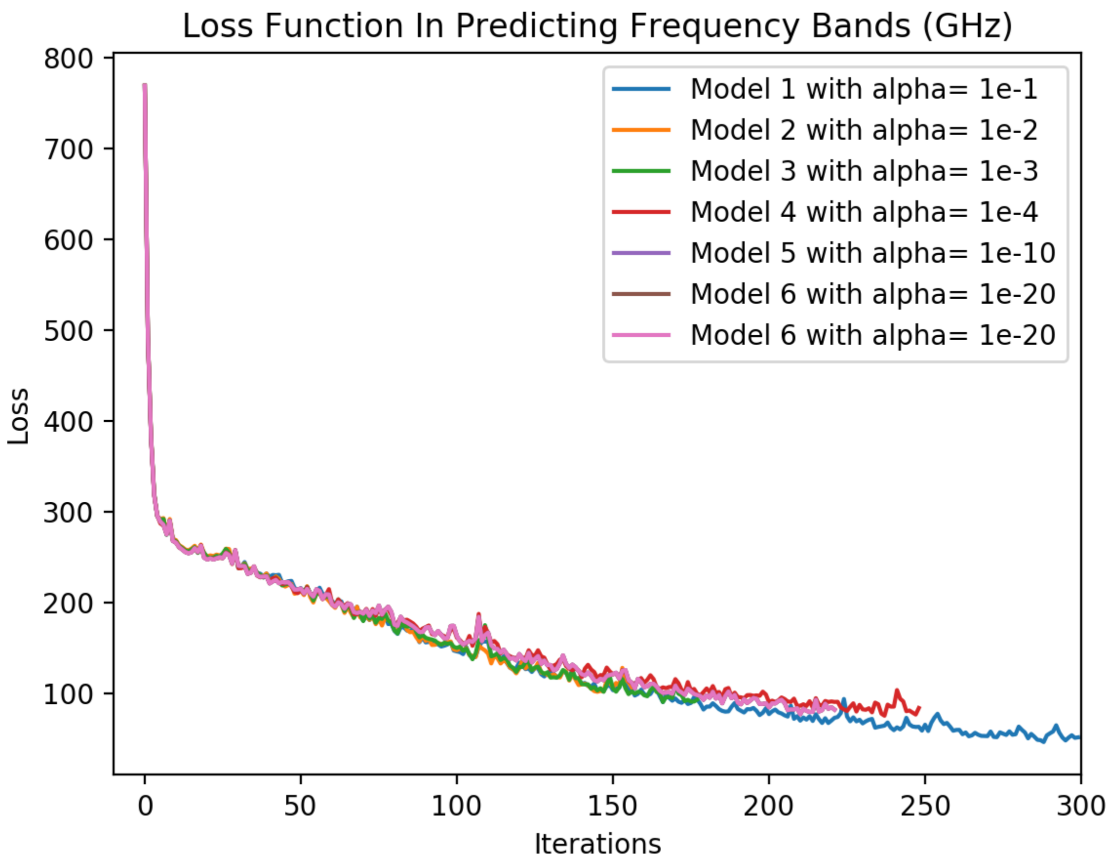

Figure 6.6: The Loss of the Proposed Models

6.6.2 Path Loss

Path loss usually obtained from the measured PDP by calculating the received power using the integration of the area under PDP. Since the transmitted power and the antenna gains are known, path loss can be obtained from the wireless channel empirically. In this manuscript, we proposing a new methodology to allow base stations to predict the path loss based on leaning using AI algorithms. In this section, investigation of the path loss in the higher frequency bands compared to the lower frequencies. A comparison between a 95 frequency to lower 28 GHz in with alike environments will be shown in this section.

Table 6.4: Path Loss Evaluation Metrics

AI Models	R-Square	MAE	MSE	RMSE
Model 1	73.63	5.94	60.81	7.79
Model 2	75.02	5.91	57.61	7.59
Model 3	73.00	6.24	62.77	7.93
Model 4	76.48	5.77	54.24	7.36
Random Forests Regress	89.18	3.34	24.93	4.99

