Enhanced Error Reduction of Signal Power Loss During Electromagnetic Propagation: Architectural Composition and Learning Rate Selection

Virginia C. Ebhota and Viranjay M. Srivastava
Department of Electronic Engineering, Howard College, University of KwaZulu-Natal, Durban-4041, South Africa
Email: ebhota@ieee.org, viranjay@ieee.org

Abstract**—**This research work analyses the effect of the architectural composition of Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN) combined with the effect of the learning rate for effective prediction of signal power loss during electromagnetic signal propagation. A single hidden layer and two hidden layers of MLP ANN have been considered. Different configurations of the neural network architecture ranging from 4 to 100 for both MLP networks have been analyzed. The required hidden layer neurons for optimal training of a single layer multi-layer network were 40 neurons with 0.99670 coefficient of correlation and 1.28020 standard deviations, while [68 72] trained two hidden layers multi-layer perceptron with 0.98880 coefficient of correlation and standard deviation of 1.42820. Different learning rates were also adopted for the network training. The results further validate better MLP neural network training for signal power loss prediction using single-layer perceptron network compared to two hidden layers perceptron network with the coefficient of correlation of 0.99670 for single-layer network and 0.9888 for two hidden layers network. Furthermore, the learning rate of 0.003 shows the best training capability with lower mean squared error and higher training regression compared to other values of learning rate used for both single layer and two hidden layers perceptron MLP networks.

Index Terms**—**Architectural composition, Learning rate, Error reduction, Signal power loss, Bayesian Regularization, MLP

I. Introduction

Over the years, different geospatial and heuristic methods have been developed for enhanced reduction of signal power loss error during electromagnetic transmission [1]-[3]. Nevertheless, the efficiency of these techniques, such as empirical models and deterministic models, have been verified experimentally with empirical models. However, popular due to their simplicity, introduce high errors during prediction while deterministic models are complex in operation [4]. There is a need for profound knowledge of the behavior of electromagnetic signal propagation in the practical wireless channel for effective radio access network planning and deployment of radio access networks in different environments. The direction and magnitude of the electromagnetic signal in feasible wireless channels are mainly random and extremely unpredictable [5]-[7]. Thus, an understanding of this phenomenon in the radio access network is required to assure excellent quality of service and high data transmission rates [1], [3], [8]-[10].

Radio signals are reflected when they collide with objects whose dimensions are large relative to the wavelength of the radiated signal. At the same time, diffraction occurs because of transmission paths being obstructed by large substances that cause the bending of signals during propagation [11], [12]. Electromagnetic signals also encounter scattering during propagation due to the object's size being less than the radio-signal wavelength. Thus, radiated electromagnetic signals are reflected in different directions. Scattering may also result due to precipitation, suspensions as well as dust particles [1], [5], [13]. The atmospheric propagation environment conditions mainly influence the propagation of electromagnetic signals.

Electromagnetic signals of higher frequencies with a few millimeter wavelengths easily get attenuated as the size of the transmitted wavelength tends towards atmospheric size [14]-[16]. Various copies of transmitted radio signals arrive at the receiver via propagation mechanism by multipath propagation in real-world propagation environments. This causes signal fading at the receiving end [1], [14]. In view of a situation where the received signal strength magnitude constantly changes within a short duration given that there is a relatively unchanged distance, such signal attenuation is small. The received signal scale significantly reduces for large-scale fading at the increase in distance. This concept is known as path-loss or loss in signal power [17]-[20].

Various propagation models have been developed for path-loss estimations under different propagation scenarios [21]. Conventionally, each propagation model can either be empirical, deterministic or semi-deterministic model w.r.t. the model technique applied. On the availability of site-specific data of the propagation terrain, deterministic models' application mainly guarantees accuracy in prediction. However, a significant challenge with the use of deterministic models is their computational complexity. They require numerous input information which may not easily be obtained [22]. In

Manuscript received March 1, 2021; revised September 15, 2021.
Corresponding author email: ebhota@ieee.org
doi:10.12720/jcm.16.10.450-456

divergent, empirical models require less computational resources for their implementation. Though easy to use, their prediction accuracy is low in comparison to the deterministic model. The effectiveness of these models has been tested in diverse environments.

