Development of Call Setup Success Rate Prediction Model using Ensemble Algorithm

T.T. Awofolaju\*1, H.O. Lasisi2, O. Oladepo3, S. Adebayo4, O.M. Bada5, A.A. Olawuyi6, P.K. Awofesobi7

1,2,3,5,6,7Department of Electrical and Electronic Engineering, Osun State University, Osogbo, Nigeria.

4Mechatronic Engineering Programme, Bowen University, Iwo, Nigeria

*\*Corresponding author: tolulope.awofolaju@uniosun.edu.ng*

**Abstract**

The poor quality of services (QoS) experienced by mobile network users may be a result of network providers relying on the troposphere for transmission of signals without first evaluating and characterizing the area of the conveyance of the signals. The knowledge of which weather variables have an impact on signal propagation is expedient to define network vulnerability. This study developed a machine learning model using bagged tree and LS Boosting algorithms after employing a suitable feature selection algorithm. Hyperparameter optimization using the Bayesian technique was employed to obtain an optimal model. The performance of the models was then compared using mean average error, mean squared error, r-squared, and prediction speed. The models were further used to predict call setup success rates using new data. Results show that the bagged tree algorithm performed better based on figures of metric as well as prediction values.

**Keywords:** Quality of Service, Mobile Network, Call set-up success rate, drop call rate, signal.

**1. INTRODUCTION**

Global System for Mobile Communications (GSM), which was introduced in Nigeria in 2001, has undoubtedly made a significant contribution to the quality of life for Nigerians [[1]](#bakare) . However, as mobile services have expanded, it has become crucial for mobile communication operators to accurately measure the quality of service (QoS) of their networks and to continue to improve them as efficiently and effectively as possible to maintain a competitive edge [[2-3]](#brooks).

Poor QoS is a result of network operators' reliance on sending signals down the troposphere without first evaluating and characterizing the troposphere [[4](#mohameed)]. A component of the atmosphere that has a direct impact on human life is the troposphere. It is an area of all-weather on Earth and the lowest layer of the atmosphere. At the poles and equator, the troposphere is located at a height of about 10 km and 17 km, respectively [[5].](#sas)

It is necessary to understand how weather parameters affect signal propagation and how changing weather conditions can seriously impair system performance to characterize the reliability of networks [[6-8]](#agbo). To forecast, simulate, and design high-performance communication systems, the exact transmission characteristics of radio waves in various environments must be known. Network planning and optimization are required for mobile network providers. There is a need to strike a balance between network operation, QoS, and radio coverage in various situations [[9]](#fad) .

The structure of radio refractive index, N, at the lower part of the atmosphere has helped in identifying the weather parameters likely to affect signal propagation and it is stated in Equation 1 given as [[10](#ten)]:

where t is the temperature in Celsius, T is the temperature in kelvin, P is the pressure (hPa), , H is the relative humidity (%), 𝑒𝑠: saturation vapor pressure (hPa) at the temperature t (0C).

Studies show that when signals are transmitted, they interact with tropospheric variables such as wind speed, relative humidity, and temperature[[11]](#twentyone) – [[12]](#twentytwo). Call Setup Success Rate (CSSR) measures the call setup success or reflects the probability of successful calls initiated by the mobile station [[13]](#twentyfive). This happens immediately after the traffic channel (TCH) assignment is done, regardless of whether the call is dropped later or not by either the calling or the called party. The CSSR is a key counter in evaluating network performance.

**2. LITERATURE REVIEW**

Call Setup Success Rate (CSSR) measures the setup success rate of the call or reflects the probability of successful calls commenced by the mobile station. This happens immediately after the traffic channel (TCH) assignment is done, regardless of whether the call is dropped later or not by either the calling or the called party. The CSSR is an important counter in evaluating the network performance. If the low value is gotten from the counter, the mobile station is not likely to successfully initiate a call, thus adversely affecting the user experience. The higher the value of the CSSR, the better the cell's performance. High CSSR is obtained when Stand-alone Dedicated Channel (SDCCH) seizure and TCH allocation are effortlessly obtained to set up a call. Equation 2 is the formula for CSSR’s calculation.