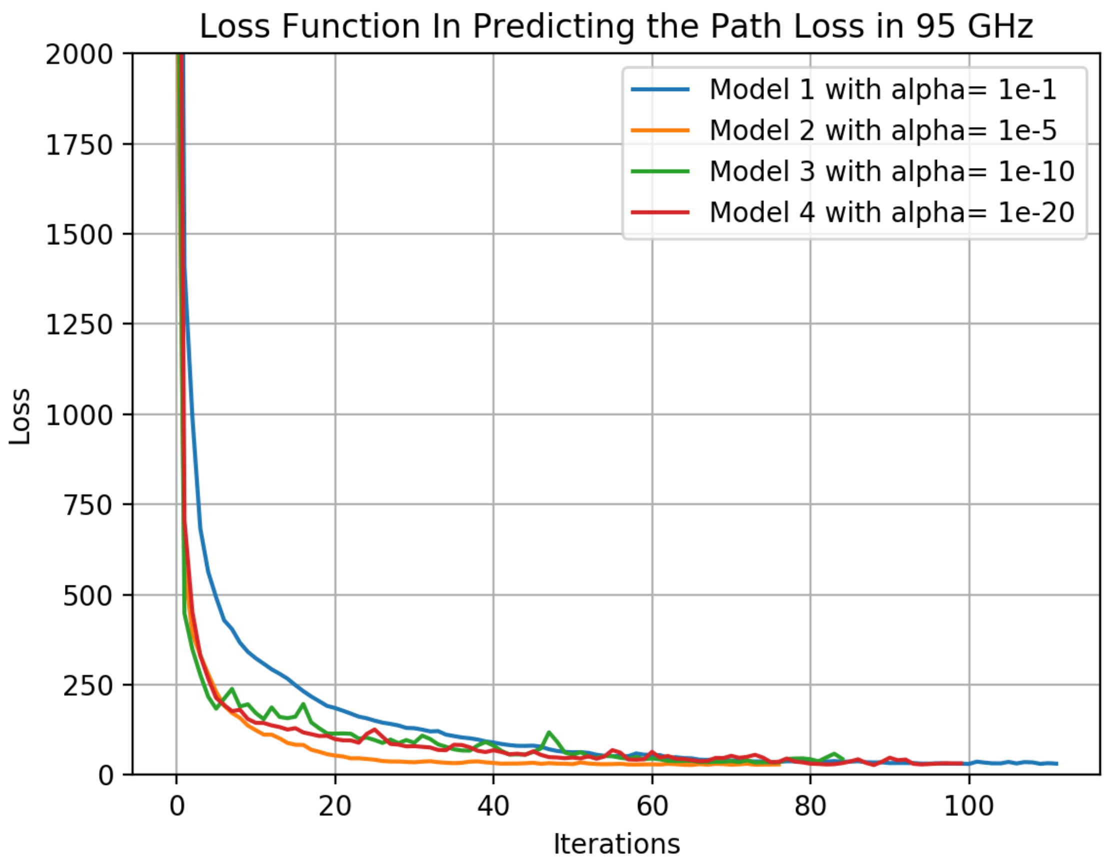

Figure 6.7: The Loss of the Deep Learning Models with Different Learning Rate

The loss function in equation 6.20 can be represented in figure 6.7 that shows how the loss of predicting the path loss decayed with the number of iterations. In terms of how the loss curve is used to investigate the path loss, three models of MLP algorithms have been used as seen below. The models are the same but with different learning rates, we can see that the low learning rate, the improvement can be seen as a linear. While with a lower learning rate, the sharper the curve became and decaying faster. While a very high learning rate can cause overfitting in the model.

Table 6.1 contains path loss predicting evaluation in the higher MmWave Bands. Where the random forests obtain the highest accuracy with 90.24% while other MLP models have lower accuracy and this is due to random forests uses ensemble technique that reduces the variance error. Moreover, you may notice that the normalization on that table was neglected due to the low difference between normalization process results and the regular procedure.

Moreover, the investigation of the effect of path loss with propagation delay in different bands. The received propagated signal appears as the sum of taps with different delay due to multipath. That stretching signals are usually casing temporal dispersion and can be characterized using the RMS delay spread of the channel impulse response $h(t)$. We confirm that the path loss increase as the frequency band increase and that can be seen from the below figures. Where figures 6.8 and 6.9 show how path loss is changing with higher frequency bands.