Researchers have recently successfully used ANN to predict losses during electromagnetic signal transmission [23]. Artificial neural networks are adaptive statistical tools that model almost the same way as the biological nervous system [24]-[26]. Due to its accuracy and litheness in adapting to different environments, various notable characteristics (for example: the ability to learn, generalization of patterns, and model nonlinear functions) application of ANN algorithms emerged for problem-solving [27], [28].

This research work carried out a comprehensive analysis of the impact of the architectural composition of MLP-ANN in unification with the effect of variable learning rates in the prediction of signal power loss using the detailed dataset from a micro-cellular environment. As a result, various variable learning rates were adopted in training a single hidden layer MLP network and two hidden layer MLP network, and their prediction performances were made. This is quite different from the author's previous work, where the effect of learning rates was only realized using a single hidden layer MLP network [7], [8], [11], [29].

In this research work, performance mean squared error and training regression have been used to analyze the required learning rate for network training. Coefficient of correlation and standard deviation have been used to analyze the adequate required architectural composition of the MLP network. Early stopping technique and Bayesian Regularization (BR) algorithm have been employed for improved network generalization during training. The required hidden layer neurons for optimal single-layer multi-layer network training were 40 neurons with a 0.99670 coefficient of correlation and 1.28020 standard deviations while trained with two hidden layers ML gave a 0.98880 coefficient of correlation and standard deviation 1.42820. Further comparison between the architectural composition required for adequate training of MLP with a single hidden layer and two hidden layers reveals better prediction of signal power loss using single layer MLP compared to two hidden layers MLP network.

Learning rates of 0.003 to 0.042 have been adopted for the network training. The results further validate better MLP neural network training using a single-layer perceptron network compared to two hidden layers with a coefficient of correlation of 0.99670 for single layer MLP and 0.9888 for two hidden layers MLP network. The learning rate of 0.003 shows the best training capability with lower mean squared error and higher training regression compared to other values of learning rate used for both single layer and two hidden layers MLP networks. However, increased convergence training time was required for training the neural network at 0.003 learning rate. This work has been organized as follows: Section II describes the architectural composition implication of MLP-ANN. The section III analyses the results of this work. Finally, Section IV concludes the work and recommends the future aspects.

II. Architectural Composition Implication of MLP-ANN and Selection of Training Set with Layer/Neuron Variation

Universal approximation theorem states that three-layered MLP-ANN approximates practically any nonlinear function [30], [31]. However, there is no specification of the size (number) of the hidden layer of neurons for specified problem complexity. Thus, the actual required number of the hidden layers and neurons has remained an open problem for effective neural network training/learning. For proper training of MLP-ANN, training requires a set of examples of network behavior i.e., the network input and the target outputs. Therefore, the values of signal power measured at different points of the considered micro-cell environment over a distance of *800 m* were employed. During the training process, MLP-ANN crams the relationship between the location of measurement points, i.e., the measured data, the link between the input vector, and the target vector for the given environment. The number of layers for the proposed MLP-ANN model is experimentally determined.

The early stopping method and Bayesian Regularization were applied during network training to improve network generalization [32]. Bayesian Regularization algorithm is applied in weight update during network training in agreement with Levenberg Marquardt (LM) algorithm and has demonstrated near better training by linear permutation of squared error and weight variables [33]-[37]. The algorithm uses back-propagation and modifies all variables in accordance with the function approximation method.

There is a need for appropriate training set selection from the real propagation path from which the MLP-ANN will learn to calculate received power, which is the most crucial factor in the training phase [38]. For training optimization, the training set involves measurement data from different routes with different propagation characteristics such as reflection, diffraction, reflection, direct rays, etc. The selected routes also include received positions that show various ranges of the input parameter. Hence the network can learn to behave in different situations and thus make an appropriate generalization on application to new cases.

The first important step in the training process is appropriate measurement points characterization in the training route according to their type of dominant path. The choice of training routes was a planned process, and a balanced number of measured data points representing different propagation conditions were supplied. Two thousand and Ten measurement data were recorded, each with different received signal power. The neural network was trained using different numbers of neurons in the hidden layers that vary from 4 to 112 neurons to extract better training points. The neuron variations are shown in Table I and Table II for training with *one* hidden layer and *two* hidden layers, respectively. More detailed analysis have been explained by the authors in [19], [24], [28], [29].

