This paper is available online at www.praiseworthyprize.org
International Journal on Communications Antenna and Propagation (I.Re.C.A.P.), Vol. 13, N. 5
ISSN 2039 – 5086
October 2023

Designing and Investigating the Effect of Single-Intermediate and 2-Intermediate Layers of Vanilla Neural Networks in Electromagnetic Path Loss Prediction

Virginia C. Ebhota, Thokozani Shongwe

Abstract – *Impact of two designed architectural compositions of Vanilla Neural Networks (VNNs): the single- intermediate layer of VNN and two (2)- intermediate layers of VNN, in the prediction of signal power loss were examined in this work using measured data from a Long Term Evolution (LTE) network collected from a micro-cell built up environment. The effect of different values of learning rate hyper-parameter were also examined on the VNN architectural models. Early stopping method at the ratio of 75%:15%:15% was adopted during the network training process to avoid over-fitting and Bayesian Regularization (BR) mathematical training algorithm was employed for the training process to ensure good network generalization. Both the intermediate layer neurons for the single- intermediate layer of the VNN and 2-intermediate layers of the VNN were carefully selected to ensure adequate and robust training with excellent predictive performance. Statistical performance metrics, the Correlation Coefficient (R), the Standard Deviation (SD) and Root Mean Squared Error (RMSE) were employed for result analysis of the performances of the two VNN architectural structures while the Mean Squared Error (MSE) and Coefficient of Regression (R) were employed for examining the effect of the various selected values of learning rates on the prediction performances of the two VNN models.The overall experimental results under similar circumstances for training both the VNN models while considering their architectural structures and the effect of learning rates show that for an efficient neural network training and prediction of signal power loss, a well-trained single-intermediate layer of VNN architectural structure with an appropriate neuron numbers gives more efficient and optimal prediction results as the results of the actual output is closest to the desired output showing optimal prediction in comparison with training using 2-intermediate layers of VNN model. The best neural networks training result outputs of a single intermediate layer VNN with 52neurons in the intermediate layer gives R of 0.9728, SD of 1.2758 and RMSE of 1.7435 while best training result of the 2-intermediate layers gives R of 0.9531, SD of 1.5407 and RMSE of 2.2754 on application of [16, 20] neurons which gives the highest prediction results of the considered neuron numbers. Training the single-intermediate layer VNN with small learning rate of 0.002 shows very high R of 0.9903 while with 2-intermediate layers VNN gives R of 0.8810 The training time required for training the single-intermediate layer VNN is considerable low in comparison to training time required to train 2-intermediate layers VNN for best prediction results. These prediction results are of utmost importance in the planning, design and upgrading of wireless network for optimal performance especially when applied in similar environment to the environment of data collection as VNNs results show adaptability and robustness.*

***Keywords**: Artificial Neural Networks, Vanilla Neural Networks, Electromagnetic Signal Power Loss, Learning Rate, Early Stopping Method, Bayesian Regularization Mathematical Training Algorithm*

Nomenclature

ANN	Artificial Neural Network
AI	Artificial Intelligence
BS	Base Station
BR	Bayesian Regularization
GPS	Global Positioning System
GSM	Global System for Mobile communication
LM	Levenberg-Marquardt
LTE	Long Term Evolution network
MATLAB	Matrix Laboratory
ML	Machine Learning
MSE	Mean Squared Error
MS	Mobile Station

Available online by October 31st, 2023
https://doi.org/10.15866/irecap.v13i5.23345

PL	Path-Loss
RF	Radio Frequency
RMSE	Root Mean Square Error
RSRP	Reference Signal Receive Power
R	Correlation coefficient
SD	Standard Deviation
TEMS	Test Mobile System software
VNNs	Vanilla Neural Networks
f	Frequency
d	Distance
C	Speed
λ	Damping factor
δ	Weight update
β	Performance index
α	Decay rate
E	Error
E_t	Sum squared weight
E_d	Sum squared error
J	Jacobian matrix
A_{tx}	Transmitting antenna
A_{rx}	Receiving antenna
p_{tx}	Transmitting power
p_{rx}	Receiving power

I. Introduction

In wireless radio communication system networks, signal power loss prediction models are unique mathematical models which are utilized by telecommunication network engineers in signal coverage appraisal of the radio signal path attenuation loss and coverage area of Base Station (BS) transmitter of areas served by a given transmitter in the course of network planning and management [1], [2]. Nonetheless, it is always a complex task in the planning of telecommunication networks to develop these signal power loss prediction models with optimal accuracy.

The traditional propagation models such as the Hata model, COST 231 Hata model, Lee model, Walfish-Ikegami model etc. have been widely studied and applied in telecommunication network planning, however they are site specific and lack accuracy most time [3] The limitations of the traditional models are well pronounced when in comparison with Artificial Intelligence (AI) models such as Artificial Neural Network (ANN) models in terms of adaptability and accuracy on application in cellular radio environment aside the environment for its development [4]. These are as a result of the distinctions in the environmental formation, weather conditions, terrain type such as open area, rural area, urban and sub-urban areas which exist in various radio propagation locations in different cities and countries [5], [6]. The traditional models are generally limited in the capturing of non-linear relationship between independent variables such as signal power loss and dependent variables such as distance. The precision of the traditional models on the prediction of signal power loss has widely been studied and reported in various previous works which ranges largely from 8-12dB in terms of Root Mean Square Error (*RMSE*), that is remarkably higher than the acceptable values [5], [7].

The traditional modeling technique is either by use of stochastic models or the deterministic models. These are based either on measurement in case of the stochastic models or based on geometry in case of the deterministic models. The architectural and topographical maps of the propagation environment is adopted in case of the deterministic models for the signal propagation characteristics prediction based on geometry optical theory and theory of electromagnetic signal propagation.

However, there are limited generalization abilities of these traditional models as their applications are either dependent on the accuracy of material and geometric information in case deterministic models or environmental dependent as different explicit types of environment is required in case of stochastic models [3], [6]. As wireless network is a time-varying, non-linear system which obscure multi-dimensional information as spatial domain, frequency or time domain, the traditional models are foists with challenges with coping with trend of diversity, time and mass varying wireless networks.

The ANNs such as Vanilla Neural Networks (VNNs) have the capability of dominant learning of the structural relationship between the complex environments data to approximating non-linear systems automatically. The VNNs has proved to effectively approximate a smooth and measurable function between the input and the output vector by appropriate selection of the connection weights and transfer functions when appropriately trained [2].

Recently, AI soft computing and modeling techniques such as ANNs has proven to solve many functional and pattern classification problems [3], [6], [8]. Artificial neural network models have shown robustness and excellent learning of pattern classification and function approximation [8]. Some of the key ANN models such as the VNNs, Generalized Regression Neural Networks (GR-NNs), etc. have numerous algorithms which are robust and have been explored to carry out improved proficient adaptive non-linear statistical modeling over classical logistic regression methods which are constantly engaged in the development of predictive models [1], [9].

The robustness of ANN such as VNNs are attributed to their learning and predictive ability as well as ability to classify non-linear data employing experience from antecedent samples introduced for input-output mapping as have been noted from studies over the years [10]-[12].

Largely, an enormous degree of freedom of structure that provides fitting for numerous datasets with linear or else non- linear correlation designs are definite underlying capability of ANNs models over the traditional log-distance-based models.

Artificial Intelligent based model such as ANNs for optimal and adaptive signal power losses predictions during electromagnetic signal transmissions have their concepts introduced to overcome the limitations of the existing traditional deterministic and empirical Path Loss (*PL*) developed models [10], [11].

This paper designed and thoroughly examines the architectural influences of a back-propagation feed forward VNNs in the prediction of signal power loss. Two architectural structure of VNNs were exerted, a single-intermediate layer VNN and VNN with 2-intermediate layers. VNN The influence of a hyper-parameter such as learning rate in the training of the VNN with single-intermediate layer and VNN with 2-intermediate layers were also examined. This examination of the VNNs is for optimal predictive modeling of signal power loss in urban micro-cell built-up radio environment.

Although there has been existence of various traditional VNN models in literature, their correct application has remain an open and a challenging task. This is because their development and correct selection of the network structural designs for required input elements and hyper-parameters in solving distinctive mapping and functional predictive problems has been by trial and error [13], [14]. While highlighting the major robust advantages of VNNs, the search to address issues such as appropriate selection of the network architectural complexity with the appropriate required input elements for signal power loss prediction during electromagnetic signal transmission is a major motivation for this research work.

A major VNNs challenging task is how to correctly select the network architecture for required input element and the hyper-parameters in solving a given problem [13]. A distinctive VNN based signal power loss prediction models with a well-structured network architectural implementation which is empowered with grid search based hyper-parameter tuning technique for optimal signal power loss prediction using dataset collected from electromagnetic transmission between the BS and the Mobile Station (MS) path length has been well investigated in this paper. The effect of the learning rate hyper-parameters by variation of its various values on both the single-intermediate and 2-intermediate layers of VNNs were examined.

The experimental dataset were acquired from a Long Term Evolution (LTE) cellular network via a drive test in an urban micro-cell built-up environment. A 2019a Matrix Laboratory (MATLAB) in neural network tool box was exerted for the training of the datasets using developed VNN models. The required user interface, the mathematical training algorithm, training, testing and validation platform were all provided by the neural network tool box.