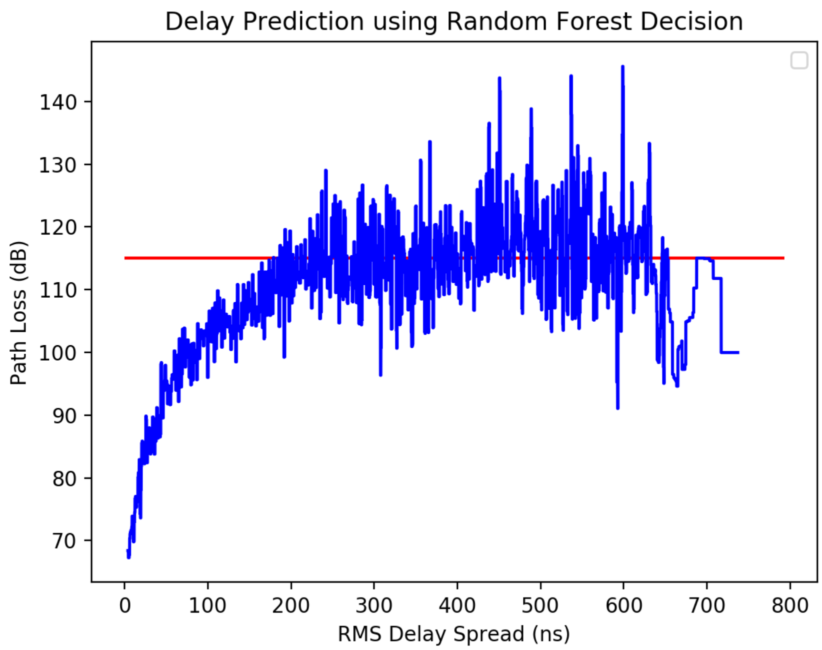

Figure 6.8: Case I: Frequency Band = 28 GHz

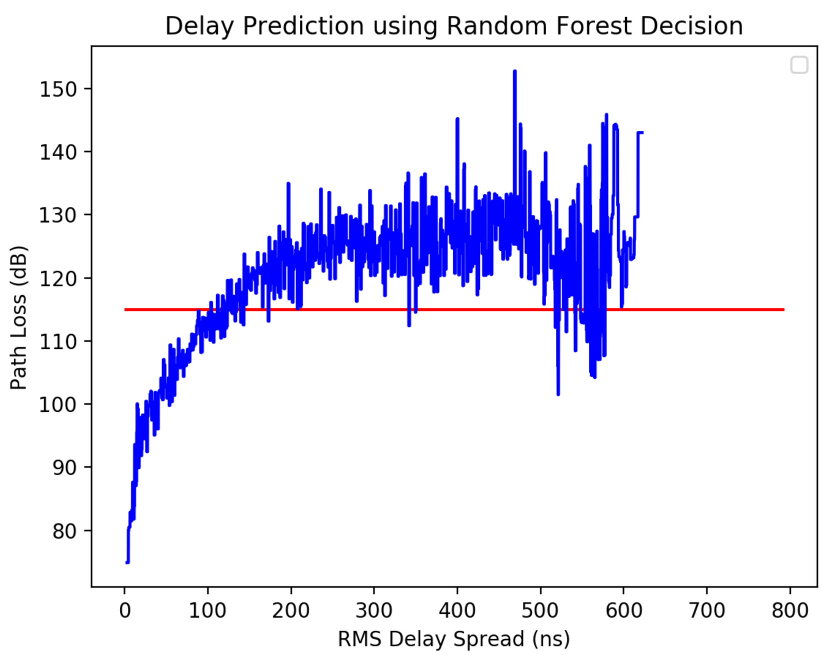

Figure 6.9: Case II: Frequency Band = 95 GHz

6.6.3 Validation Results

Table 6.5: Frequency Bands Prediction Evaluations Using an Alternative Case study

AI Models	R-Square A	R-Square B
Model 1	65.6	70.6
Model 2	63.3	70.9
Model 3	65.0	63.9
Model 4	67.1	63.8
Model 5	65.0	77.3
Model 6	66.2	76.9
Model 7	73.0	79.4
Random Forests Regress	90.23	94.9

To validate the proposed methodology to assist base station to be capable to predict the frequency bands, we generated another dataset. The second environment is a micro urban environments where the generated bands were from 10 to 100 GHz with an increment of 5. The results conclusion is almost match. Table 6.5 shows the R-square only of both cases for the same applied models. The above procedures were implemented on a cellular systems and following the same methodology can be applied to other systems such wireless local area network (WLAN), wireless sensor networks (WSN) and etc.

6.7 Conclusion

The high MmWave bands for beyond 5G have been examined in this manuscript with the methods of applying artificial intelligence techniques. The superiority of AI towards wireless communication is proposed in these manuscripts based on data-driven. By utilizing wireless channel modeling data and other deep learning and machine learning approaches, we were able to assist the base stations to predict the frequency bands and the path loss. The deep leaning MLP technique was associated with the performance of other machine learning algorithms such as random forests regressor to predict path loss and frequency bands. Optimization methods were used to assess and update the internal variables such as the weight Where various types of optimization algorithms can improve the system performance. The random forest as supervised learning was modified by an unsupervised PCA algorithm to assist the base station to enhance the prediction of the MmWave bands.

Chapter 7: Future Works and Conclusion

7.1 Challenges and Open Issues in Wireless Channel Modeling Using AI

Applying machine learning to enhance or to enable channel modeling is still in the infant stage. We summarize a list, surely not an exhaustive list, in the following:

- Channel Modeling under overfitting and underfitting with massive data that leads to estimation erroneous [81] [125].
- Analysis massive data of channel modeling with advanced machine learning algorithms [55].
- Applying machine learning techniques to MmWave channel challenges such as Blockages [12].
- Interference such as co-channel interference (CCL) and inter-symbol interference (ISI) [11] [12] [31].
- Securing the data during channel estimation in V2V.
- Power control and handover where decision supposed to be built based on previous states or results and making a decision using reinforcement in the current stage such to execute a handover [126].
- Accurate channel models with dynamic behavior with mobility [127].
- Switching between dual bands system (MmWave $\Longleftrightarrow$ Microwave) [128].

- New standardization supposes to include machine learning and other data formats.

Solving these issues requires more investigation form researchers in academia and industries. Resolving this issue is essential for future smart cities and artificial intelligence life.

