A Comparative Analysis of Support Vector Machine and Decision Tree Algorithm for Predicting Fault in Uninterruptible Power Supply Systems

Isaac M. Doe, John K. Annan, Benjamin Odoi

Abstract: ***Power supply systems can have problems, and Ghana Gas Limited is not an exception. Ghana Gas Limited uses an intricate Uninterruptible Power Supply (UPS) system, which is made up of several parts such as electromechanical components, PCB boards, and electrolytic capacitors. The majority of components have technical lifespans that are governed by usage, operational environment, and working conditions, such as electrical stress, working hours, and working cycles. Most of the time, these errors affect the integrity and power supply of the product after it is manufactured. The issue is that it takes longer for the professionals who operate this machine to recognise these flaws, which makes it difficult for them to predict errors quickly or anticipate the likelihood of faults occurring in system components at an early stage, allowing for effective corrective action to be taken. Support vector machines (SVMs) and decision trees were used in this study to predict faults for the technical data scheduling of uninterruptible power supply systems for Ghana Gas Limited in an efficient manner. Based on a comparative analysis using these two techniques, faults in Ghana Gas Limited's power supply system were predicted using a four-hour daily interval dataset on UPS recordings, including input voltage, battery voltage, battery current, and alarm, spanning from August 2017 to October 2023. The findings showed that the support vector machine was more efficient in detecting fault locations in the power supply system, with an accuracy of 96.80%, a recall of 99.80%, a precision of 100%, and an F1-score of 93.15%. The results from the error metrics also validate the measures in assessing the predictive ability of the model with MAE of 0.42%, MSE of 1.18%, RMSE of 4.45%, R^2 of 99.97%, RMSLE of 0.036%, and MAPE of 0.21%.***

Keywords: ***Power Supply System, Support Vector Machine, Decision Tree Algorithm, Precision, Accuracy, Error Metrics***

I INTRODUCTION

The majority of engineers now find that their ability to distribute an uninterruptible electrical power supply and safeguard their operations is limited by how sophisticated they are at handling power supply outages [1].

Manuscript received on 18 April 2024 | Revised Manuscript received on 03 May 2024 | Manuscript Accepted on 15 May 2024 | Manuscript published on 30 May 2024.

**Correspondence Author(s)*

Isaac M. Doe, Department of Electrical and Electronic Engineering, University of Mines and Technology, UMaT, Tarkwa, Ghana. E-mail: bodoi@umat.edu.gh, isaacmawulidoe@gmail.com, ORCID ID: 0009-0000-8451-0064

Dr. John K. Annan, Department of Electrical and Electronic Engineering, University of Mines and Technology, UMaT, Tarkwa, Ghana. E-mail: jkannan@umat.edu.gh, ORCID ID: 0000-0001-8056-7880

Dr. Benjamin Odoi*, Department of Mathematical Sciences, University of Mines and Technology, UMaT, Tarkwa, Ghana. E-mail: bodoi@umat.edu.gh, ORCID ID: 0000-0001-6759-0744

Larger industries utilise uninterruptible power supplies (UPS) as an emergency power source to ensure business activities continue as usual in the event of a primary power supply failure. When it comes to providing almost instantaneous safety from input power failures, a UPS is different from a backup or supplemental generator since it uses electricity produced by batteries, ultracapacitors, or flywheels [2]. A variety of scenarios need the use of uninterruptible power supply (UPS) to provide consistent and well-regulated AC voltages for critical workloads, including computer servers, medical equipment, air traffic control systems, and communication networks [3]. It has been noted in numerous deployments that when system load increases over time, an upgraded UPS with a larger capacity is needed [4]. As power grids expand and loads increase, the primary goal of a UPS is to maintain the stability and reliability of the electrical supply. Furthermore, as the system serves as a precaution or protection against any unanticipated power outages that the companies may suffer, it must be reliable, secure, and less prone to malfunctions [5] and [6]. However, overloading, overvoltage, power fluctuations, and outdated components can result in malfunctions in UPS systems, which can cost the industry a lot of money in terms of lost productivity and system replacement [4]. Fault prediction is the process of tracking and evaluating historical data to identify the presence of a failure in the power system, so that actions may be taken to prevent accidents and assure system recovery [7]. As a result, it is essential to implement fault prediction models that are efficient and accurate in determining when faults may occur [1]. Fault prediction is an essential technology and maintenance security technique that is more advanced than fault diagnosis, which is frequently carried out after issues have occurred [8]. Making wise decisions to avoid errors and minimize their negative consequences is made easier with the help of fault prediction. To reduce the frequency and duration of power outages, utility workers can find and eliminate persistent defects with the use of high-accuracy fault prediction in power systems [9].