For training optimization, the training set involves measurement data from different routes with different propagation characteristics such as reflection, diffraction, reflection, direct rays, etc. Moreover, the selected routes also include received positions that show various ranges of the input parameter. Hence the network can learn to behave in different situations and make an appropriate generalization on application to new cases.

From Table I of training of MLP-ANN using a single hidden layer, there were variations of the hidden layer neurons from 4 to 100 numbers during the neural network training using BR training algorithm. On application of 40 neuron numbers in the hidden layer, the highest correlation coefficient is 0.99670, and a standard deviation of 1.28020 was recorded as the least values showing minimal deviation of the training dataset from the actual dataset, thus high prediction accuracy. This indicates the closeness of the prediction of the training dataset to the actual dataset. The training time and standard deviation for applying the different hidden layer neurons were recorded as shown in Table II. As the neuron numbers in the MLP-ANN hidden layer increase, there is a rapid decrease in the correlation coefficient and an increase in the standard deviation. This implies that for signal power loss prediction using single layer MLP-ANN, there is difficulty predicting the losses in signal power as the network architecture becomes complex.

TABLE I. ANALYSIS OF NEURON VARIATION IN MLP-ANN WITH ONE HIDDEN LAYER

Neuron number	Training time (s)	Epoch (1000)	Coefficient of Correlation (r)	Standard Deviation (SD)
4	00:00:01	89	0.80550	2.10440
8	00:00:06	164	0.82290	1.96000
12	00:00:14	330	0.90000	1.89000
16	00:00:22	521	0.92430	1.86060
20	00:00:24	664	0.92460	1.60850
24	00:00:27	1000	0.93600	1.60010
28	00:00:29	1000	0.93850	1.59800
32	00:00:32	1000	0.94800	1.55060
36	00:00:36	1000	0.96010	1.48030
*40	00:00:57	1000	0.99670	1.28020
44	00:00:39	1000	0.99430	1.29900
48	00:00:39	1000	0.96220	1.36040
52	00:00:44	1000	0.95589	1.89010
56	00:00:46	1000	0.94300	2.23050
60	00:00:47	1000	0.94220	2.60400
64	00:00:49	1000	0.92800	2.86900
68	00:01:01	1000	0.92100	2.94540
72	00:01:04	1000	0.90100	3.20300
76	00:01:10	1000	0.86890	3.60800
80	00:01:11	1000	0.84000	3.90910
84	00:01:18	1000	0.83900	3.91210
88	00:04:24	1000	0.83440	3.93300
92	00:06:28	1000	0.83100	3.98000
96	00:18:28	1000	0.82000	3.96700
100	00:19:30	1000	0.70430	4.12090

TABLE II. ANALYSIS OF NEURON VARIATION IN MLP-ANN WITH TWO HIDDEN LAYERS

Neuron number	Training time	Epoch (1000)	Coefficient of Correlation (r)	Standard Deviation (SD)
[4 8]	00:00:31	1000	0.81840	1.92010
[12 16]	00:00:38	1000	0.86600	1.97210
[20 24]	00:56	1000	0.91420	1.97340
[28 32]	00:01:10	1000	0.94400	1.95110
[36 40]	00:09:10	1000	0.94820	2.74040
[44 48]	00:10:42	1000	0.94490	2.63360
[52 56]	00:11:11	1000	0.95180	1.91440
[60 64]	01:10:10	1000	0.95310	1.94800
*[68 72]	02:50:00	1000	0.98880	1.42820
[76 80]	03:02:15	1000	0.94320	1.91900
[84 88]	03:14:23	116	0.91600	1.96040
[92 96]	01:20:33	112	0.91300	2.24040
[100 104]	01:30:40	102	0.88900	2.89020

III. RESULTS ANALYSIS AND DISCUSSIONS

In Table II, two hidden layers were utilized for the MLP-ANN training while varying the hidden layer neurons from [4 8] to [100 104] in the first and second hidden layers, respectively. The neuron variation with [68 72], i.e. 68 neurons in the first hidden layer and 72 neurons in the second hidden layer, gives the highest correlation coefficient of correlation of 0.98880 and the least standard deviation of 1.42820. As the network gets more complex by increasing the number of the hidden layer neurons, the coefficient of correlation decreases rapidly while the standard deviation increases.