7.2 Future Works

Getting involved in a measurement campaign will provide us with confirming results. We will investigate the feature of wireless channels in other MmWave frequency bands will be measured under very close conditions and will be compared to other measurements. Moreover, these measured data will lead us to build a new channel model free to design optimized wireless communication systems.

7.3 Conclusion

The propagation channel of communication systems fundamentals based on channel modeling particularly for a new technology era such as MmWave. The irregularity and complexity of the wireless channel lead to solid ways to achieve accurate models where more trails are always required to accomplish the precise results mainly with new technologies.

Fading is identified as the time variation of the received multipath components MPCs power that propagates through the medium channel where it can be categorized into two parts as Small Scale Fading and Large Scale Fading. Propagation in wireless channel models can result in either large or small scale fading base on the variations of the signal where large scale fading can be either path loss or shadowing due to shadowing by objects that's larger than the carrier wave. Whereas, small scale fading is due to amplitude variations due to multipath time delay or Doppler spread. In this dissertation, we focused more on the large scale fading.

Switching the traditional channel modeling to machine learning channel modeling still in its early stage. One of the main issues in current communication is to accurately predict the channel parameters, whereas using machine learning techniques could enhance the prediction and reduce the complexity.

Artificial Intelligence (AI) emerges to revolutionize system design for new radio 5G. The subcategories of AI involve machine learning, deep learning techniques such as Multilayer Perceptrons Neural Network have been used to investigate the wireless channel modeling to predict channel state information and assist base stations to select the optimum propagated signals. Forevermore, investigating the higher MmWave bans was involved in this dissertation and predicting the frequency bands was performed.

This dissertation focused on predicting channel state information based on data-driven and elaborates on how to overcome some wireless issues in the new era 5G. To conclude, based on our investigation in this dissertation, we confirm that applying Artificial intelligence towards wireless channel modeling is a promising technique and should be implemented in current and future wireless communication systems.

References

[1] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 2595–2621, Fourthquarter 2018.

[2] S. M. Aldossari and K.-C. Chen, "Machine learning for wireless communication channel modeling: An overview," *Wireless Personal Communications*, vol. 106, no. 1, pp. 41–70, 2019.

[3] Y. Zhang, J. Wen, G. Yang, Z. He, and J. Wang, "Path loss prediction based on machine learning: Principle, method, and data expansion," *Applied Sciences*, vol. 9, no. 9, p. 1908, 2019.

[4] S. Aldossari and K. C. Chen, "Predicting the path loss of wireless channel models using machine learning techniques in mmwave urban communications," *The 22ND International Symposium on Wireless Personal Multimedia Communications (WPMC)*, vol. pp. 1–5, Nov 2019.

[5] S. S. Das and R. Prasad, *Evolution of Air Interface Towards 5G: Radio Access Technology and Performance Analysis*. River Publishers Series in Communications, 2019.

[6] Z. Zhong, B. Ai, G. Zhu, H. Wu, L. Xiong, F. Wang *et al.*, "Advances in high-speed rail technology," 2018.

[7] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5g) wireless networks—with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, Dec 2017.

[8] F. Hu, *Vehicle-to-vehicle and vehicle-to-infrastructure communications : a technical approach*. Taylor and Francis Group: San Francisco, CA,, November 30, 2005.

[9] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, 2005, vol. 95.

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[11] T. S. Rappaport, S. DiPierro, and R. Akturan, "Analysis and simulation of interference to vehicle-equipped digital receivers from cellular mobile terminals operating in adjacent frequencies," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1664–1676, May 2011.

[12] K. Sultan, H. Ali, and Z. Zhang, "Big data perspective and challenges in next generation networks," *Future Internet*, vol. 10, no. 7, p. 56, 2018.

[13] E. Feukeu, S. Ngwira, and T. Zuva, "Doppler shift signature for bpsk in a vehicular network: Ieee 802.11 p," in *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2015, pp. 1744–1749.

[14] M. P. et al, "Deliverable d2.1 measurement campaigns and initial channel models for preferred suitable frequency ranges,," University Bristol, Tech. Rep., 2016.

[15] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, and J. Jarvelainen, "Propagation path loss models for 5g urban micro-and macro-cellular scenarios," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2016, pp. 1–6.

[16] N. DOCOMO, "5g channel model for bands up to100 ghz."

[17] G. R. MacCartney and T. S. Rappaport, "Rural macrocell path loss models for millimeter wave wireless communications," *IEEE Journal on selected areas in communications*, vol. 35, no. 7, pp. 1663–1677, 2017.

[18] G. R. MacCartney Jr, S. Sun, T. S. Rappaport, Y. Xing, H. Yan, J. Koka, R. Wang, and D. Yu, "Millimeter wave wireless communications: New results for rural connectivity," in *Proceedings of the 5th workshop on all things cellular: operations, applications and challenges*, 2016, pp. 31–36.

[19] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.

[20] S. Sun, T. S. Rappaport, T. A. Thomas, A. Ghosh, H. C. Nguyen, I. Z. Kovács, I. Rodriguez, O. Koymen, and A. Partyka, "Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5g wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 2843–2860, 2016.