The prediction of the state of a channel on a given link was done by [[14]](#eleven) taking measurements on other links, thus causing a decline in the signaling overhead. The first representative approach considered was Random Dot Product Graphs while the second approach was Graph Neural Network.

The proposed graph-based machine learning methods outperformed traditional methods in predicting the state of the channel on a given link by taking measurements on other links, achieving an RMSE of 10 dB and 73% accuracy by making use of a dataset of RSSI measurements of real-world Wi-Fi operating service providers. The paper should have discussed the computational complexity of the proposed methods, which could be a potential limitation in practical implementations.

The study by [[15]](#twelve) made use of an automatic artificial neural network (ANN) predictive quality of service model to evaluate the efficiency of services rendered by the GSM network in Nigeria, after which the evaluation results of the developed GSM QoS prediction model showed that the results of the developed model could perform favorably well but not at its best, compared to how the Nigerian Communications Commission (NCC) approaches it manually. Hence, there is a need to employ more advanced machine learning algorithms or techniques to develop the QoS prediction model, such as deep learning or ensemble learning, to boost the accuracy of the model.

[[16]](#thirteen) used a walk-test methodology to measure KPIs for accessibility of the internet on the 4G network by the users who subscribe to different mobile network operators (MNOs) within the University of Ilorin. Data was gathered using TEMS Investigation 16.3.4 and processed using TEMS Discovery Device 10. The walk test involved uploading files, downloading data, and streaming videos online at various test areas. MNO4 had the best overall quality and throughput, while MNO1 had the poorest service, although it still provided some service in all test locations. The 4G test did not yield exceptional results, but students reported specific locations with optimum 4G speed. Expanding the study to include other universities or public areas to determine if the results are similar or if there are differences in service quality and throughput will be a great advantage.

The study by [[17]](#fourteen) proposed a traffic congestion prediction model using machine learning techniques used for the prediction of the traffic congestion existence in LTE networks as appraised by users. The model was divided into several phases: data preparation, splitting, modeling, classification, model evaluation, tuning, and result. The four machine-learning algorithms were compared and conclusions were hinged on the output of the Jupyter Notebook for the classifiers. Out of all the techniques used in predicting the traffic congestion existence, k-Nearest Neighbour had the best performance. Online machine learning techniques will be considered for future studies, and they can constantly obtain data from network operators to gather the necessary features and prediction performance of the traffic congestion in real-time to aid the traffic providers in engaging mechanisms to reduce traffic congestion to the barest minimum.

[[18]](#fifteen) measured and analyzed the Key Performance Indicators (KPIs) of a 4G/LTE Telecom of Kosovo (TK) network 24-cell cluster. The results of the analysis show that the availability KPI has fewer values than the threshold (>99%). He stated that future studies would focus on the analysis of the QoS in the overall 4G/LTE network executed in TK and that the major challenges the operators will encounter during the transition process from 4G to 5G technologies would also be addressed.

[[19]](#sixteen) used network statistics to evaluate the QoS of a cellular network service provider of an enclosed area during a church event. The CSSR and Percentage Drop Call Rate were investigated and compared with the benchmark defined by the Nigerian Communications Commission (NCC). The study results showed that the cellular network service provider's Key Performance Indicators (KPIs) fell below the NCC recommendation, especially during high traffic intensity. The quality of service requires improvement to ensure better service delivery to subscribers. Comparing the QoS of different cellular network service providers in the same area and evaluating the QoS of cellular networks in different geographical locations should be thoroughly examined.

[[20]](#seventeen) evaluated the 4G-LTE communications base station’s quality parameters in the rural part of Peru using Key Performance Indicators (KPIs) defined for 4G LTE Technology, including signal level, Signal to Noise Ratio and quality. The study confirmed that the KPIs comply with the recommendations of the ITU in its E-800 recommendation. The study found that the 4G-LTE communications base station’s quality parameters in a rural part of Peru comply with the ITU's E-800 recommendation, which ensures maximum mobile phone coverage in the rural parts and accessibility to the complete mobile phone network nationwide. Future studies could focus on analyzing the impact of environmental factors, such as weather conditions and terrain, on the communications’ base station’s parameters.