Contribution to knowledge via this paper are as follows: investigation of the effect of a distinct VNNs in the prediction of signal power loss with a well-designed VNN architectural structural implemented for an optimal results. The training, testing and validation were carried out using 1st order statistical performance indices, *SD* and r for result analysis. Optimization of the examined VNN models through hyper-parameter tuning leveraging grid search algorithm analyses of the experimental signal power loss dataset and the optimal prediction effectiveness of the designed VNN models in comparison with the well-structured implementation architecture over log-distance models using various 1st order statistical performance indices are also contribution to knowledge in this research work.

The continuing part of the research paper is organized as: Section II is the background of study where relevant areas such radio propagation mechanism, log-distance based signal power loss prediction models and the foundation of ANNs were studied. Section III shows the methodology while detailing the implementation of both the single-intermediate layer VNN and the 2-intermediate layers VNN for predictive modeling of signal power loss exerting acquired dataset from LTE micro-cellular built-up environment. Section IV is the result and analysis while Section V is conclusion.

II. Background Study

This section discusses the mechanism of radio propagation, traditional signal power loss prediction models and ANNs.

II.1. Radio Propagation Mechanism

Electromagnetic radio signal in course of transmission interact with media and objects, the consequences of such interaction being a weaker radio signal as a result of reflection, refraction, absorption, diffraction and various other propagation phenomenon [15]. This resultant effect of the phenomenon on transmitted signals lead to loss in signal power called propagation loss. The features of the medium by which the radio signals transmits determines the amount of transmission loss or propagation loss, as well as the distinction of the received signal which is attainable at the receiving terminal. Electromagnetic signal propagation losses are also governed by other diverse elements such as receiver sensitivity, transmitter power, and the overall antenna parameters such as the antenna gain, the receiver location and the antenna height.

The major factors which influences the number of signal power loss in a medium comprises of reflection, diffraction, refraction, absorption and scattering, etc.

When radio signals collide with large amount of obstacles compared to the propagating signal wavelength, it give rise to diffraction. This mainly occur when radio signals bend around objects, mostly those with sharp edges. This changes repeatedly authorizes the energy of the received signal to spread round the boundaries of the obstructing object [15]. Also, amplitude, phase, frequency and pathway of the transmitted signal influences diffraction. There will be undoubtedly negative effect from the environment of the radio frequency signal transmission. Factors such as terrain landscape, population density, building structures lead to huge variation in signal transmission as well as propagation losses. Damp, sandy and marshy terrain attenuate radio signals, primarily low frequency signal with fast signal transmission in conducive terrain over damp, sandy or marshy terrains.

II.2. Traditional models for Signal Power Loss Prediction

Largely, signal power loss prediction models are set of mathematical models expression and algorithms exerted for the prediction of signal attenuation losses between the path of BS and MS. These models are useful planning tools which assists the radio network engineers in the design of cellular telecommunication systems by optimal positioning of the BS transmitters to meet the required signal coverage level and the quality of service required by the communication network. Predictive power and the performance of every signal power loss model is based on the resultant prediction accuracy with the actual field measurement data. The average power loss of traditional signal power loss is arithmetically dependent on distance i.e. length of the transmission path, knotted with transmission exponent modeling parameters. The propagation exponent is mostly used to account for particular radio environment. These can be termed as simplified models which endeavor to model fluctuations, attenuations, variations in the received signal power.

Various exemplars of traditional based prediction models include the COST234 Hata model, the Hata Okuruma model, the Lee model, the Ikegami model, the Egli model, the SUI model, the Walficsh-Ikegami model, Free space path loss, etc. [16].

Free Space Path Loss (*FSPL*) is loss in signal strength of electromagnetic signal as a result of Line of Sight (LOS) path through free space. It is the loss between two isotropic radiators in free space which is expressed as a power ratio. For instance, the free space path loss model is based on the assumption that there is a single path without any obstruction between the transmitter and receiver [14], [16]. It assumes a power ratio of 1.0 or 0dB for the antenna gain and no loss associated with hardware imperfections or the effects of any antenna gains is included:

$$FSPL = \left(\frac{4\pi d}{\lambda}\right)^2 = \left(\frac{4\pi df}{c}\right)^2 \quad (1)$$

where λ is signal wave length (meters), f is frequency of the signal (hertz), d is distance from the transmitter (meters), c is speed of light in a vacuum, (2.99792458×10^8 meter per second).

However this equation is only accurate for far field where spherical spreading can be assumed not to the transmitter:

$$FSPL(dB) = 10\log_{10}\left(\left(\frac{4\pi df}{c}\right)^2\right) = = 20\log_{10}\left(\frac{4\pi df}{c}\right) \quad (2)$$

$$FSPL(dB) = 20\log_{10}(d) + 20\log_{10}(f) + +20\log_{10}\left(\frac{4\pi}{c}\right) \quad (3)$$

$$FSPL(dB) = 20\log_{10}(d) + 20\log_{10}(f) - 147.55 \quad (4)$$

f is measured in units of GHz and d in km for a typical radio application, thus:

$$FSPL(dB) = 20\log_{10}(d) + 20\log_{10}(f) + 92.45 \quad (5)$$

The constant becomes 32.45 for d, f in kilometers and megahertz respectively and -87.55 for d, f in meters and kilohertz respectively and -27.55 for d, f in meters and megahertz respectively

However, most time the application of Free space model is as part of Friis transmission equation that includes antenna gain. Friis proposed a formula for the free space transmission loss that defines the ratio between the received power P_{rx} and the transmitted power P_{tx} in terms of the efficient transmitting antenna area A_{tx} receiving antenna A_{rx} distance d (m) and damping factor λ.

Modeling exerting this type of traditional model, the received power is a function of transmitted power, antenna gain and distance between the transmitter and the receiver [16]:

$$Transmission\ loss = \frac{P_{rx}}{P_{tx}} = \frac{A_{rx}A_{tx}}{d^2\lambda^2} \quad (6)$$

The above can further be simplified as:

$$\frac{P_{rx}}{P_{tx}} = \left(\frac{\lambda}{4\pi d}\right)^2$$

for an ideal isotropic antenna.

Converting Equation (6) to a distance in km instead of m, frequency in MHz instead of wavelength in m, and converting the linear domain power units (W) to log domain unit (dBm), a usual reference equation of path loss as a function of carrier frequency and distance is gotten, thus:

$$P_{rx} = P_{tx} - \left(20\log_{10}(d) + 20\log_{10}(f) + 32.45\right) \quad (7)$$

$$P_{dBm} = 10\log_{10}(P_{mw}) \quad (8)$$

Although the traditional path loss prediction models possess changing frequency rationality thresholds, various correction factors have been employed for easiness of their applicability at the tested frequency band. One robust way to address the limitations of the traditional models is by use of ANNs as they cater for stochastic signal attenuation phenomenon and for the inhomogeneity of the spatial propagation channels in different radio communication environment

II.3. Concepts of Artificial Neural Networks

Artificial neural networks are distinct and robust non-linear statistical data modeling networks generally defined as interconnected neurons which are pre-arranged in various layers [17]. The processing elements are known as the neurons which are processor that operates on the inputs they are fed through the connections. The processors are vast parallel-spread out and comprises of simple processing units with natural tendency of storing and making available exponential knowledge [18]. Artificial neural networks acquire knowledge from the environment by means of learning process and the acquired knowledge is stored in the network by inter-neuron connection strength called synaptic weights [17], [18]. It derives its computing power from its architecture and its learning ability to generalize well [18]. The neurons are distinct mathematical functions which captures and organizes information in accordance to the neural network architecture.

The ANNs have assorted family of networks with the functionality of each type of network determined by the architectural structure, adopted mathematical training algorithm, neuron characteristics etc. It derives its computing power from its architecture and its learning ability to generalize [19]. The use of ANNs offer the following capabilities and properties:

i. Non-linearity: Artificial neural network can either be linear or non-linear network.
ii. Input-Output mapping: When input data which is desired output is presented to the ANN, there is modification of network synaptic weight to reduce the difference between desired and actual output to minimal according to the appropriate statistical principles. The network training is repeatedly carried out until a steady state where no observed changes in the synaptic weight adjustment are seen. Thus, the network learns by constructing input-output mapping for the considered problem.
iii. Adaptivity: Artificial neural networks possess in-built capacity of their synaptic weight adaptation to changes in proximate environments. However, its adaptive capability may not always result to robustness.
iv. Contextual information: The architectural design and ANNs activation state is used in knowledge representation. Each neuron in network architecture is probably affected by global activities of other network neurons.
v. Fault tolerance: Artificial neural network which is implemented in hardware form has the capability of robust computation.
vi. Analysis design and uniformity: The universality of the ANNs processing unit i.e. the neuron, permit it to share theories and training algorithm in various ANNs applications.

II.4. The Basic of Vanilla Neural Networks

II.4.1. Architectural Arrangements of Vanilla Neural Networks

The theorem of universal approximation asserts that the three-layered VNN approximates practically any non-linear function [20]. Nonetheless, there is no stipulation of the needed number of intermediate layers or number of neurons for solving definite problem. The required numbers of the intermediate layers and neurons for solving problems are dependent on the complexity of the problem and has stayed as an open enigma for effectual neural network training. So many intermediate layer neurons could result in over-learning of the dataset during training of network, and insufficient neurons in the intermediate layer may result in inadequate training of the dataset [18].

Artificial neural network can be measured in terms of its mapping and generalization ability. Practically VNNs with one intermediate or two intermediate layers is normally utilized in Radio Frequency (RF) and also microwave utilization. In minimum, an intermediate layer is necessary for the estimation of non-linear function. Nonetheless, four-layered VNN could show superiority in non-linear problems modeling with repeated subsistence of definite localized sections conducts in diverse areas of problematic space [18]. A three-layered VNN though can model such problems, it could necessitate copious intermediate layer neurons.