[21] J. Riihijarvi and P. Mahonen, "Machine learning for performance prediction in mobile cellular networks," *IEEE Computational Intelligence Magazine*, vol. 13, no. 1, pp. 51–60, Feb 2018.

[22] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2018.

[23] T. Lu, J. Sun, K. Wu, and Z. Yang, “High-speed channel modeling with machine learning methods for signal integrity analysis,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 6, pp. 1957–1964, 2018.

[24] Y. Zhang, J. Wen, G. Yang, Z. He, and X. Luo, “Air-to-air path loss prediction based on machine learning methods in urban environments,” *Wireless Communications and Mobile Computing*, vol. 2018, pp. 8 489 326:1–8 489 326:9, 2018.

[25] Z. Jiang, Z. He, S. Chen, A. F. Molisch, S. Zhou, and Z. Niu, “Inferring remote channel state information: Cramér-rae lower bound and deep learning implementation,” in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–7.

[26] D. A. Freedman, *Statistical models: theory and practice*. cambridge university press, 2009.

[27] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, “Thirty years of machine learning: The road to pareto-optimal wireless networks,” *IEEE Communications Surveys & Tutorials*, 2020.

[28] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, “Support vector clustering,” *J. Mach. Learn. Res.*, vol. 2, pp. 125–137, Mar. 2002. [Online]. Available: http://dl.acm.org/citation.cfm?id=944790.944807

[29] J. H. Lee, J. Kim, B. Kim, D. Yoon, and J. W. Choi, “Robust automatic modulation classification technique for fading channels via deep neural network,” *Entropy*, vol. 19, no. 9, p. 454, 2017.

[30] S. Haykin, *Neural Networks and Learning Machines, 3/E*. Pearson Education India, 2010.

[31] H. Ye, G. Y. Li, and B. Juang, “Power of deep learning for channel estimation and signal detection in ofdm systems,” *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, Feb 2018.

[32] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014.

[33] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, “Machine learning paradigms for next-generation wireless networks,” *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[34] K. Sultan and H. Ali, “Where big data meets 5g?” in *Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing*, ser. ICC ’17. New York, NY, USA: ACM, 2017, pp. 103:1–103:4. [Online]. Available: http://doi.acm.org/10.1145/3018896.3025151

[35] D. Neumann, T. Wiese, and W. Utschick, "Learning the mmse channel estimator," *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2905–2917, June 2018.

[36] Z. Yang, Y. Zhang, J. Yu, J. Cai, and J. Luo, "End-to-end multi-modal multi-task vehicle control for self-driving cars with visual perceptions," in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018, pp. 2289–2294.

[37] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric mimo channel data including path powers," in *IEEE Vehicular Technology Conference*. IEEE, 2006, pp. 1–5.

[38] Y. Li, J. Zhang, and Z. Ma, "Clustering in wireless propagation channel with a statistics-based framework," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.

[39] P. Dayan and Y. Niv, "Reinforcement learning: the good, the bad and the ugly," *Current opinion in neurobiology*, vol. 18, no. 2, pp. 185–196, 2008.

[40] M. Ghavamzadeh, S. Mannor, J. Pineau, A. Tamar *et al.*, "Bayesian reinforcement learning: A survey," *Foundations and Trends® in Machine Learning*, vol. 8, no. 5-6, pp. 359–483, 2015.

[41] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.

[42] S. Bi, R. Zhang, Z. Ding, and S. Cui, "Wireless communications in the era of big data," *IEEE communications magazine*, vol. 53, no. 10, pp. 190–199, 2015.

[43] S. Kumar and R. Miikkulainen, "Dual reinforcement q-routing: An on-line adaptive routing algorithm," in *Proceedings of the artificial neural networks in engineering Conference*, 1997, pp. 231–238.

[44]

[45] J. Zhang, "The interdisciplinary research of big data and wireless channel: A cluster-nuclei based channel model," *China communications*, vol. 13, no. 2, pp. 14–26, 2016.

[46] L. Ericsson, "More than 50 billion connected devices," *White Paper*, vol. 14, no. 1, p. 124, 2011.

[47] C.-X. Wang, J. Bian, J. Sun, W. Zhang, and M. Zhang, "A survey of 5g channel measurements and models," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3142–3168, 2018.

[48] A. Mahmood, K. Shi, S. Khatoon, and M. Xiao, “Data mining techniques for wireless sensor networks: A survey,” *International Journal of Distributed Sensor Networks*, vol. 9, no. 7, p. 406316, 2013.

[49] X. Ma, J. Zhang, Y. Zhang, and Z. Ma, “Data scheme-based wireless channel modeling method: motivation, principle and performance,” *Journal of Communications and Information Networks*, vol. 2, no. 3, pp. 41–51, 2017.

[50] G. Hackeling, *Mastering Machine Learning With scikit-learn*, 2014.

[51] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler, “Challenges in visual data analysis,” in *Tenth International Conference on Information Visualisation (IV’06)*. IEEE, 2006, pp. 9–16.