**3.0 MATERIALS AND METHODS**

**3.1 DATASET DESCRIPTION**

Historical data of CSSR, and meteorological parameters such as temperature, wind speed, relative humidity, and surface pressure were collected from the archives of the Nigerian Meteorological Agency (NIMET) and Nigerian Communication Commission (NCC) in comma-separated values (CSV) using Modern Era Retrospective Analysis for Research and Applications version 2 (MERRA – 2). The data collected spanned over seven years from January 1, 2016 to December 31, 2022, and the summary dataset is represented in Table 1.

Table 1: Dataset Summary

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Statistics** | **CSSR** | **Surface Pressure**  **(kPa)** | **Relative Humidity (%)** | **Avg Temp (K)** | **Avg**  **Wind**  **Speed**  **(m/s)** |
| count | 432 | 432 | 432 | 432 | 432 |
| mean | 99.04 | 98.49 | 84.37 | 25.93 | 4.71 |
| std | 0.73 | 1.61 | 7.29 | 7.08 | 8.67 |
| min | 91.9 | 96.63 | 48.56 | 18.12 | 2.40 |
| 25% | 98.95 | 96.94 | 81.56 | 24.43 | 3.22 |
| 50% | 99.3 | 98.12 | 86.88 | 25.54 | 3.62 |
| 75% | 99.48 | 100.16 | 89.44 | 26.66 | 4.11 |
| max | 99.7 | 101.08 | 91.69 | 82.44 | 82.44 |

**3.2 Feature Selection**

Feature selection is the main step in machine learning-based model development [[21]](#eighteen). To obtain a good model, the feature selection algorithm chooses the most relevant features from the feature vector and discards the irrelevant ones.

**3.3 Ensemble Model**

Ensemble modeling involves the use of different basic machine learning models, although as a single model, to forecast an outcome. Reduction of the generalization error of the forecast is the inspiration for using ensemble models. The general principle of ensemble methods is to construct a linear combination of some model fitting methods, instead of using a single fit of the method.

**3.3.1** **Bagging Tree Algorithm**

The Bagging Tree (BT) algorithm creates a bootstrapped sample, on which either a regression algorithm or classification algorithm is applied. For regression, an average is taken and computed over all the outputs forecasted by the individual learners. For classification, the most voted class (hard-voting) is considered as the output, else the highest average of all the class probabilities (soft-voting). Mathematically, BT prediction can be represented as in Equation 3.

(3)

where is the output of the bagging tree and are the input.

**3.3.2 LSBoost Algorithm**

Unlike bagging, the LSBoost (LSBT) algorithm trains the basic machine-learning models consecutively and gives weights to all the training records. The training set for the subsequent iteration will be overrepresented by the training records that are difficult to categorize thanks to the boosting process.

Every training record has a weight assigned by boosting, and depending on how tough the classification is, boosting must adaptively adjust the weight. This results in the creation of a basic learners’ group, skilled at classifying both simple and complex records. By using a straightforward voting aggregation, the model's basic learners are all pooled.

**3.4 Hyperparameter Optimization**

Hyperparameter optimization aims at finding the optimal collection of hyperparameters for a model to reduce the loss function and, ultimately, increase accuracy on given independent data. Hyperparameter tuning is important in machine learning-based model development and several researchers have employed various optimization strategies [[22- 25]](#nineteen) including the Bayesian method employed in this study. Bayesian optimization is a form of sequential model-based optimization (SMBO) strategy that improves the sampling approach for subsequent tests. Table 2 shows the model optimization using Bayesian hyperparameter tuning. Finding the highest value at the sample point for an unknown function f is the goal of Bayesian optimization and it is represented mathematically in Equation 4 as:

 (4)

where A represents the x search space.