Neurons are switches that has information input and output which are triggered with sufficient stimuli of other neurons bugging intelligence input and conveying a rhythm at the information output [21]. There are conveyance of information from synapses to the neuron dendrites that are distinct connection. Mathematically, an intermediate layer of VNN can randomly estimate functions, nonetheless, is only alongside first derivations and limited discontinuities [18]. There is however no constructive proof as there is breach as to the precise neurons numbers and the weight required. An n-layer has an n-variable weight layers then n+1 neuron layer.

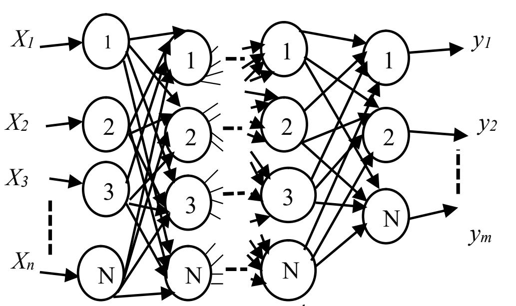

Fig. 1. Vanilla Neural Network with multiple intermediate layers (Adapted from V.C Ebhota et.al, 2018)

For L total number of layers, the 1st layer is the input layer, 2nd layer which is the 1st intermediate layer is L-1, while L^{th} layer is the output layer. For N_L numbers of neuron in L^{th} layer, then L = 1, 2, 3... If w^1_{ij} signifies weight of jth neuron of $(L\text{-}1)^{th}$, then the layer and i^{th} neuron of the L^{th} is $1 \le j \le N_{L-1}$, $1 \le i \le N_L$. If x_i signifies i^{th} external input, VNN and z_i^L is the output of the i^{th} neuron of L^{th} layer. Introduction of extra weight parameter for every neuron $w_{i0}{}^L$ signifies the i^{th} neuron of L^{th} layer bias. So, weight w of VNN comprises $w^L{}_{ij}$ s, j = 0, 1, 2, 3..., $N_{L\text{-}1}$, and i = 1, 2 3, ...N_L, L = 1, 2, 3, ..., thus:

$$w = \left(w_{10}^2 w_{11}^2 w_{12}^2 ....w^L N_L N_{L-1} \right)^T \tag{9}$$

Basically, the optimal weight value w is gotten during the VNN training, there is weight modifications in a way that there is minimal error between the output VNN models and original problem output.

The VNN is a feed-forward Neural Network (NN) that exerts back-propagation for network training and capable of solving complex non-linear problem with quick training and prediction while handling huge amount of input dataset [18], [21]. The same level of accuracy achieved on training with smaller input dataset can also be achieved with large dataset. However, VNNs unlike some other NNs are purely feed-forward network with no loop or cycles as seen in other Deep Neural Network (DNN) like in Recurrent Networks. A DNN has certain level of complexity and adopts sophisticated mathematical models for data processing in certain way.

The VNN can also be said to be under the umbrella of DNN but a feed forward NN with few intermediate layers capable of handling both small input dataset as well as large dataset with same prediction accuracy.

Deep neural network unlike the VNNs are mainly exerted in dealing with un-structural data while VNNs are easily exerted for simple input/output protocols [21].

II.5. Artificial Neural Network Training Algorithm

A major characteristics of ANNs are their ability to get familiarized with unknown problems and thereafter solves problems of similar class. They actualize this by using learning rules known as learning/training algorithms [29]. The aim of the training algorithm is for the establishment of weight combination that gives a minimal error, thus selection of appropriate training algorithms are critical for adequate network training.

There are various phenomenon ANNs learn from and there are changes in the learning system for situation adaptation such as environmental changes, weight modification for assurance of fast and reliable learning etc. [30]. The aim of the training algorithms are for global error reduction by adjustment of weight and bias during network training and training algorithm that has proven to give the least prediction error are employed.

II.5.1. Artificial Neural Network Training Using Bayesian Regularization Mathematical Training Algorithm

Bayesian Regularization (BR) mathematical training algorithm is harnessed for weight update in course of training the network which is in accordance with Levenberg-Marquardt (LM) mathematical algorithm. The BR mathematical training algorithm has established an improved training by linear permutation of weight variables and squared error [31]. There is modification using BR mathematical training algorithm in accordance with LM function approximation techniques [32], [33]:

$$\left[J^t J + \lambda I \right] \delta = J^t E \tag{10}$$

where J = Jacobian matrix and δ = weight update vector, E = error and λ= damping factor. For the optimization process, the damping factor is modified at every iteration.

Customarily, there is approximation of the Hessian utilizing Jacobian matrix [34]:

$$H = J^t J \tag{11}$$

Levenberg-Marquardt mathematical training algorithm centers on initial weight of the network, there could be occurrence of at local minima, or there could be non-occurrence of convergence at all. Jointly, there is no consideration for the first weigh and the data outliers which subsequently gives rise to poor network generalization [34]. Consequently, there is application of the BR mathematical training algorithm in order to evade poor generalization of the network during training. This is achieved by allowing adequate weights which are necessary for resolving the correct problem [35]. This ensures cost function increase for the smallest error detection while the smallest weight is applied. There is initiation of two hyper-parameters which are beta and alpha. They are initiated to announce the learning process direction. The cost function is stated as [36]:

$$F = \beta E_d + \alpha E_s \tag{12}$$

E_d = sum squared error, E_s = sum squared weight. BR mathematical training algorithm addition to LM mathematical training algorithm sums to a small overhead in the network training, β = performance index and α = decay rate parameter.

II.6. Vanilla Neural Networks Implementation

Vanilla neural network models implementation are accomplished by the application of Matrix Laboratory (MATLAB). Matrix laboratory is a distinct programming language that has multi-exemplar numerical computing environmental and user interface. It offers an easy matrix calculation and graphical multi-domain simulation, creating function plotting, figurative computing, simple algorithm simulation, brilliant data mining, etc. Matrix laboratory permits access to the use of optional toolbox with the neural network toolbox having distinct tools for model designs, visualization, implementation and simulation of neural networks [37].

In this research work, the MATLAB was used to encode the script files for the predictive training of the VNN models, for testing and quantitative evaluation. The proposed VNNs consists of various numbers of neurons and layers to ascertain the best architectural structure for signal power loss prediction using measured dataset collected from a LTE micro-cell built-up environment.

The architectural structure of the VNNs is one of the determinant factor in its optimal prediction ability and this has been an open problem as it can only be determined by trial and error. Selection of network processing elements such as the adequate number of neurons and the layers as well as an adequate training algorithm and transfer function are very vital for an optimal prediction. Insufficient neurons in the intermediate layers of VNN may result to failure in capturing the complex link between the input variables and the target output. Equally, too many neurons in the intermediate layer of the VNN may result to over-parameterization leading to poor network generalization and poor predictive modeling of the initial dataset [38], [39]. To improve on generalization during network training i.e. preventing over-fitting, techniques such as early stopping method was adopted. The inputs and the target output dataset were scaled to reside in the range of -1 to +1 to improve training and testing speed. The early stopping techniques were adopted to avoid network over-training and to eliminate contemptuous initial values influence and to develop robust adaptive predictive ability. Though different learning algorithms are available for neural network training such as for the training of VNNs, the BR mathematical training algorithm has proved as the best training algorithm for predictive modeling problems in VNN as concerns convergence speed and accuracy [40]. A thorough search is therefore employed in our study to accelerate the convergence and also evaluate the effect of the VNN training algorithm during neural network training.

II.7. Previous Research Work Using Vanilla Neural Networks for Predictive Modeling

Earlier studies have examined the fitness of various machine learning algorithms and models in the prediction of signal power loss [41]-[45]. The need to overcome the shortcoming of traditional models when applied for signal power loss prediction led to application of machine learning techniques such as ANNs [41].

Artificial neural network signal power loss prediction models are more efficient and has shown ability for easier deployment than the traditional prediction models [45]. The analysis of empirical models with various propagation features were carried out while the model performance with the lowest RMS values compared with prediction from ANN model [46], [47]. The neural network based signal power loss prediction models gives much lower value of *MSE* upon validation. A VNN was introduced for *PL* prediction in [47], where the VNN network was trained using back-propagation algorithm, this was compared with prediction from analytical models with the results showing that the VNN model trained the network more efficiently than the analytical model used in radio network optimization [48], [49].

Artificial neural network models were explored for *PL* prediction in an urban areas [49]. The research work explored the effect of different input parameters and environmental terrain on robustness of the path-loss prediction. The study demonstrates increase in the accuracy of the signal power loss prediction model at increased input parameters, greater number of features results to increased system accuracy. Training of the model was carried out with the assistance of dataset inform of input features. In [43], [44], an ANN based *PL* model at 800 MHz and 1800 MHz were introduced with input for latitude, longitude, distance, clutter height and elevation. The ANN model shows superior performance over Support Vector Machine (SVM) in [50]. In [51], VNN based model was employed for the prediction of *PL* at 1800MHz in smart campus environment. The VNNs with 2-intermediate layers was used and the prediction performance of output results shows superiority over prediction made using RF. Various machine learning based prediction models were used for signal power loss prediction for wireless sensor network [52]-[54]. In [55], [56], the machine learning based prediction models were employed for signal power loss prediction in wireless sensor network which produced lowest values of *RMSE* in comparison to the analytical models.