Previously, UPS system maintenance was reactive, initiated only when a problem was identified. Recent years have seen an increase in the popularity of preventive maintenance, which involves changing out components to increase system reliability [10].

The simplest kind of preventive maintenance is to target consumable parts before they are expected to reach the end of their useful life. Fans, AC and DC capacitors, and of course, the battery are all consumable parts of UPS systems [11]. Preventative maintenance is the failure prediction based on prediction models created from gathered data or log files in more sophisticated systems [12]. Making the right decisions to prevent power system problems and estimating their likelihood can both be aided by analysing historical data [13]. In general, utilising electrical measurement data to its fullest potential will enhance the accuracy of failure prediction and ensure the stability and dependability of the UPS system. Studies using artificial intelligence (AI) and machine learning have been developed in the past several years to predict faults. Right now, it's a worthwhile and pressing topic [13] [14]. Provided a more accurate prediction strategy for optimised Artificial Neural Networks (ANNs) based on multilayer evolutionary algorithms to improve the fault forecasting model's accuracy [15]. employed ANN to obtain likelihoods of success for five fault prediction techniques, ranging from 87% to 100% using 33 data sets [16] and looked at the use of convolutional neural networks (CNNs) to forecast refrigerant charge failures. Two classification and regression predictive models were proposed to predict the quantitative refrigerant in both cooling and heating applications. In summary, the recommended tasks were finished with a 3.1% mistake rate and 99% accuracy. To predict power converter failures, [17] developed a fault prediction model using Markov Chain Analysis, based on data collected from several UPS installations. Furthermore, classification is a crucial component of the fault prediction process. The Support Vector Machine (SVM) is a hyperplane-configured discriminant classifier. SVM-based applications are feasible in [17] and [18].

This study aims to estimate the failure rate of UPS systems using Support Vector Machine (SVM) and Decision Tree (DT) Algorithms to create prediction models that the Ghana Gas engineering team can utilise. The Ghana National Gas Company, also known as Ghana Gas, was established with the responsibility of creating, acquiring, and overseeing the natural gas infrastructure required for the processing, transportation, and marketing of gas to satisfy the nation's needs for both household and commercial electricity. Most importantly, a consistent and dependable power source is essential to Ghana Gas's operations and activities. Predicting UPS failures is therefore necessary to enhance power supply performance and lower the company's total operational expenditure (OPEX).

The UPS employed by Ghana Gas is a complex system comprising several components, including PCB boards, electromechanical components (such as relays and fans), and electrolytic capacitors. The majority of components' lifespans are determined by their technical attributes and are influenced by their usage, operational environment, and working conditions, that is, working hours, working cycles and electrical stress [19]. Currently, preventative maintenance is conducted over a predetermined period without considering the level of stress experienced or the overall health of the UPS system. For instance, fans are typically replaced every five years without considering whether the UPS was, perhaps, in a clean room or a harsher, dustier environment [20]. Costly on-site maintenance is carried out irrespective of the device's status and may therefore be too late or too early. The latter situation could result in the servicing of a healthy component, thereby increasing the company's financial costs and decreasing the UPS systems' reliability [21].

A model for accurate fault prediction and forecasting will help to improve the level of UPS system reliability and reduce power quality disturbances as well as equipment damages [21]. Furthermore, the prediction model will enable Ghana Gas to better manage its engineering resources by forecasting failures and enabling precautionary actions to be taken, thereby reducing operating expenses caused by unnecessary component replacement and additional charges. An essential step in understanding the reliability of the UPS system as a whole is determining the significance of various UPS parameters [22]. Hence, by monitoring UPS parameters such as output voltage, output current, frequency, power factor, working hours, and active and reactive powers, a runtime lifespan calculation of the components can be conducted to determine their health state and predict potential UPS failures. To perform such a task, an intelligent data methodology has to be employed. This study emphasises the utilisation of machine learning algorithms for fault evaluation and prediction in the UPS system. The research discusses fault prediction models with a particular emphasis on UPS operations. Particularly, operational data from UPS installations has been recorded. To create the failure prediction models, the data is then processed using Support Vector Machines (SVM) and Decision Tree (DT) algorithms.