Further comparison between prediction performance of MLP-ANN with one hidden layer and two hidden layers, respectively, show a better network training using one hidden layer as the coefficient of correlation recorded is 0.99670 in comparison to 0.98880 with two hidden layers and standard deviation 1.28020 of the MLP-ANN with a single hidden layer in comparison 1.42820 of MLP with two hidden layers. The best prediction values are highlighted with an asterisk (*) in Table I and Table II.

A. Effect of the Learning Rate for Neural Network Optimization During Network Training

Different learning rates were applied for neural network optimization during network training to examine the efficiency of the MLP-ANN in signal prediction. The impact of the learning rates during the network training was examined. A comparison of the results from the different learning rate values was made using the Mean Squared Error (MSE) and the regression on the data applied for training. The most well-trained network gave minimal MSE, showing its closeness to zero. The implication is that the desired output and the neural network training set output become closer. Also, the network regression that indicates the strength of the relationship between the training dataset i.e. the dependent variable, and series of changing variables i.e. the learning rates, were considered. At a regression point closer to +1, the training, testing, and validation performances were considered excellent.

The obtained MSE and regression result from MLP-ANN training on applying different learning rates are shown in Table III and Table IV, respectively. Learning rates from 0.003 to 0.042 (with gapping 0.03) were used in training the neural network using the Bayesian Regularization training algorithm. The training results show that at a learning rate of 0.003, the MSE was very minimal, and the training regressions, which encompass the training, testing, and validation of the training dataset, i.e. it is closest to +1 at 0.03 learning rate. These are highlighted with an asterisk (*) in the following tables, respectively.

The results from Table III and Table IV show the performance results of MSE and training regression of the dataset neural network training on the application of early stopping training technique using BR training algorithm for single layer MLP-ANN. The results demonstrate that training the neural network with a small learning rate of 0.003 gives the least MSE and the highest training regression. Furthermore, the MSE increases as the learning rate increases to 0.042. At the same time, the training regression gradually decreases on the increase of learning rate showing a substantial deviation of the actual dataset from the training dataset.

TABLE III. TRAINING PERFORMANCE AT DIFFERENT LEARNING RATE USING BR ALGORITHM FOR SINGLE HIDDEN LAYER MLP-ANN

Learning Rate	Epoch (Iteration) Maximum= 1000	Time (seconds)	Performance (MSE)	Validation check
*0.003	1000	00:01:11	1.610	0
0.006	1000	00:00:47	1.920	0
0.009	1000	00:00:39	1.940	0
0.012	1000	00:00:32	1.960	0
0.015	1000	00:00:41	1.965	0
0.018	1000	00:00:41	1.970	0
0.021	1000	00:00:46	1.977	0
0.024	1000	00:00:46	1.990	0
0.027	1000	00:00:47	2.020	0
0.030	1000	00:00:47	2.025	0
0.033	1000	00:00:51	2.145	0
0.036	1000	00:00:53	2.175	0
0.039	1000	00:00:53	2.200	0
0.042	1000	00:00:56	2.240	0

TABLE IV. REGRESSION AT DIFFERENT LEARNING RATES USING BR TRAINING ALGORITHM FOR SINGLE HIDDEN LAYERS MLP-ANN

Learning Rate	Training	Test	Validation	All
*0.003	0.9981	0.9901	0.9985	0.9965
0.006	0.9930	0.9921	0.9982	0.9932
0.009	0.9926	0.9885	0.9963	0.9911
0.012	0.9922	0.9846	0.9890	0.9886
0.015	0.9918	0.9820	0.9886	0.9845
0.018	0.9916	0.9816	0.9872	0.9832
0.021	0.9890	0.9780	0.9865	0.9827
0.024	0.9886	0.9772	0.9861	0.9824
0.027	0.9881	0.9730	0.9842	0.9816
0.030	0.9874	0.9721	0.9730	0.9789
0.033	0.9866	0.9695	0.9711	0.9775
0.036	0.9859	0.9680	0.9690	0.9760
0.039	0.9842	0.9660	0.9685	0.9755
0.042	0.9836	0.9658	0.9655	0.9760

Thus, at the selection of a small learning rate in combination with early stopping training technique for the neural network training, there is a reduction in the performance of the MSE, thus reducing network over-fitting of the neural network during training. Furthermore, this prevents over-drawing of the excess information needed from the dataset, thereby resulting in a biased training network. Therefore, there is a need to select the adequate learning rate during neural network training correctly.