II.8. Vanilla Neural Network Parameters and Search Space

Hyper-parameters are distinct set of regulating parameters which the neural network models utilizes for adaptive learning process in data training, testing and validation. These distinct parameters could be categorical, integer variables or continuous variables that has value range normally lower and upper bounded. The hyper-parameters include the number of the intermediate layers, the number of the neurons in the intermediate layer, the learning rate, transfer function etc.

i. *The Intermediate Layer-* Determining the required number of the intermediate layer for architectural structure of VNN is a vital issue while choosing VNN model for predictive modeling. Adopting of various intermediate layers during the construction of VNN can result to poor generalization during network training as a result of complexity in network training. Two intermediate layers in combination with an output neurons are suitable for ANN to learn N data samples and deliver insignificant errors [41], [42].

ii. *Neuron*- Determination of the neuron numbers remains very integral part of the comprehensive ANN architecture. Inadequate neurons in the intermediate layer of the VNN can result to under-fitting problem during network training. This is incapability of the intermediate layer neurons to learn or predicts signals satisfactorily during network training [18], [23]. Also, too many neurons in the intermediate layer can result to over-fitting problem. This is too many information processing capability problems. This leads to excessive neural network training time, inadequate or impossible training of the neural network [23].

iii. *Transfer Function*- The transfer function is a distinct monotonically increasing and differentiable function employed in translating input data signals to give the final output signal of a neuron [50]. This is fundamental to the concrete concept of artificial neural network for two major reasons, firstly, the entire organization of the ANN will be similar to a typical linear function which cannot learn non-linear relationships without activation function and secondly, transfer function styles the major computation achieved by the neural networks.

iv. *Learning Rate*- Learning rate is a hyper-parameter which standardizes the weights of neural networks in relation to loss gradient [31], [42]. It has to be cautiously selected to robustly support both generalization and optimization. A large learning rate value could lead the complete learning process to hurdle over minima while small learning rate value could lead to a learning process which requires long convergence time resulting in it being stuck in negative and erroneous local minima. Hyper-parameter optimization such as in selection of the appropriate learning rate conveys the robust procedure of finding best feasible values of the hyper-parameters for machine learning model to attain desired result or modeling outcome. Common hyper-parameter tuning algorithms include Bayesian optimization, random search and grid search. Cycling of different selected values of learning rates for the VNNs training and studying their effects during network training were examined in this work.

III. Methodology

The method employed in the design of the VNN models with a well-structured network architecture which is enabled with adequate hyper-parameter tuning algorithm for optimal prediction of signal power loss are listed in the steps below. The steps explored for the prediction of signal power loss using VNN models is underlined:

i. Acquiring the dataset.
ii. Pre-processing of the dataset.
iii. Designing of the VNN models: the single-intermediate layer and the 2-intermediate layers of VNN model.
iv. Application of BR mathematical training algorithms for the neural network training, testing and validation.
v. Training of the designed VNN models by cycling the neuron numbers and selected learning rate for both the single-intermediate VNN and 2-intermediate layers VNN to obtain the best training, testing and validation results.
vi. Appraise and validation of the accuracy of the training process and training results.

The VNNs trainings were carried out using the dataset of Fig. 2 where time represents the covered distance during data capturing and *t* represents the signal power loss (dBm). The dataset were normalized exerting excel spreadsheet for easy training in deep learning tool box of MATLAB 2022b. These were exerted as input to the neural network and the VNNs target is on the prediction of the signal power pattern with minimal error.

III.1. Data Measurement and Collection, Tools and Procedure

The conducted field measurement was to acquire signal data across LTE BS antennas over a period of time. This is to enable study of the location seasonal variations and the LTE transceiver BS antennas which are operating at 2600 MHz and 10 MHz bandwidth [37].

The transceiver BS antennas are sectionalized with 17.5 dBi gain and 43 dBm transmit power.

The measurement were carried out for radio spectrum analysis and field test tools such as scanners, LTE proficient mobile handsets, Global Positioning System (GPS) device, a car for the derive test, a laptop installed with a Test Mobile System (TEMS) software, compass etc. were exerted in the field measurement for data collection. The dataset generated was pre-processed employing computer software such as excel spreadsheet for dataset normalization.

The TEMS software entrenched in the mobile handsets and the laptop enables access, recording and extraction of signal data along all measurement routes.

```
time = [36   71   87   44   85   94   103  113  121  130  140  149
        167  176  184  193  197  196  201  209  219  228  234  247
        254  273  285  297  319  341  355  374  385  389  392  400
        410  420  428  441  456  455  428  424  422  430  448  461
        477  502  519  531  554  562  600  632  655  698  624  434
        440  462  470  485  500  507  515  530  545  561  577  584
        592  607  599  614  629  637  755  763  771  779  796  804
        812  820  827  858  866  873  882  898  905  914  922  939
        948  955  964  971  979  988  996  1012 1020 1030 1046 1054
        1069 1076 1084 1094 1102 1109 1118 1125 1134 1141 1150 1166
        1175 1183 1191 1200 1208 1224 1232 1239 1247 1272 1279 1303
        1311 1328 1336 1354 1361 1370 1402 1409 1417 1426 1432 1442
        1463 1472 1480 1488 1497 1505 1628];

   t = [-83.5 -86.88 -90.13 -78.69 -89.69 -90.25 -85.19 -81.25 -89.44   -88.25  -94.25  -89.5 -90
        -91.69 -87.31 -89.94 -92.81 -86.25 -82.06 -92.19 -95.31 -103.13 -90.5 -101.19 -100 -99.88
        -96.38  -95.25  -93.81  -102.38 -105.56 -102.44 -94.63  -93.19  -96.38  -93.94  -93.69  -95.69
        -94.13  -99.63  -96.13  -97.06  -104.13 -99.06  -94.31  -99.94  -97.38  -100.19 -100.56 -101.31
        -90.13  -99.94  -101.5  -94.19  -100.63 -97.38  -100.19 -99.63  -95.75  -103.56 -99.31  -102.5
        -101.44 -100.06 -97.88  -95.94  -93.06  -98.06  -99.13  -94.75  -98.81  -100.25 -98.63  -100.5
        -100.5  -101.38 -101.06 -104.81 -102.06 -103.13 -106.5  -105.13 -105.19 -105.44 -104.56 -107.88
        -103.75 -104.75 -101.06 -102.06 -107.69 -105.19 -107.31 -109    -108.5  -107.94 -109    -105.13
        -100.56 -103.25 -102.56 -104.44 -106.63 -111.69 -108.06 -110.13 -110.56 -106.56 -106.56 -108 -
107.94 -108.38 -108.25 -107.13 -104.56 -107.5 -110.31 -110.75 -108.63 -107.06 -102.12  -106 -104.44 -
103.8 -106.25 -107.13 -107.94 -110 -105.81 -104.75 -108.31 -103.81 -106.81      -110 -104.31 -111 -
109.06 -110.5 -106.56 -107.94 -112.81 -110.39 -109.25 -108.94 -111.13  -113.06 -112.44 -116.44 -112.75
        -112.69 -110.75];
```

Fig. 2. Acquired dataset exerted for VNNs neural network training

The GPS and the compass were engaged in matching up MS i.e. user equipment measurement locations in correspondent to BS transmitter and the field test environment. The extracted signal parameters from the derive test log files for prediction analysis is the Reference Signal Received Power (*RSRP*) (dBm) at received terminal. The signal power measurement were conducted in a built-up environment. The *PL* dataset for prediction are related to radio signal data by the measured *PL* data in relation to the acquired radio signal data. PL_{mea} (dB) values are gotten from the measured signal, *RSRP* (dBm) as expressed in Equation (13):

$$PL_{mea}\left(dB\right) = EIRP + GA - RSRP_{meas} \tag{13}$$

EIRP is calculated as:

$$EIRP = PTX + GTX - CLTX \tag{14}$$

GTX = BS transmit antenna gain and *GA* = MS antenna gain. *PTX* =transmitted power, *CLTX*= transmission cables loss, all is in dB. Some of the key parameters for the BS antenna obtained during the derive test are listed in Table I.

III.2. Training of Designed Vanilla Neural Network Models

For optimization of the neural network during network training, the training dataset encompass measured dataset collected from various propagation routes. Radio propagation features with instances like diffraction, reflection, direct rays etc. were adequately considered.

The picked routes comprise of the received positions which shows different input parameter ranges. Therefore, the neural network learns how to behave in diverse situations while making appropriate generalization during application on new instances. Appropriate measurement point characterization of acquired datasets from the training routes is an important step in the training process to their type of dominant path.

At total number of two thousand, eight hundred and sixty (2,860) measured dataset were detailed with all having distinctive received signal power. The training exerting the acquired dataset was carried out on the two designed VNN, the single-intermediate layer VNN and the 2-intermediate layers VNN. Variations of the neurons were done in fours up to hundred for the single-intermediate layer VNN and in fours for the first intermediate layer of the 2-intermediate layers VNN and in eights for the second intermediate layer. These variations were sequentially carried out while cycling the neuron numbers up to 100 numbers for the single-intermediate layer VNN and one hundred and four (104) for the 2-intermediate layers VNN to ensure a thorough investigation of appropriate number of neurons required in the single-intermediate layer VNN and appropriate number of neurons required in the intermediate layers of the 2-intermediate layers of VNN for optimal prediction of signal power loss.

The results of the single-intermediate layer VNN and the 2-intermediate layer VNN models trainings are shown in Table II and Table III respectively.