[52] V. Kristem, C. U. Bas, R. Wang, and A. F. Molisch, “Outdoor wideband channel measurements and modeling in the 3–18 ghz band,” *IEEE Transactions on Wireless Communications*, vol. 17, pp. 4620–4633, 2018.

[53] H. Ye, G. Y. Li, B.-H. F. Juang, and K. Sivanesan, “Channel agnostic end-to-end learning based communication systems with conditional gan,” in *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2018, pp. 1–5.

[54] T. O’Shea, K. Karra, and T. C. Clancy, “Learning approximate neural estimators for wireless channel state information,” in *2017 IEEE 27th international workshop on machine learning for signal processing (MLSP)*. IEEE, 2017, pp. 1–7.

[55] H. He, C.-K. Wen, S. Jin, and G. Y. Li, “Deep learning-based channel estimation for beamspace mmwave massive mimo systems,” *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.

[56] Z. Jiang, Z. He, S. Chen, A. F. Molisch, S. Zhou, and Z. Niu, “Inferring remote channel state information: Cramér-rae lower bound and deep learning implementation,” in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–7.

[57] T. O’Shea and J. Hoydis, “An introduction to deep learning for the physical layer,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[58] T. J. O’Shea, T. Erpek, and T. C. Clancy, “Deep learning based mimo communications,” *arXiv preprint arXiv:1707.07980*, 2017.

[59] V. Raj and S. Kalyani, “Backpropagating through the air: Deep learning at physical layer without channel models,” *IEEE Communications Letters*, vol. 22, no. 11, pp. 2278–2281, 2018.

[60] S. Navabi, C. Wang, O. Y. Bursalioglu, and H. Papadopoulos, "Predicting wireless channel features using neural networks," in *2018 IEEE international conference on communications (ICC)*. IEEE, 2018, pp. 1–6.

[61] D. Karagiannis and A. Argyriou, "Jamming attack detection in a pair of rf communicating vehicles using unsupervised machine learning," *Vehicular Communications*, vol. 13, pp. 56–63, 2018.

[62] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and doa estimation based massive mimo system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8549–8560, 2018.

[63] M. Ibnkahla, "Applications of neural networks to digital communications–a survey," *Signal processing*, vol. 80, no. 7, pp. 1185–1215, 2000.

[64] B. J. Cavalcanti, G. A. Cavalcante, L. M. d. Mendonça, G. M. Cantanhede, M. M. de Oliveira, and A. G. D'Assunção, "A hybrid path loss prediction model based on artificial neural networks using empirical models for lte and lte-a at 800 mhz and 2600 mhz," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 16, no. 3, pp. 708–722, 2017.

[65] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Improved adaptive signal power loss prediction using combined vector statistics based smoothing and neural network approach," *Progress In Electromagnetics Research*, vol. 82, pp. 155–169, 2018.

[66] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5663–5678, 2018.

[67] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[68] D. Kim *et al.*, "A mimo mobile-to-mobile channel model: Part i-the reference model," in *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1. IEEE, 2005, pp. 573–578.

[69] L.-C. Wang, W.-C. Liu, and Y.-H. Cheng, "Statistical analysis of a mobile-to-mobile rician fading channel model," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 32–38, 2008.

[70] L. Cheng, D. D. Stancil, and F. Bai, "A roadside scattering model for the vehicle-to-vehicle communication channel," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 449–459, 2013.

[71] A. G. Zajic and G. L. Stuber, "Three-dimensional modeling, simulation, and capacity analysis of space–time correlated mobile-to-mobile channels," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2042–2054, 2008.

[72] X. Cheng, C.-X. Wang, D. I. Laurenson, S. Salous, and A. V. Vasilakos, "An adaptive geometry-based stochastic model for non-isotropic mimo mobile-to-mobile channels," *IEEE transactions on wireless communications*, vol. 8, no. 9, pp. 4824–4835, 2009.

[73] D. W. Matolak and Q. Wu, "Vehicle-to-vehicle channels: Are we done yet?" in *2009 IEEE Globecom Workshops*. IEEE, 2009, pp. 1–6.

[74] S. Sun, G. R. MacCartney, and T. S. Rappaport, "Millimeter-wave distance-dependent large-scale propagation measurements and path loss models for outdoor and indoor 5g systems," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2016, pp. 1–5.

[75] Y. Wang, M. Narasimha, and R. W. Heath, "Mmwave beam prediction with situational awareness: A machine learning approach," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, pp. 1–5.

[76] D. Burghal, R. Wang, and A. F. Molisch, "Band assignment in dual band systems: A learning-based approach," in *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*. IEEE, 2018, pp. 7–13.

[77] E. Ostlin, H.-J. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2735–2747, 2010.

[78] X. Zhao, C. Hou, and Q. Wang, "A new svm-based modeling method of cabin path loss prediction," *International Journal of Antennas and Propagation*, vol. 2013, 2013.

[79] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, *Automatic clustering of MIMO channel parameters using the multi-path component distance measure.* WMPC, 2005.