Table 2: Model optimization using Bayesian Hyperparameter Tuning

|  |  |  |
| --- | --- | --- |
| Hyper-parameter | Range value | Optimized value |
| Ensemble method | Bagged, LSBoost | Bagged |
| Number of learners | 10-500 | 12 |
| Learning Rate | 0.001-1 | 0.01 |
| Minimum leaf size | 1-346 | 344 |
| Number of Predictors to Sample | 1-8 | 6 |

**4.0 RESULTS AND DISCUSSION**

**4.1 Exploratory Data Analysis**

Figure 1 shows the distribution of CSSR for different regions under different weather conditions namely: relative humidity, average temperature, average wind speed, and surface pressure. A significant effect can be observed from the change in relative humidity. Figure 2 shows the Feature Importance of the CSSR Prediction Model.

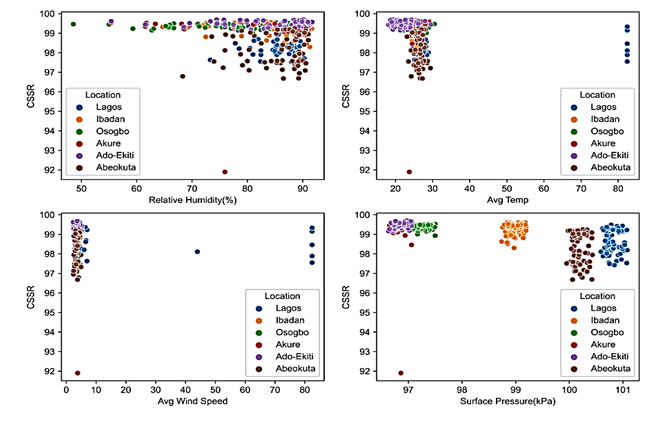
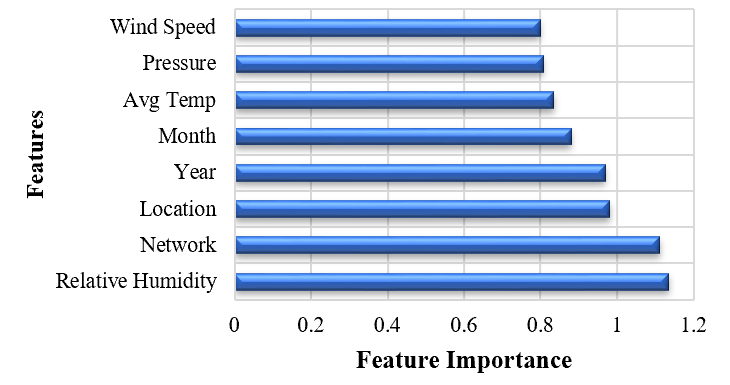


Figure 1: Exploratory Data Analysis (a) CSSR versus change in relative humidity for several regions (b) CSSR versus average temperature (c) CSSR versus average wind speed and (d) CSSR versus surface pressure.

 Figure 2: Feature Importance on CSSR Prediction Model

4.2. Experimental Results

The experimental environment of this study was based on MATLAB R2022a (Regression learning Application). In this study, two ensemble-based algorithms trained using historical datasets to predict CSSR were applied. To analyze the performance of the proposed models, Figure 3, Figure 4, Figure 5 and Figure 6 compared the performance of the models based on RMSE, MAE, R-squared, and prediction speed respectively. BT algorithm performed better than LSBT based on all the figures of metrics considered. It can also be seen that the effect of Principal Component Analysis (PCA) on the model is very insignificant. However, optimizing the model by Bayesian hyperparameter turning produces a better model. Figure 7 and Figure 8 present the prediction of CSSR using the two models: LSBT and BT respectively. The BT model predicted the CSSR accurately for the 2nd 7th, 18th, and 20th months. The overall prediction performance of BT in Figure 7 is better than that of LSBT in Figure 8.