Table V and Table VI show values of MSE and training regression of the dataset neural network training on the application of early stopping training technique using BR training algorithm for two hidden layers MLP-ANN. The results still validate 0.003 learning rate as the lowest MSE and the highest training regression for the network training. The MSE increases as the learning rate increases to 0.042 while the training regression gradually decreases on the increase of the learning rate showing a substantial deviation of the actual dataset from the training dataset.

However, comparison of the performances of single layer and two layers MLP-ANN show that, for the prediction of signal power loss, training with a single hidden layer MLP-ANN gives the least MSE performance and the highest coefficient of regression in comparison to training with two hidden layers MLP-ANN using the same learning rates. These are clearly seen in Table III to Table VI, respectively. For example, MSE is 1.610 for single hidden layers and 1.820 for two hidden layers, and training regression is 0.9981 for a single hidden layer and 0.9881 for two hidden layers MLP-ANN. However, increased training time was seen with neural network training using the 0.003 learning rate, with 00:01:11 seconds used for training single layer MLP-ANN and 00:05:33 required to train two hidden layers MLP-ANN. This validates information from literature implying required increased convergence training time using a small learning rate.

TABLE V. TRAINING PERFORMANCE AT DIFFERENT LEARNING RATE USING BR ALGORITHM FOR TWO HIDDEN LAYERS MLP-ANN

Learning Rate	Epoch (Iteration) Max. 1000	Time (seconds)	Performance (MSE)	Validation check
*0.003	1000	00:05:33	1.820	0
0.006	1000	00:04:56	2.240	0
0.009	1000	00:03:10	2.330	0
0.012	1000	00:01:18	2.840	0
0.015	1000	00:01:58	2.915	0
0.018	1000	00:02:01	2.970	0
0.021	1000	00:02:20	2.177	0
0.024	1000	00:02:45	2.050	0
0.027	1000	00:03:18	2.220	0
0.030	1000	00:03:36	2.825	0
0.033	1000	00:03:47	2.835	0
0.036	1000	00:03:51	2.905	0
0.039	1000	00:03:58	2.960	0
0.042	1000	00:04:03	2.965	0

TABLE VI. REGRESSION AT DIFFERENT LEARNING RATES USING BR TRAINING ALGORITHM FOR TWO HIDDEN LAYERS MLP-ANN

Learning Rate	Training	Test	Validation	All
*0.003	0.9881	0.8601	0.8585	0.8820
0.006	0.9860	0.8590	0.8360	0.8810
0.009	0.9760	0.8545	0.8225	0.8760
0.012	0.8960	0.8400	0.8150	0.8500
0.015	0.8775	0.8395	0.8146	0.8488
0.018	0.8600	0.8255	0.8110	0.8400
0.021	0.7980	0.7690	0.7446	0.7800
0.024	0.7679	0.7440	0.7244	0.7780
0.027	0.7650	0.7401	0.7180	0.7550
0.030	0.7445	0.7345	0.7120	0.7480
0.033	0.7225	0.7300	0.7118	0.7330
0.036	0.7221	0.7286	0.7112	0.7300
0.039	0.6870	0.6110	0.6550	0.6885
0.042	0.6775	0.6080	0.6456	0.6760

IV. CONCLUSIONS AND FUTURE ASPECTS

This research work examines the impact of the architectural composition of multi-layer perceptron ANN in combination with the effect of the learning rate for effective prediction of signal power loss during neural network training using training data from a built-up terrain. A single hidden layer and two hidden layers multi-layer perceptron artificial neural network were considered. Different configurations of the neural network architecture ranging from 4-100 for the single hidden layer multi-layer perceptron network and [4 8] to [108 112] for the two hidden layers multi-layer perceptron network were explored.

The result demonstrates improved prediction of signal power loss in built-up areas using a single layer MLP network compared to two hidden layers MLP network while considering the architectural composition of the neural network and on the application of varied learning rates. For optimal training, the learning rate of 0.003 was more adequate, however, it required more convergence training time.

Further extension of this work will explore the performance of other architectural compositions of ANN in the prediction of electromagnetic signal power loss.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS CONTRIBUTIONS

Virginia C. Ebhota (VCE) and Viranjay M. Srivastava (VMS) conducted this research. VCE trained and analyzed the model with data and wrote the paper, VMS has verified the result with the designed model. All authors had approved the final version.