A written neural network program was exerted for the simulation and the computation which were realized using artificial neural network toolbox in MATLAB 2022b. Over-fitting during network training which normally lead to reduction in the predictive abilities and the modeling capability of ANN was outmaneuver by exerting early stopping method achieved by division of the dataset into three parts in the ration of 70%:15%:15% for the each of the VNN training, testing and validation respectively. The learning rate hyper-parameters were cyclically varied from 0.002 to 0.03 to ascertain the best learning rate for optimal training of the single-intermediate layer VNN and the 2-intermediate layers VNN respectively. The neural network training for the two models are shown in Table IV for the single-intermediate layer VNN and Table V for the 2-intermediate layers VNN. The BR mathematical training algorithms which updates bias and weight in accordance with the LM mathematical training algorithm during neural network training was exerted for the network training of the two VNN models. The input dataset were normalized in Excel spreadsheet as expressed [31]:

$$d_n = \frac{d}{\sqrt{\sum_{l=1}^{n}(d_o)}} \tag{15}$$

d_n = normalized data value, d_o= original data value.

III.3. Performance Metrics

Two statistical performance metrics were employed for result analysis of neuron variation in the intermediate layers of the exrted VNNs. These are:

i. Standard Deviation (*SD*): This measures the amount of variation between measured and prediction values. Low standard deviation shows closeness of data points while the reverse is the case with high standard deviation shows [57]:

$$SD = \sqrt{\frac{1}{N_{\exp}}\left(\sum_{d=1}^{N_{\exp}} |\, l_d - y_d \,| - \text{MAE}\right)^2} \tag{16}$$

ii. Coeffient of Regression (*R*): This measures the statistical relationship between measured and prediction values. It returns a value between -1 and +1. The +1 signifying a strong positive connection and -1 signifying a negative connection [58] (17):

$$R = \frac{N_{\exp}\sum l(p) - y_0(p) - \left(\sum l(p)\left(\sum y_0\right)\right)}{\sqrt{N_{\exp}\left(\sum l(p)\right)^2} - \left(\sum l(p)\right)^2 \sqrt{N_{\exp}\left(\sum y_0(p)\right)^2}}$$

TABLE I
FIELD PARAMETERS

Parameter	Site
BS Operating transmitting Frequency (MHZ)	2600
BS antenna Height (m)	28
MS Antenna Height (m)	1.5
BS Antenna gain (dBi)	17.5
MS Antenna gain (dBi)	0
BS Transmitting Power	30
MS transmitting power	43
Transmitter cable loss	0.5
Feeder loss	3

iii. Root Mean Squred Error

The *RMSE* suggests the mean error magnitude between the actual values and the prediction values. The *MSE* is the mean average of the square difference between the actual recorded values and the prediction values:

$$RMSE = \sqrt{MSE} = \frac{1}{K_{measured}} \sqrt{\sum_{k=1}^{k_{measured}} [t_k - y_k]^2} \quad (18)$$

N_{exp} = number of experimental measured data, $l(p)$= values of measured signal power loss, p^{th} = input pattern, and $y_o(p)$ = neural network output.

IV. Experimental Results and Discussions

It is a challenging task applying VNN models for neural network training and prediction and the choice of apposite network architectural structure for the prediction of signal power loss. This is because designing and formation of apposite VNN architecture depends largely on trial and error and thus has remain an open problem.

In this research work, the effect of two hyper-parameters, the learning rate and the number of neurons in the intermediate layers of exerted VNNs in the prediction of signal power loss were considered. The designed VNN models for the prediction of signal power loss were trained using measured dataset collected via a derive test from LTE network microcell built-up environment. Bayesian Regularization mathematical training algorithms was exerted for the training purposes.

The effect of the two architectural structures of the VNNs: the single-intermediate layer VNN and the 2-intermediate layers VNN were examined during neural network training for efficient optimal prediction of the training dataset.

Different number of neurons in the single-intermediate layer and in the 2-intermediate layers as well as different values of learning rates were applied during the neural network training and their performances examined. The respective results for the different architectural structures of the VNNs in network training for both neuron variations and learning rate cycling are shown in the Table II, Table III, Table IV and Table V respectively.

The graphical representation of the VNN prediction results for NN training with single intermediate layer and 2-intrmediate layers of VNNs are presented in Figs. 3, 4, 5 and 6.

Figures 3 and 4 are present the single-intermediate layer VNNs for worst and best prediction results respectively while Figures 5 and 6 present the training results for 2-intermediate layers VNNs for worst and best prediction results respectively.

The worst prediction results for both the single-intermediate layer and 2-intermediate layers VNNs are graphically presented in Figures 7 and 8 respectively.

TABLE II
ANALYSIS OF NEURON VARIATION OF THE SINGLE-INTERMEDIATE LAYER VNN,
TRAINING NETWORK USING BR MATHEMATICAL TRAINING ALGORITHM

Neuron numbers	Training Time (s)	Epoch (1000)	Coefficient of Regression (R)	Standard Deviation (SD)	Root Mean Squares Error (RMSE)	Mean Absolute Error (MAE)
4	00:01:06	197	0.8896	2.2274	3.4267	2.6041
8	00:00:22	315	0.9104	1.9490	3.1046	2.4166
12	00:00:02	1000	0.9168	1.8733	2.9950	2.3368
16	00:00:08	1000	0.9258	1.8369	2.8500	2.1791
20	00:00:06	1000	0.9444	1.6567	2.4684	1.8299
24	00:00:06	1000	0.9508	1.5914	2.3356	1.7095
28	00:00:09	1000	0.9553	1.4856	2.2170	1.6457
32	00:00:85	1000	0.9594	1.4783	2.1315	1.5356
36	00:00:08	1000	0.9620	1.4636	2.0619	1.4523
40	00:00:08	1000	0.9620	1.3425	1.9356	1.4125
44	00:00:11	1000	0.9661	1.3143	1.9295	1.3943
48	00:00:13	1000	0.9705	1.2808	1.8065	1.2740
52	00:00:18	1000	0.9728	1.2788	1.7435	1.1851
56	00:00:14	1000	0.9625	1.6055	2.0479	1.1986
60	00:00:14	1000	0.9580	1.8064	2.1527	1.1188
64	00:00:14	1000	0.9417	2.3731	2.6230	1.1937
68	00:01:16	1000	0.9390	2.4010	2.6576	1.1386
72	00:00:16	1000	0.9304	2.5721	2.8131	1.1394
76	00:00:16	1000	0.8927	3.4549	3.7083	1.3473
80	00:00:18	1000	0.8749	3.9178	4.1789	1.4542

TABLE III
ANALYSIS OF NEURON VARIATIONS OF THE 2-INTERMEDIATE LAYERS VNN, TRAINING NETWORK USING BR MATHEMATICAL TRAINING ALGORITHM

Neuron numbers	Training time (s)	Epoch (1000)	Coefficient of Regression (R)	Standard Deviation (SD)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)
[4, 8]	00:00:06	1000	0.9097	1.9126	3.1141	2.4575
[8, 12]	00:00:07	1000	0.9103	1.9243	3.1113	2.4448
[12, 16]	00:00: 21	1000	0.9468	1.5904	2.4194	1.8232
[16, 20]	00:00:58	1000	0.9531	1.5407	2.2754	1.6744
[20, 24]	00:01:38	1000	0.9264	2.0305	2.8312	1.9730
[24, 28]	00:03:04	1000	0.9182	2.2407	2.9931	1.9844
[28, 32]	00:03:20	1000	0.9191	1.8468	2.9552	2.3071
[32, 36]	00:10:41	1000	0.8433	4.3415	4.6688	2.0666
[36, 40]	02:17:06	1000	0.4515	11.6717	11.8532	2.4171

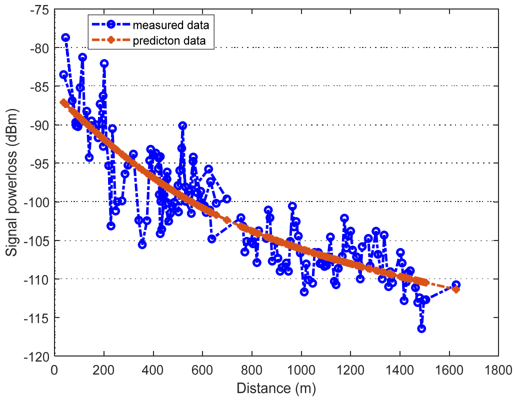

Fig. 3. First prediction result from training of single intermediate layer VNN with 4 neurons

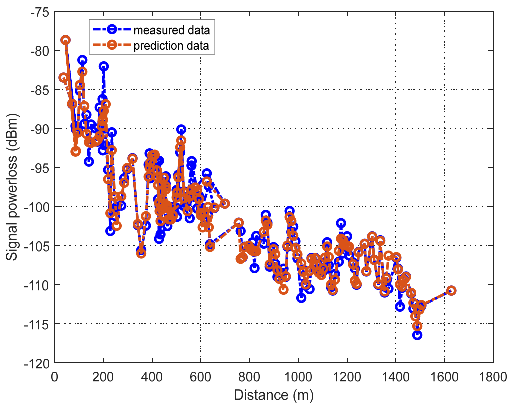

Fig. 4. Best prediction result from training of single intermediate layer VNN with 52 neurons