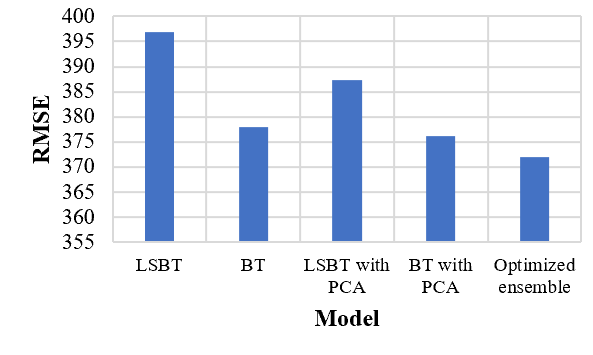


Figure 3: Performance of Models using RMSE

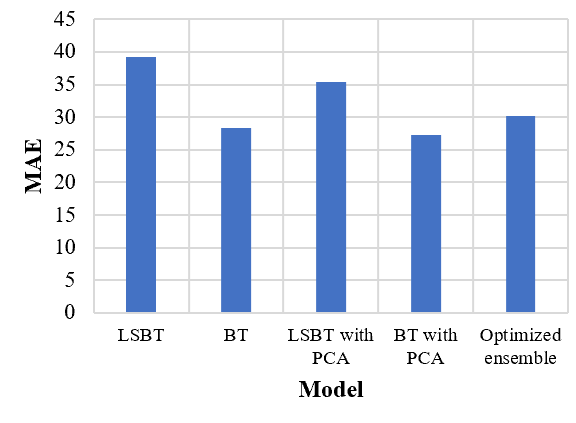


Figure 4: Performance of Models using MAE

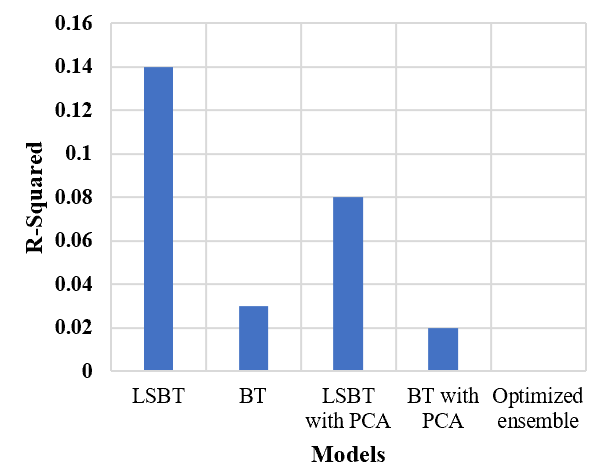


Figure 5: Performance of Models using R-squared

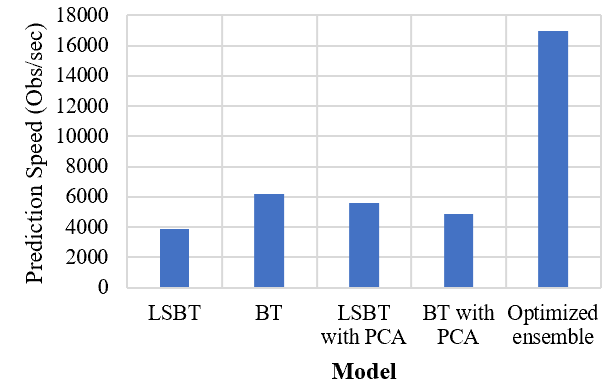


Figure 6: Performance of Models using Prediction Speed

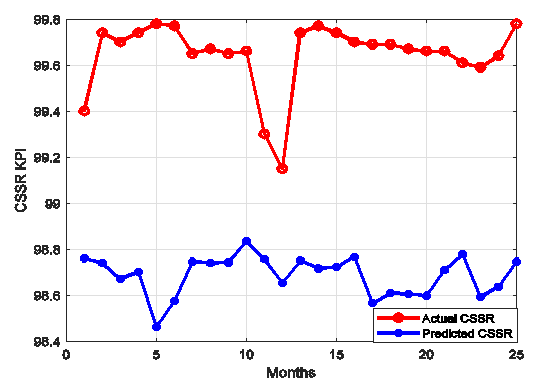


Figure 7: CSSR Prediction using LSBT

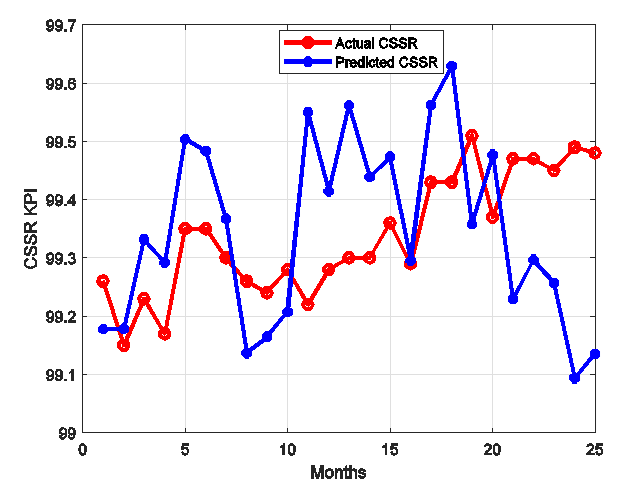


Figure 8: CSSR Prediction using BT

**5. CONCLUSION**

Quality of service is a main feature in the behavioral patterns in the telecommunications industry and one of the key performance indicators is the CSSR. Mobile communication engineers have faced the challenges of understanding these parameters which is attributed to achieving optimum quality of end-user experience (QoE) of communication networks.

This study presented a machine-learning model capable of predicting CSSR using the historical dataset of meteorological parameters such as temperature, wind speed, relative humidity, and surface pressure. Two algorithms BT and LSBT were used to learn the features from the dataset after several data preprocessing and hyperparameter optimization.

The actual and predicted values of the CSSR were compared and found that BT performed better than LSBT. The proposed model can establish CSSR, an important KPI in QoS and QoE measurement. Finally, this model has a great impact on the effective design, development, and deployment of radio communication systems.

**REFERENCES**

[1] Bakare, B., Ekanem, I. and Allen, I. 2017. "Appraisal of global system for mobile communication (GSM) in Nigeria," *American Journal of Engineering Research,* vol. 6, pp. 97-102.