REFERENCES

[1] N. Faruk, S. I. Popoola, N. T. Surajudeen-Bakinde, *et al.*, "Path loss predictions in the VHF and UHF bands within urban environments: Experimental investigation of empirical, heuristics and geospatial models," *IEEE Access*, vol. 7, pp. 77293 – 77307, June 2019.

[2] Y. Wang, Y. Deng, F. Ren, R. Zhu, P. Wang, T. Du, and Q. Du, "Analysing the spatial configuration of urban bus networks based on the geospatial network analysis method," *Cities*, vol. 96, pp. 1-16, Jan. 2020.

[3] A. Aditian, T. Kubota, and Y. Shinohara, "Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia," *Geomorphology*, vol. 318, pp. 101-111, Oct. 2018.

[4] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," *Neural Networks: Tricks of the Trade*, vol. 7700, pp. 437-478, 2012.

[5] T. K. Sarkar, Z. Ji, K. Kim, A. Medour, and M. S. Palma, "A survey of various propagation models for mobile

communication," *IEEE Antennas and Propagation Magazine,* vol. 45, no. 3, pp. 51-82, June 2003.

[6] E. Amaldi, A. Capone, and F. Malucelli, "Radio planning and coverage optimization of 3G cellular networks," *Wireless Networks*, vol. 14, pp. 435-447, 2008.

[7] 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," *Int. J. on Progress in Electromagnetics Research C*, vol. 82, pp. 155-169, March 2018.

[8] V. C. Ebhota and V. M. Srivastava, "Prediction of signal power loss (at micro-cell environments) using generalized adaptive regression and radial basis function ANN models built on filter for error correction," *International Journal on Communications Antenna and Propagation*, vol. 11, no. 2, pp. 94-105, April 2021.

[9] W. Jin, Z. J. Li, L. S. Wei, and H. Zhen, "The improvements of BP neural network learning algorithm," in *Proc. 5th International Conference on Signal Processing*, Beijing, China, 21-25 Aug. 2000, pp. 1647-1649.

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

[11] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Base line knowledge on propagation modelling and prediction techniques in wireless communication networks," *Journal of Engineering and Applied Sciences*, vol. 13, no. 7, pp. 1919-1934, 2018.

[12] A. B. Tammam, A. Rabie, and K. S. Mustafa, "Neural network approach to model the propagation path loss for great tripoli area at 900, 1800, and 2100 MHz bands," in *Proc. 16th International Conference on Sciences and Techniques of Automatic Control & Computer Engineering*, Monastir, Tunisia, 21-23 Dec. 2015, pp. 793-798.

[13] J. P. Kiran, D. N. Vishal, and G. Rajkot, "Comparative analysis of path loss propagation models in radio communication," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, 2015.

[14] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*, 1st Ed., Cambridge, Massachusetts: MIT Press, 1986.

[15] Z. M. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7315-7327, Dec. 2018.

[16] M. S. Mollel and M. Kisangiri, "Comparison of empirical propagation path loss models for mobile communication," *Computer Engineering and Intelligent Systems,* vol. 5, no. 9, pp. 1-10, 2014

[17] J. Isabona and V. M. Srivastava, "Coverage and link quality trends in suburban mobile broadband HSPA network environments," *Wireless personal Communication*, vol. 95, no. 4, pp. 3955-3968, Aug. 2017.

[18] K. Hornik, M. Stinchcombe, and H. White, "Multi-layer feedforward networks are universal approximators," *Neural Networks,* vol. 2, pp. 359-366, 1989.

[19] V. C. Ebhota and V. M. Srivastava, "Effect of architectural composition of MLP ANN in neural network learning for signal power loss prediction," *Journal of Communications,* vol. 16. no. 1, pp. 20-29, Jan. 2021.

[20] 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, 2018.

[21] A. Alatishe, O. Adu, A. Atayero, and F. Idachaba, "A performance review of the different path loss models for LTE network planning," *World Congress on Engineering,* London, UK, 2-4 July 2014, pp. 1-4.

[22] I. Khan, T. C. Eng, and S. A. Kamboh, "Performance analysis of various path loss models for wireless network different environments," *International Journal of Engineering and Advanced Technology*, vol. 2, no. 1, pp. 161-165, Oct. 2012

[23] J. Wei, "Forget the learning rate, decay loss," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 3, pp. 267-272, June 2019.