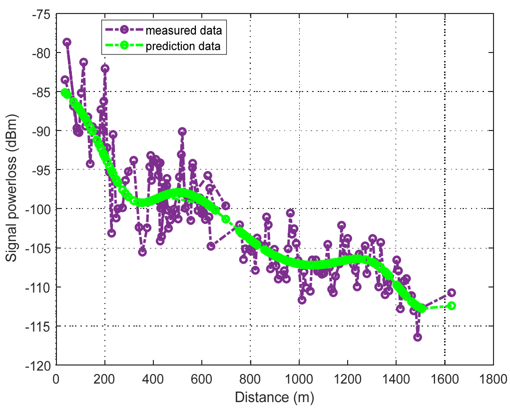

Fig. 5. First prediction result from training of 2- intermediate layers VNN with [4, 8] neurons

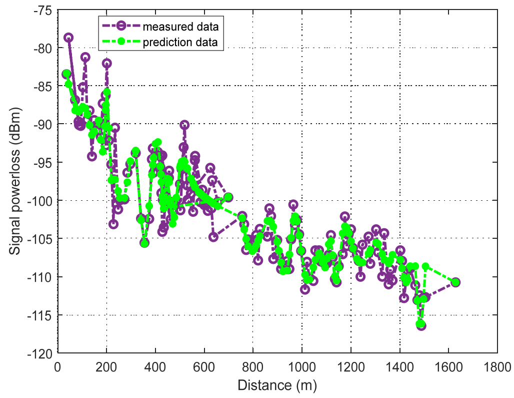

Fig. 6. Best prediction result from training of 2- intermediate layers VNN with [16, 20] neurons

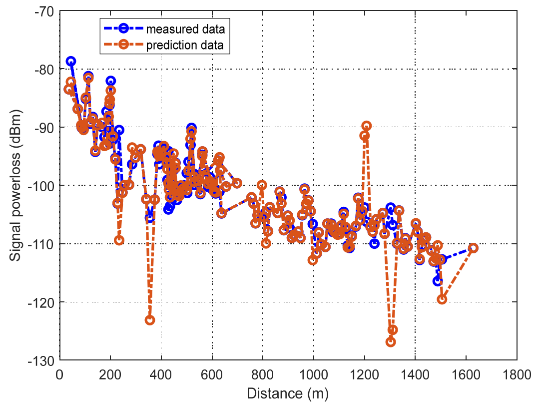

Fig. 7. Worst prediction result from training of single intermediate layers VNN with 80 neurons

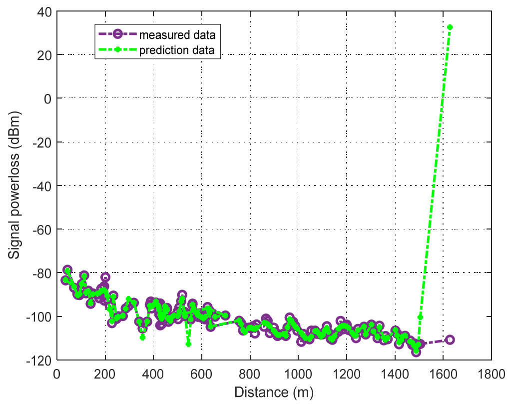

Fig. 8. Worst prediction result from training of 2- intermediate layers VNN with [36, 40] neurons

TABLE IV
TRAINING PERFORMANCE OF SINGLE-INTERMEDIATE LAYER VNN AT VARIED LEARNING RATE USING BR MATHEMATICAL TRAINING ALGORITHM

Varied Learning Rate	Epoch (1000)	Time (s)	Performance (MSE)	Regression (R)
0.002	1000	00:01:09	1.480	0.9930
0.004	1000	00:00:47	1.810	0.9912
0.008	1000	00:00:38	1.890	0.9905
0.010	1000	00:00:39	1.940	0.9846
0.012	1000	00:00:43	1.954	0.9841
0.014	1000	00:00:44	1.977	0.9833
0.016	1000	00:00:46	1.985	0.9829
0.018	1000	00:00:49	1.990	0.9815
0.020	1000	00:00:51	2.020	0.9800
0.022	1000	00:00:52	2.021	0.9790
0.024	1000	00:00:52	2.135	0.9780
0.026	1000	00:00:53	2.165	0.9750
0.028	1000	00:00:56	2.190	0.9755
0.030	1000	00:00:58	2.230	0.9780

TABLE V
TRAINING PERFORMANCE OF 2-INTERMEDIATE LAYERS VNN AT VARIED LEARNING RATE USING BR MATHEMATICAL TRAINING ALGORITHM

Varied Learning Rate	Epoch (1000)	Time (s)	Performance (MSE)	Regression (R)
0.002	1000	00:04:44	1.720	0.8810
0.004	1000	00:04:55	2.250	0.8809
0.006	1000	00:03:20	2.350	0.8750
0.008	1000	00:02:18	2.840	0.8500
0.010	1000	00:01:45	2.895	0.8460
0.012	1000	00:02:00	2.975	0.8320
0.014	1000	00:02:30	2.977	0.7820
0.016	1000	00:02:35	2.290	0.7785
0.020	1000	00:03:20	2.320	0.7560
0.022	1000	00:03:40	2.825	0.7480
0.024	1000	00:03:44	2.830	0.7440
0.026	1000	00:03:50	2. 990	0.7300
0.028	1000	00:03:55	2.995	0.6890
0.030	1000	00:04:03	2.998	0.6990

IV.1. Discussion of Experimental Results

The experimental results as recorded in the tables and graphs are discussed below.

IV.1.1. Effect of Architectural Structure of Vanilla Neural Network Models in the Prediction of Signal Power Loss

Table II describes the results from the neural network training using a single-intermediate VNN architectural structure. The single-intermediate layer neurons were sequentially varied from four (4) neurons to eighty (80) neurons during the training of the neural network utilizing BR mathematical training algorithm. There is to ascertain clearly the prediction performances exerting the various neuron numbers. The VNN gradually learns the dataset pattern on application of 4 neurons giving *R* of 0.8896, *SD* of 2.2274 and *RMSE* of 3.4267. There was steady increase in the learning capability of the VNN as the neuron numbers were increased and the training time was also very fast. On application of 52 neurons, the best performance result was recorded with *R* of 0.9728, *SD* of 1.2788 and *RMSE* of 1.7435. Thereafter, there became reduction in the learning capability and increase in training time with the worst training recorded on training the VNN with 80 neurons which gave *R* of 0.89749, *SD* of 3.9178 and *RMSE* of 4.1789.These are within the context of the considered number of neurons. The training results for the single-intermediate layer VNN are shown in Table II, Figures 3, 4 and 7 respectively for first training, best and worst trainings exerting different number of neurons. The same training exerting different variations of neurons for 2-intermediate layers VNN was carried out and the results are recorded in Table III and Figures 5, 6 and 8 for the first VNN prediction using [4, 8] neurons, best VNN prediction exerting [16, 20] neurons in the first and second intermediate layers and worst VNN prediction exerting [36, 40] neurons respectively.

The first VNN prediction for the 2-intermediate layers VNN gives *R* of 0.9077, *SD* of 1.9126 and *RMSE* of 3.1141. The best prediction result of the 2-intermediate layers VNN gives *R* of 0.9531, *SD* of 1.5407 and *RMSE* of 2.8312 while the worst prediction result gives *R* of 0.4515, *SD* of 11.6717 and *RMSE* of 11.8532. The training results show a very good prediction result for a take-off of VNN training with [4, 8] neurons giving *R* of 0.9097 with very fast training time, thereafter, after the best prediction result was recorded with training with [16, 20] neurons with *R* of 0.9531, there become rapid decline in the prediction ability of the 2-intermediate layers VNN with very increase in training time giving the worst prediction result within the context of consideration as *R* of 0.4515, *SD* of 11.6717 and *RMSE* of 11.8532. The graphical results for prediction exerting single-intermediate layer VNN is represented with measured data represented using blue colour and prediction data represented using orange colour while the graphical results for the 2-intermediate layers VNN is presented with measured dataset represented with purple colour and prediction with 2-intermediate layers VNN represented with green colour.

IV.1.2. Effect of Learning Rate in Vanilla Neural Networks for the Prediction of Signal Power Loss Prediction

Different values of learning rates ranging from 0.002 to 0.030 were exerted in examining the efficiency of both single and 2-intermediate layers VNN architectural Structures during neural network training for signal power loss prediction. Mean Squared Error (*MSE*) and Regression (*R*) on the training dataset during network training of the 2-VNN models are the two metrics used in the analysis of their performances. A well trained VNN gives a low *MSE* i.e. it's close to zero shows that the actual output optimally predicted the desired output making both closer to one another. Regression demonstrates the strength of the relationship between the actual output and the desired output. A regression values closest to +1 shows strong relationship between the actual output and the desired output. Table IV and Table V show the results of the training conducted on a single-intermediate layer VNN and 2- intermediate layers VNN respectively while training with different values of learning rates using BR mathematical training algorithm and exerting early stopping method for efficient network generalization.

On application of different variations of learning rates ranging from 0.002 to 0.030, the results show the lowest *MSE* and highest *R* at 0.002 learning rate with *MSE* of 1.480 and *R* of 0.9930 from training of single-intermediate layer VNN and *MSE* of 1.720 and regression of 0.8810 from training of 2-intermediate layers VNN model. The training time at 0.002 learning rate for single-intermediate layer VNN was 00:01:09 while it was 00:04:44 for training at the same learning rate of 0.002 for VNN with 2-intermediate layers VNN.

This shows that in comparison of the minimal learning rates with the least *MSE* and the highest *R* for both single-intermediate layer VNN and 2-intermediate layers VNN models, the single-intermediate layer VNN model demonstrates better performance in prediction of signal power loss than the 2-intermediate layers VNN model.