[2] Brooks, P. and Hestnes, B. 2010. "User measures of quality of experience: why being objective and quantitative is important," *IEEE network,* vol. 24, pp. 8-13.

[3] Banović-Ćurguz, N. and Ilišević, D. 2019. "Mapping of QoS/QoE in 5G networks," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 404-408.

[4] Borralho, R., Mohamed, A., Quddus, A. U., Vieira, P., and Tafazolli, R. 2021. "A survey on coverage enhancement in cellular networks: Challenges and solutions for future deployments," *IEEE Communications Surveys & Tutorials,* vol. 23, pp. 1302-1341.

[5] Sassen, K. and Wang, Z. 2012. "The clouds of the middle troposphere: Composition, radiative impact, and global distribution," *Surveys in geophysics,* vol. 33, pp. 677-691.

[6] Agbo, E. P., Ekpo, C. M., and Edet, C. O. 2021 "Analysis of the effects of meteorological parameters on radio refractivity, equivalent potential temperature and field strength via Mann-Kendall test," *Theoretical and Applied Climatology,* vol. 143, pp. 1437-1456.

[7] Vadukkoot, A., Ravi Shankar, K., and Karakkad, S. 2022 "Influence of meteorological parameters on the propagation of lightning electromagnetic fields," *Journal of Electromagnetic Waves and Applications,* vol. 36, pp. 1323-1338.

[8] Jeremiah Ekah, U. and Onuu, M. U. 2022. "Tropospheric influence on call setup in mobile networks," *Journal of Engineering Research and Reports,* vol. 22, pp. 14-26.

[9] Fadlullah, Z. M., Mao, B., and Kato, N. 2022. "Balancing QoS and security in the edge: Existing practices, challenges, and 6G opportunities with machine learning," *IEEE Communications Surveys & Tutorials*.

[10] Thayer , G. D. 1974. "An improved equation for the radio refractive index of air," *Radio Science,* vol. 9, pp. 803-807.