[80] J. Ko, Y.-J. Cho, S. Hur, T. Kim, J. Park, A. F. Molisch, K. Haneda, M. Peter, D.-J. Park, and D.-H. Cho, "Millimeter-wave channel measurements and analysis for statistical spatial channel model in in-building and urban environments at 28 ghz," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 5853–5868, 2017.

[81] A. R. Guraliuc, P. Barsocchi, F. Potortì, and P. Nepa, "Limb movements classification using wearable wireless transceivers," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 474–480, 2011.

[82] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "Improving clustering performance using multipath component distance," *Electronics Letters*, vol. 42, no. 1, pp. 33–35, 2006.

[83] D.-J. Kim, Y.-W. Park, and D.-J. Park, "A novel validity index for determination of the optimal number of clusters," *IEICE Transactions on Information and Systems*, vol. 84, no. 2, pp. 281–285, 2001.

[84] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The cost259 directional channel model-part i: Overview and methodology," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3421–3433, 2006.

[85] R. H. Qingbi Zheng and C. Huang, "A tracking-based multipath components clustering algorithm," *2018 2nd URSI Atlantic Radio Science Meeting*, 28 May – 1 June 2018.

[86] C. Huang, R. He, Z. Zhong, Y.-A. Geng, Q. Li, and Z. Zhong, "A novel tracking-based multipath component clustering algorithm," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2679–2683, 2017.

[87] J. Li, B. Ai, R. He, M. Yang, Q. Wang, B. Zhang, and Z. Zhong, "Cluster-based 3-d channel modeling for massive mimo in subway station environment," *IEEE Access*, vol. 6, pp. 6257–6272, 2017.

[88] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 12, pp. 1650–1654, 2002.

[89] M. Piacentini and F. Rinaldi, "Path loss prediction in urban environment using learning machines and dimensionality reduction techniques," *Computational Management Science*, vol. 8, no. 4, pp. 371–385, 2011.

[90] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of research and development*, vol. 3, no. 3, pp. 210–229, 1959.

[91] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2012, vol. 34.

[92] S. Sun, G. R. MacCartney, and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5g wireless communications," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–7.

[93]

[94] K. Sankhe, C. Pradhan, S. Kumar, and G. Ramamurthy, "Machine learning based cooperative relay selection in virtual mimo," *arXiv preprint arXiv:1506.01910*, 2015.

[95] K. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative communications and networking*. Cambridge university press, 2009.

[96] J. Qiao, L. X. Cai, X. S. Shen, and J. W. Mark, "Enabling multi-hop concurrent transmissions in 60 ghz wireless personal area networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3824–3833, 2011.

[97] S. Wu, R. Atat, N. Mastronarde, and L. Liu, "Improving the coverage and spectral efficiency of millimeter-wave cellular networks using device-to-device relays," *IEEE Transactions on Communications*, vol. 66, no. 5, pp. 2251–2265, 2017.

[98] M. L. Attiah, M. Isa, A. Awang, Z. Zakaria, N. F. Abdullah, M. Ismail, and R. Nordin, "Adaptive multi-state millimeter wave cell selection scheme for 5g communications." *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 8, 2018.

[99] M. Amjad, L. Musavian, and M. H. Rehmani, "Effective capacity in wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3007–3038, 2019.

[100] V. K. Shah and A. P. Gharge, "A review on relay selection techniques in cooperative communication," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2, no. 4, pp. 65–69, 2012.

[101] J. D. Kelleher, B. Mac Namee, and A. D'arcy, *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies.* MIT press, 2015.

[102] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," 2011.

[103] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," 2017.

[104] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum machine learning for 6g communication networks: State-of-the-art and vision for the future," *IEEE Access*, vol. 7, pp. 46 317–46 350, 2019.

[105] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, April 2017.

[106] J.-M. Kang, C.-J. Chun, and I.-M. Kim, "Deep-learning-based channel estimation for wireless energy transfer," *IEEE Communications Letters*, vol. 22, pp. 2310–2313, 2018.

[107] J. Liu, K. Mei, X. Zhang, D. Ma, and J. Wei, "Online extreme learning machine-based channel estimation and equalization for ofdm systems," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1276–1279, July 2019.

[108] T. Zhang, S. R. Liu, W. Xiang, L. Xu, K. Qin, and X. Y. Yan, "A real-time channel prediction model based on neural networks for dedicated short-range communications," in *Sensors*, 2019.

[109] H. Ye, G. Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in ofdm systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, Feb 2018.

[110] S. Aldossari and K. C. Chen, "Relay selection for 5g new radio via artificial neural networks," *The 22ND International Symposium on Wireless Personal Multimedia Communications (WPMC)*, vol. pp. 1–6, Nov 2019.

[111] E. Kurniawan, L. Zhiwei, and S. Sun, "Machine learning-based channel classification and its application to ieee 802.11ad communications," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–6.

[112] F. Report, "Fcc takes steps to open spectrum horizons for new services and technologies," *Report*, vol. WASHINGTON, March 2019.