The best trained intermediate layers of VNNs are highlighted in red colour in Table IV and Table V respectively where it can be seen that the *MSE* for the single-intermediate layer VNN is lower than that of the 2-intermediate layers VNN and the *R* of the single-intermediate layer VNN is higher than that of the 2-intermediate layers VNN. This approves that for an efficient neural network training and signal power loss prediction, a well- developed single-intermediate layer VNN architectural structure with adequate neuron number gives efficient and optimal prediction results.

V. Conclusion

This work designed investigated two VNN models: a Single-intermediate layer VNN and a 2-intermediate layers VNN architectural network models and examines their distinct performances and effect in signal power loss prediction using measured data collected from a LTE micro-cell built-up environment. The effect of learning rate hyper-parameters on the two VNN models and the impact of different numbers of neurons in their intermediate layers while training them with BR mathematical training algorithm were investigated.

The modeling of the VNNs were carried out such that the neurons of the single-intermediate layer VNN were repeatedly increased in fours (4) up to eighty (80) neurons to ensure robust training and ascertain clearly their predictive performances during neural network training. The 2-intermediate layers VNN model was also trained while varying the neuron numbers in the first and second intermediate layers respectively. The first intermediate layer of the 2-intermediate layers VNN was repeatedly increased in fours (4) while the second intermediate layer of the VNN was repeatedly increased in eights (8).

The results from best training performances for the Single-intermediate layer VNN and the 2-intermediate layers VNN shows best training performance on application of 52 neurons in the single-intermediate layer of VNN with *R* of 0.9728, *SD* of 1.2788 and *RMSE* of 1.7435 and prediction declined with worst prediction results of *R* of 0.8749, *SD* of 3.9178 and *RMSE* of 4.1789 on application of 80 neurons, within the training framework.

However, the 2-intermediate layers VNN gives the best training performance of *R* as 0.9531, *SD* of 1.5407 and *RMSE* of 2.2754 on training the network with [16, 20] and on training the network with [36, 40], R becomes 0.4515, *SD* becomes 11.6717 and *RMSE* becomes 11.8532. the training time required for the network training rapidly increased up to 00:00:58 to get best training performance on training with [16, 20] neurons and continuously increased up to 00:17:16 on training with [36, 40] neurons with low *R* of 0.4515. In comparison with training with single-intermediate layer VNN, where the best prediction at 52 neurons was trained at 00:00:18 training time with *R* of 0.9728 and at 80 neurons, the training time was still 00:00:18, it demonstrates swift network learning, training and prediction exerting single-intermediate layer VNN in comparison to 2-intermediate layers VNN.

Also, at learning small learning rate of 0.002, the Single intermediate layer VNN recorded *R* of 0.9930 and 0.8810 for 2-intermediate layer VNN respectively and as the learning rate increased up to 0.030, *R* of Single-intermediate layer VNN becomes 0.9780 while that of 2-intermediate layers VNN becomes 0.6990. This still buttress expeditious performance of Single-intermediate layer VNN in comparison to 2-intermediate layers VNN in the learning, training and prediction of signal power pattern. Future works will investigate the impacts of other hyper-parameters and other ANN architectures in prediction of signal power loss using measurement data for effective prediction. Also, the role of path loss modelling in 6G and how path loss or channel modelling can help challenges in 6G in will be considered in future works.

References

[1] J. Isabona, A.L. Imoize, S. Ojo, O. Karunwi, Y. Kim, L. Cheng-Chi, L. Chun-Ta, “Development of a Multilayer Perceptron Neural Network for Optimal Predictive Modeling in Urban Microcellular Radio Environments” *Appl.Sci.* vol. 12, pp.1-27, 2022.

[2] Ebhota, V., Isabona, J., Srivastava, V., Effect of Learning Rate on GRNN and MLP for the Prediction of Signal Power Loss in Microcell Sub-Urban Environment, (2019) *International Journal on Communications Antenna and Propagation (IRECAP)*, 9 (1), pp. 36-45. doi:https://doi.org/10.15866/irecap.v9i1.15329

[3] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Modelling, Simulation and Analysis of Signal Pathloss for 4G Cellular Network Planning" *Journal of Engineering and Applied Sciences*, vol. 9, pp. 1907-1918, 2018.

[4] M.M. Rathore, A. Paul, S. Rho, M. Khan,. S. Vimal,. S.A.Shah,” Smart traffic control: Identifying driving-violations using fog devices with vehicular cameras in smart cities”. *Sustain Cities*

Soc., vol.71, pp.102986, 2021.
[5] H. Tataria, K. Haneda, A.F. Molisch, M. Shafi, F.Tufvesson,"Standardization of Propagation Models for Terrestrial Cellular Systems: A Historical Perspective". *Int. J. Wirel. Inf. Networks*, vol.28, pp.20–44, 202.
[6] B.Y. Hanci, I.H. Cavdar,"Mobile radio propagation measurements and tuning the path loss model in urban areas at GSM-900 bandinIstanbul-Turkey". In *Proceedings of the IEEE 60th VehicularTechnologyConference, 2004,* VTC2004-Fall, Los Angeles, CA, USA, vol. 1, pp. 139–143, 26–29 September 2004.
[7] S. Ojo, A.Imoize, D. Alienyi, "Radial basis function neural network path loss prediction model for LTE networks in multitransmitter signal propagation environments". *Int. J. Commun. Syst*. vol. 34 2021.
[8] V. C. Ebhota, V. M. Srivastava, "Performance Analysis of Learning Rate Parameter on Prediction of Signal Power Loss for Network Optimization and Better Generalization" *Springer Wireless Personal Communications,* vol.119. pp.1111-1128. 2021.
[9] J, Isabona, "Wavelet Generalized Regression Neural Network Approach for Robust Field Strength Prediction*" Wirel. Pers. Commun*., vol. 114, pp. 3635–3653, 2020.
[10] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Environment-Adaptation Based Hybrid Neural Network Predictor for Signal Propagation Loss Prediction in Cluttered and Open Urban Microcells*" Wireless Personal Communications,*. pp.1-14, 2018.
[11] B. Turan,. S Coleri, "Machine learning based channel modeling for vehicular visible light communication". *IEEE Trans. Veh. Technol.,* vol.70, pp. 9659–9672, 2021.
[12] Y .Zhang, J. Wen, G Yang, Z He,. J. Wang, "Path loss prediction based on machine learning: Principle, method, and data expansion". *Appl. Sci*., vol 9, pp.1908-1920, 2019.
[13] S.P Sotiroudis, S.K Goudos, K.A.Gotsis, K.Siakavara, J.N. Sahalos, "Modeling by optimal artificial neural networks the prediction of propagation pathloss in urban environments", *In Proceedings of the 2013 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, Turin, Italy, pp. 599–602, 9-13 September 2913.
[14] J. Isabona, A.I. Osaigbovo, "Investigating predictive capabilities of RBFNN, MLPNN and GRNN models for LTE cellular network radio signal power datasets". *FUOYE J. Eng. Technol*. vol. 4, pp. 155-159, 2019.
[15] A.L.Imoize,O.A.Ajibola,T.R.Oyedare,J.O.Ogbebor,S.O.Ajose,".DevelopmentofanEnergy Efficient Wireless Sensor Network Model for Perimeter Surveillance" . *Int. J. Electr. Eng. Appl. Sci*., vol. pp. 1–17, 2021.
[16] A. L. Imoize, "*Analysis of Propagation Models for Mobile Radio Reception at 1800MHz*." Master's Thesis, University of Lagos, Lagos, Nigeria, 2011.
[17] S. Haykin, "*Neural networks*". A comprehensive foundation, 2nd ed.: Prentice Hall, 1999.
[18] Ebhota, V., Isabona, J., Srivastava, V., Investigation and Comparison of Generalization Ability of Multi-Layer Perceptron and Radial Basis Function Artificial Neural Networks for Signal Power Loss Prediction, (2019) *International Journal on Communications Antenna and Propagation (IRECAP)*, 9 (1), pp. 46-54.
doi:https://doi.org/10.15866/irecap.v9i1.15330
[19] A. R. Barron, R. L. Barron, and E. J. Wegman, "Statistical learning networks: A unifying view," *Computer Science and Statistics: Proceedings of the 20th Symposium on the Interface* pp. 192-203, 1992.
[20] M. Hassoun, *Fundamentals of artificial neural networks*. Massachusetts: MITPress, 1995.
[21] M. Ayadi, A. B. Zineb, and S. A. Tabbane, "UHF path loss model using learning machine for heterogeneous networks," *IEEE Trans. Antennas Propagation*, vol. 65, pp. 3675-3683, 2017.
[22] M. K. S. Alsmadi, K. B. Omar, and S. A. Noah, "Back propagation algorithm: The best algorithm among the multi-layer perceptron algorithm," *International Journal of Computer Science and Network Security (IJCSNS),* vol. 9, pp. 378 - 383, 2009.
[23] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Improved adaptive signal power loss prediction using combined vector smoothing and neural network approach," *Progress in Electromagnetic Research C,* vol. 82, pp. 155-169, 2018.
[24] A. B. Zineb and M. Ayadi, "A multi-wall and multi-frequency indoor path loss prediction model using artificial neural networks," *Arabian Journal of Science and Engineering,* vol. 41, pp. 987-996, 2016.
[25] S. Haykin, *Neural networks: A comprehensive foundation*, 2nd ed. United States of America: Prentice Hall, 2001.
[26] M. T. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transaction on Neural Networks,* vol. 5, pp. 989-993, 1994.
[27] M. T. Hagan and H. B. Demuth, "Neural networks for control," in *Proceedings of the 1999 American Control Conference*, San Diego, USA, 1999, pp. 1642-1656.
[28] M. T. Hagan, H. B. Demuth, and M. H. Beale, "*Neural network design"*. Boston: PWS Publishing, 1996.
[29] D. Graupe, "Principles of artificial neural networks" *Advanced series on circuits and systems*, 2nd ed. vol. 6. Singapore World Scientific Publishing, 2007.
[30] H. White, "Learning in neural networks: A statistical perspective," *Neural Computation,* vol. 1, pp. 425- 464, 1989.
[31] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Signal power loss prediction based on artificial neural networks in microcell environment," 3rd *IEEE International Conference on Electro-Technology for National Development,* Owerri, Nigeria, 2017.
[32] R. Haelterman, "*Analytical study of the least squares quasi-Newton method for interaction problems*," PhD, Ghent University, 2009.
[33] Y. Hao and B. M. Wilamowski, Levenberg-marquardt training*. The Industrial Electronics Handbook* vol. 5, 2011.
[34] C. Kanzow, N. Yamashita, and M. Fukushima, "Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints*," Journal of Clinical and Analytical Medicine*, vol. 172, pp. 375-97, 2004.
[35] D. J. C. MacKay, "Bayesian interpolation*," Neural Computation,* vol. 4, pp. 415-447, 1992.
[36] J. Isabona and V. M. Srivastava, "A neural network based model for signal coverage propagation loss prediction in urban radio communication environment," *International journal of applied engineering research,* vol. 11, pp. 11002-11008, 2016.
[37] A.L. Imoize, K . Orolu, A.A. Atayero, " Analysis of key performance indicators of a 4G LTE network based on experimental data obtained from a densely populated smart city". *Data Br*. vol. 29, pp. 105-304, 2020.
[38] V.C. Ebhota, J. Isabona, V.M. Srivastava, "Environment-Adaptation Based Hybrid Neural Network Predictor for Signal Propagation Loss Prediction in Cluttered and Open Urban Microcells*". Wirel. Pers. Commun*, vol. 104, pp. 935–948, 2019.
[39] J. Isabona, M. Babalola, " Statistical Tuning of Walfisch-Bertoni Pathloss Model based on Building and Street Geometry Parameters in Built-up Terrains". *Am. J. Phys. Appl.* vol. 1, pp. 10-16, 2013.
[40] B.Y. Hanci, I.H. Cavdar, "Mobile radio propagation measurements and tuning the path loss model in urban areas at GSM-900 bandinIstanbul-Turkey". In *Proceedings of the IEEE 60th Vehicular Technology Conference,* 2004, VTC2004- Fall, LosAngeles, CA, USA, vol. 1, pp. 139–143, 26–29 September 2004.
[41] S. Ojo, M. Akkaya, J.C. Sopuru, "An ensemble machine learning approach for enhanced path loss predictions for 4G LTE wireless networks". *Int. J. Commun. Syst*. vol. 35, e5101, 2022.
[42] V.C Ebhota, J Isabona,.; V.M. Srivastava, "Base line knowledge on propagation modelling and prediction techniques in wireless communication networks*". J. Eng. Appl. Sci*., vol. 13, pp. 1919–1934, 2018.
[43] M. A.Salman,. S.I. Popoola, N. Faruk, N.T. Surajudeen-Bakinde, A.A. Oloyede, L.A. Olawoyin, "Adaptive Neuro-Fuzzy model for path loss prediction in the VHF band"*. In Proceedings of the 2017 International Conference on Computing Networking and Informatics (ICCNI),* Lagos, Nigeria, pp. 1–6, 29–31 October 2017.
[44] S.Aldossari, K.Chen, "Predicting the pathloss of wireless channel models using machine learning techniques in mmwave urban communications, "In *Proceedings of the 2019 22nd International*