[11] Ali ,Y. A., Awwad, E. M., Al-Razgan, M., and Maarouf, A. 2023. "Hyperparameter search for machine learning algorithms for optimizing the computational complexity," *Processes,* vol. 11, p. 349.

[12] Mantovani, R. G., Rossi, A. L., Vanschoren, J., Bischl, B and De Carvalho, A. C. 2015. "Effectiveness of random search in SVM hyper-parameter tuning," in *International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1-8.

[13] Awofolaju T.T., Lasisi H.O., Olawuyi A.A., Oladepo O., Adebayo S. (2023): Artificial NeuralNetwork Based Prediction of Key Performance Indicators for Mobile Telecommunications. *Journal of Science and Technology, 15(2) (2023):43 – 53.*

[14] Rattaro, C., Larroca, F., and Capdehourat, G. 2021. "Predicting wireless rssi using machine learning on graphs," in *2021 IEEE URUCON*, 2021, pp. 372-376.

[15] Popoola, J. J and Areo A. E. 2020. "Modeling and development of a novel quality of service prediction model for global system for mobile communications network using artificial neural networks," *Journal of Applied Science & Process Engineering,* vol. 7, pp. 510-523, 2020.

[16] Oje, A. and Edeki, S. 2021."KPI measurement on the 4G network within the University of Ilorin," in *Journal of Physics: Conference Series*, 2021, p. 012031.

[17] Kuboye, B. M., Adedipe, A. I., Oloja, S. V., and Obolo, O.A. 2023. "Users’ Evaluation of Traffic Congestion in LTE Networks Using Machine Learning Techniques," *Artificial Intelligence Advances,* vol. 5.

[18] Krasniqi, F., Gavrilovska, L., and Maraj, A. 2019. "The analysis of key performance indicators (KPI) in 4G/LTE networks," in *Future Access Enablers for Ubiquitous and Intelligent Infrastructures: 4th EAI International Conference, FABULOUS 2019, Sofia, Bulgaria, March 28-29, 2019, Proceedings 283*, 2019, pp. 285-296.

[19] Eghonghon, U. K. 2017. "Evaluation of the quality of service of a cellular network using the network statistics," *Evaluation,* vol. 1, pp. 1-7.

[20] Chamorro Atalaya, O. F., 2020 "Evaluation of the quality parameters of a 4G-LTE communications base station, installed in a rural area of Peru,".

[21] Naheed, N., Shaheen, M., Khan S. A., Alawairdhi, M., and Khan M. A. 2020. "Importance of Features Selection, Attributes Selection, Challenges and Future Directions for Medical Imaging Data: A Review," *CMES-Computer Modeling in Engineering & Sciences,* vol. 125.

[22] Raji, I. D., Bello-Salau, H., Umoh, I. J., Onumanyi, A. J., Adegboye, M. A., and Salawudeen, A.T. 2022. "Simple deterministic selection-based genetic algorithm for hyperparameter tuning of machine learning models," *Applied Sciences,* vol. 12, p. 1186.

[23] Liao, L., Li, H., Shang, W., and Ma, L. 2022. "An empirical study of the impact of hyperparameter tuning and model optimization on the performance properties of deep neural networks," *ACM Transactions on Software Engineering and Methodology (TOSEM),* vol. 31, pp. 1-40.

[24] Awofolaju T.T., Lasisi H.O., Ajewole T.O., Adebayo S., Olawuyi A.A., Adeagbo F.M. and Asafa M.O. (2024): Validation of Key Performance Indicators of Mobile TelecommunicationOperators Using Ensemble Models and Artificial Neural Networks with Nigerian Communication Commission’s Thresholds, *UNIOSUN Journal of Engineering and Environmental Sciences.6(1):61- 72.*

[25] Awofolaju T.T., Lasisi H.O., Ajewole T.O., Olawuyi A.A., Adebayo S., Alawode K.O., Oladepo O. (2024): Performance Evaluation of Predictive Models for Key Performance Indicators in Telecommunication Networks. *University of Pietsi Scientific Bulletin, Faculty of Electronics Communications and Computers.* *24(1): 27- 34.*