[113] F. Hossain, T. K. Geok, T. A. Rahman, M. N. Hindia, K. Dimyati, S. Ahmed, C. P. Tso, A. Rahman, and N. Ziela, "An efficient 3-d ray tracing method: prediction of indoor radio propagation at 28 ghz in 5g network," *Electronics*, vol. 8, no. 3, p. 286, 2019.

[114] Y. Xing and T. S. Rappaport, "Propagation measurement system and approach at 140 ghz-moving to 6g and above 100 ghz," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec 2018, pp. 1–6.

[115] M. Ayadi, A. Ben Zineb, and S. Tabbane, "A uhf path loss model using learning machine for heterogeneous networks," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 7, pp. 3675–3683, July 2017.

[116] S. Sun, T. S. Rappaport, T. A. Thomas, A. Ghosh, H. C. Nguyen, I. Z. Kovács, I. Rodriguez, O. Koymen, and A. Partyka, "Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5g wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 2843–2860, May 2016.

[117] H. T. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, 1946.

[118] S. R. Saunders and A. Aragón-Zavala, *Antennas and propagation for wireless communication systems*. John Wiley & Sons, 2007.

[119] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE access*, vol. 6, pp. 32 328–32 338, 2018.

[120] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, and J. Jarvelainen, "Propagation path loss models for 5g urban micro- and macro-cellular scenarios," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–6.

[121] H.-N. Dai, R. C.-W. Wong, H. Wang, Z. Zheng, and A. V. Vasilakos, "Big data analytics for large-scale wireless networks," *ACM Computing Surveys*, vol. 52, no. 5, p. 1–36, Sep 2019. [Online]. Available: http://dx.doi.org/10.1145/3337065

[122] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[123] R. Fletcher, *Practical methods of optimization (2nd ed.)*. John Wiley & Sons, 1987.

[124] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature," *Geoscientific model development*, vol. 7, no. 3, pp. 1247–1250, 2014.

[125] V. L. Berardi and G. P. Zhang, "An empirical investigation of bias and variance in time series forecasting: modeling considerations and error evaluation," *IEEE Transactions on Neural Networks*, vol. 14, no. 3, pp. 668–679, 2003.

[126] Z. Lan, C.-S. Sum, J. Wang, T. Baykas, F. Kojima, H. Nakase, and H. Harada, "Relay with deflection routing for effective throughput improvement in gbps millimeter-wave wpan systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1453–1465, 2009.

[127] Y.-H. Nam, Y. Li, and J. C. Zhang, "3d channel models for elevation beamforming and fd-mimo in lte-a and 5g," in *2014 48th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2014, pp. 805–809.

[128] D. Burghal and A. F. Molisch, "Rate and outage probability in dual band systems with prediction-based band switching," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 872–875, 2018.

Appendix A: Copyright Permissions

Below are the permissions for the usage of published materials in this dissertation.

SPRINGER NATURE

Thank you for your order!

Dear Mr. Saud Aldossari,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	University of South Florida
Order Date:	Apr 1, 2020
Order Number:	4800341182791
Publication:	Wireless Personal Communications
Title:	Machine Learning for Wireless Communication Channel Modeling: An Overview
Type of Use:	Thesis/Dissertation
Order Total:	0.00 USD

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
https://myaccount.copyright.com

Copyright Clearance Center RightsLink®

Figure A1: A Copyright Permission from Springier Nature

.

Saud Aldossari <saldossari@mail.usf.edu> Fri, Mar 13, 8:09 PM
to contact

Greetings,

I already presented two papers in the WPMC 2019 shown below:

1- Predicting the Path Loss of Wireless Channel Models Using Machine Learning Techniques in MmWave Urban Communications
2- Relay selection for 5g new radio via artificial neural networks.

I'm using these papers for my dissertation, would please give me the permission to include these materials in my dissertation?

Regards,

Saud

Saud Aldossari. Ph.D. Student
University of South Florida
ENB 237
e-mail: saldossari@mail.usf.edu

contact@wpmc-home <contact@wpmc-home.com> Fri, Mar 13, 9:41 PM
to me

Dear Saud

Thank you for joining the WPMC2019.

It is no problem to include your own papers in your dissertations.

We hope you will get the degree, and do god works in communications area.

Best regards,
Secretary of WPMC Steering Board.

Figure A2: An Email Permission for Both WPMC Papers

.

About the Author

Saud Aldossari earned his Bachelor and Master of Science in Electrical Engineering from Ohio Northern University, USA, and Florida Institute of Technology, USA, in 2011 and 2015, respectively. On the same year, he started Ph.D. studies in Electrical Engineering and that is focusing on Wireless Communication Systems at the University of South Florida. Prior to his master's degree, he worked as a lecturer at Sattam bin Abdulaziz University, Saudi Arabia. His areas of research interest encompass wireless communication systems, measurement, and modeling of wireless radio channels, millimeter- wave, artificial intelligence, big data, and data analytics in the wireless communication domain.