Symposium on Wireless Personal Multimedia Communications (WPMC)", pp. 1–6, Lisbon, Portugal, 24–27 November 2019.
[45] O.Ahmadien,. H.F. Ates, T. Baykas, B.K. Gunturk, *"*Predicting Path Loss Distribution of an Area From Satellite Images Using Deep Learning*". IEEE Access*, vol. 8, pp. 64982–64991, 2020.
[46] H.F. Ates, S.M. Hashir, T. Baykas, B.K. Gunturk, "Path loss exponent and shadowing factor prediction from satellite images using deep learning". *IEEE Access*, vol. 7, pp. 101366–101375, 2019.
[47] J .Huang, Y. Cao, X. Raimundo, A. Cheema, S. Salous, "Rain statistics investigation and rain attenuation modeling for millimeter wave short-range fixed links". *IEEE Access*, vol. 7, pp. 156110–156120, 2019.
[48] Y. Zhang, J. Wen, G. Yang,.; Z .He, J. Wang, "Path loss prediction based on machine learning: Principle, method, and data expansion". *Appl. Sc*i. vol. pp. 19089, 2019.
[49] D.C, Nguyen, P . Cheng, M. Ding, D. Lopez-Perez, P.N.Pathirana, J. Li.. A. Seneviratne, Y. Li, H.V. Poor, "Enabling AI in future wireless networks: A data life cycle perspective". *IEEE Commun. Surv. Tutor*., vol. 23, pp. 553–595, 2020.
[50] G.P. Ferreira, L.J. Matos, J.M.M. Silva, "Improvement of outdoor signal strength prediction in UHF band by artificial neural network" *IEEE Trans. Antennas Propag*., vol.64, pp.5404–5410, 2016.
[51] H.; Singh, S. Gupta, C. Dhawan, A. Mishra, "Path loss prediction in smart campus environment: Machine learning-based approaches". In *Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, 25–28 May 2020; pp. 1–5.
[52] J.Thrane, D. Zibar, H.L Christiansen, "Model-aided deep learning method for path loss prediction in mobile communication systems at 2.6 GHz". *IEEE Access*, vol.8, pp.7925–7936, 2020.
[53] A. Shrestha, A. Mahmood, "Review of deep learning algorithms and architectures". *IEEEAccess*, vol. 7, pp.53040–53065, 2019.
[54] C. Nguyen, A.A. Cheema, "A deep neural network-based multi-frequency path loss prediction model from 0.8 GHz to 70 GHz". *Sensors* vol. 21, pp5100, 2021.
[55] U. Challita, L. Dong, W. Saad, "Deep learning for proactive resource allocation in LTE-U networks". In *Proceedings of the European Wireless Technology Conference*, Dresden, Germany, 15 May 2017.
[56] W. Song, F. Zeng, J. Hu, Z. Wang, X. Mao, "An unsupervised-learning-based method for multi-hop wireless broadcast relay selection in urban vehicular networks". In *Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring),* Sydney, NSW, Australia, 4–7 June 2017; pp. 1–5.
[57] S. Ghahramani, *Fundamentals of probability*, 2nd Ed. New Jersey: Prentice Hall, 2000.
[58] A. Buda and A. Jarynowski, "Life time of correlations and its applications," *Wydawnictwo Niezależne,* pp. 5-21, 2010.

Authors' information

Department of Electrical and Electronic Engineering Science,
University of Johannesburg, P.O. Box 524, Auckland Park, 2006,
Johannesburg, South Africa.
E-mails: engrchikaugwu@yahoo.com
tshongwe@uj.ac.za

Dr. **Virginia C. Ebhota** received Bachelor of Engineering degree in Electronic Engineering from University of Nigeria, Nsukka (UNN), Master of Engineering from Enugu State University of Science and Technology, Enugu, Nigeria and a Doctoral degree in Electronic Engineering from Kwazulu-Natal, Durban, South Africa. Presently, she is a post-doctoral research fellow with the department of Electrical and Electronic Engineering, University of Johannesburg, South Africa. She has published over seventeen (17) peer-reviewed articles in Journals and IEEE Conferences. Her major research area is system modeling using Artificial Intelligence (AI) such as Artificial Neural networks, etc.

Thokozani Shongwe received the B. Eng degree in Electronic engineering from the University of Swaziland, Swaziland, in 2004 and the M. Eng degree in Telecommunications Engineering from the University of the Witwatersrand, South Africa, in 2006 and the D. Eng degree from the University of Johannesburg, South Africa, in 2014. He is currently an Associate Professor of Telecommunications and the Head of the School of Electrical and Electronic Engineering at the University of Johannesburg. He is a recipient of the 2014 University of Johannesburg Global Excellence Stature (GES) award, which was awarded to him to carry out his postdoctoral research at the University of Johannesburg. In 2016, Prof T. Shongwe was a recipient of the TWAS-DFG Cooperation Visits Programme funding to do research in Germany.

Other awards that he has received in the past are: the post-graduate merit award scholarship to pursue his master's degree at the University of the Witwatersrand in 2005, which is awarded on a merit basis; In the year 2012, He (and his co-authors) received an award of the best student paper at the IEEE ISPLC 2012 (power line communications conference) in Beijing, China. His research fields are in Digital Communications and Error Correcting Coding. His research interests are in power-line communications; cognitive radio; smart grid; visible light communications; Machine Learning and Artificial Intelligence.

He is the co-founder of a research group at the University of Johannesburg called Artificial Intelligence for Electrical Engineering Applications (AI for EE Applications). This research group is currently composed of, five staff members; two postdoctoral researchers working in the fields of Communications and Machine Learning; five Doctoral students and ten master's degree students working in the fields of power line communications, visible light communications, application of ML in PLC, power systems, agriculture and object tracking.