[24] 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*, vol. 104, no. 3, pp. 935-948, Feb. 2019.

[25] A. Neskovic and N. Neskovic, "Microcell electric field strength prediction model based upon artificial neural networks," *AEU-International Journal Electronics and Communications*, vol. 64, no. 8, pp. 733-738, Aug. 2010.

[26] H. Anysz, A. Zbiciak, and N. Ibadova, "The influence of input data standardization method on prediction accuracy of artificial neural networks," *Procedia Engineering*, vol. 153, pp. 66-70, 2016.

[27] S. P. Sotiroudis and K. Siakavara, "Mobile radio propagation path loss prediction using artificial neural networks with optimal input information for urban environments," *AEU - International Journal of Electronics and Communications*, vol. 69, no. 10, pp. 1453-1463, Oct. 2015.

[28] V. C. Ebhota and V. M. Srivastava, "An adaptive vector median filtering approach to enhance the prediction efficiency of signal power loss," *Journal of Communications*, vol. 15, no. 12, pp. 866-875, Dec. 2020.

[29] V. C. Ebhota and V. M. Srivastava, "Performance analysis of learning rate parameter on prediction of signal power loss for network optimization and better generalization," *Wireless Personal Communications*, vol. 116, no. 2, pp. 1111-1128, May 2021.

[30] A. Torre, F. Garcia, I. Moromi, P. Espinoza, and L. Acuna, "Prediction of compression strength of high-performance concrete using artificial neural networks," *Journal of Physics*, vol. 582, pp. 1-6, 2015.

[31] R. Caruana, S. Lawrence, and C. L. Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and

early stopping," in *Proc. 13th Int. Conf. on Neural Information Processing Systems*, 2000, pp. 381-387.
[32] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415-447, May 1992.
[33] N. M. Nawi, W. H. Atomi, and M. Z. Zehman, "The effect of data pre-processing on optimized training of artificial neural networks," *Procedia Technology*, vol. 11, pp. 32-39, 2013.
[34] H. Yu and B. M. Wilamowski, *Levenberg-Marquardt Training: The Industrial Electronics Handbook*, Routledge Handbook, CRC, April 2016.
[35] F. D. Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian regularization," in *Proc. International Conference on Neural Networks*, Houston, TX, USA, 12 June 1997, pp. 1930-1935.
[36] 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*, vol. 9, pp. 378-383, 2009.
[37] M. A. Costa, A. P. Braga, and B. R. Menezes, "Improving generation of MLPs with sliding mode control and the Levenberg-Marquardt algorithm," *Neurocomputing*, vol. 70, no. 7-9, pp. 1342-1347, March 2007.
[38] I. F. Anitzine, J. A. R. Argota, and F. P. Fontan, "Influence of training set selection in artificial neural network-based propagation path loss predictions," *International Journal of Antennas and Propagation*, vol. 2012, pp. 1-7, 2012.

Dr. Virginia C. Ebhota received a Bachelor of Engineering degree in Electronic Engineering from the 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 working towards a post-doctoral research fellow with the Department of Electronic Engineering, University of Kwazulu-Natal, Durban South Africa.

She has published various peer-reviewed articles in Journals and IEEE Conferences. Her major research area is system modeling using Artificial Intelligence (AI). She is author/co-author of more than 12 scientific contributions including articles in international refereed Journals and Conferences

Prof. Viranjay M. Srivastava is a Doctorate (2012) in the field of RF Microelectronics and VLSI Design, Master (2008) in VLSI design, and Bachelor (2002) in Electronics and Instrumentation Engineering. He has worked for the fabrication of devices and development of circuit design. Presently, he is working in the Department of Electronic Engineering, Howard College, University of KwaZulu-Natal, Durban, South Africa. He has more than 18 years of teaching and research experience in the area of VLSI design, RFIC design, and Analog IC design. He has supervised various Bachelors, Masters and Doctorate theses.

He is a Professional Engineer of ECSA, South Africa and Senior member of IEEE(USA), SAIEE, and member of IEEE-HKN, IET(UK), IITPSA. He has worked as a reviewer for several Journals and Conferences both national and international. He is author/co-author of more than 280 scientific contributions including articles in international refereed Journals and Conferences and also author of various books.