II MATERIALS AND METHODS

A. Data Collection

UPS obtained the data for the study (recordings include input voltage, battery voltage, battery current, and alarm), which records an observation every four hours. The data was gathered between August 2017 and October 2023. The Python programming language was used to do the analysis.

B. Support Vector Machine (SVM)

Given a set of training data $\{(x_1, y_1), \dots, (x_1, y_N)\}$, where $x_i \in R^D$ Are the input vectors and $y_i \in \{-1,1\}$ Are the corresponding class labels, an SVM seeks to construct a hyperplane that separates the data with the maximum margin of separability [23]. N Is the number of observations, and D Is the dimension of the input vectors. The decision function can be written as

$$f(x) = sign\{\sum_{j=1}^{N^{sv}} \alpha_j y_j^{sv}\left(\Phi(x) \cdot \Phi(x_j^{sv})\right) + b\} \quad (1)$$

Where x_j^{sv} The support vectors are $\Phi(x)$ A nonlinear vector function that maps the input vector onto a higher-dimensional feature space [24], y_j^{sv} Is the label corresponding to the th support vector? N^{sv} Is the number of support vectors, b is a bias term, and α_j The Lagrangian multipliers are the inner product. $\left(\Phi(x) \cdot \Phi(x_j^{sv})\right)$ Called the kernel function.s

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

C. Decision Tree (DT):

In the shape of a tree structure, Decision Tree creates models. The process gradually creates a decision tree for each dataset by breaking it down into smaller and smaller sections. A measure used for segmentation is information gain. To partition the data into the most informative features, we establish an objective function.

$$IG(S,A) = \sum_{v\in V(A)} \frac{|S_v|}{|S|} Entropy(S_v) \qquad (2)$$

S_v Is a subset of set S equal to the attribute value of attribute v, and the range of attribute A is represented by V(A). The measure of impurity or randomness in a dataset is called entropy. Entropy is a quantity that is constantly between 0 and 1.

$$Entropy(S) = \sum_{i=1}^{c} P_i \log \log 2^{P_I} \qquad (3)$$

Where P_i Is the ratio of the sample size of the subset and the $i-th$ Attribute value.

D. Evaluation Metrics

This study compares various machine learning techniques to forecast the location of a wire drawing process problem. Therefore, six (6) well-known assessment measures that are frequently utilized in fault prediction applications were used to gauge how well these algorithms performed. Since a model's effectiveness cannot be determined by a single metric, these evaluation metrics were selected [33]-[35]. These evaluation metrics conform to the literature, and a discussion of the various evaluation and error metrics can be found in the works done by [25]–[32] and [37]-[39].

III RESULTS AND DISCUSSIONS

This section presents the results and discussion of the study. The exploratory and ML models were discussed, along with their development. The results demonstrate a novel approach that can be used as an alternative method for detecting faults in uninterruptible power supply systems.

A. Results

The UPS data includes four-hour periods for each of the equipment's four shifts. In all, 3912 time frames were created over the course of three years. The objective is to generate statistical features for each window to define each daily signal better and reduce the data's dimensionality. To further describe the signal and understand its evolution over time, eight different characteristics were extracted from each data frame.

The features generated included Maximum, Minimum, Mean, Standard Deviation, Root Mean Square (RMS), Skewness, Kurtosis, and Mean Absolute Deviation (MAD). These eight variables were chosen to limit the study's resources and determine whether they are sufficient to achieve identification. As none of the variables require frequency analysis, all features are analysed in the time domain. Furthermore, according to [36], time domain statistical resources provide a high performance to characterize trends and changes.

Table 1: Extracted Statistical Features for the Input Voltage Attribute

Date	Max	Min	Mean	Std	RMS	Skew.	Kurt.	MAD
04/01/2021	423	418	419.83	1.9407	419.837	0.46611	-1.5259	1.5
05/01/2021	427	409	421.5	7.14843	421.5505	-0.75421	-1.2867	5.6667
06/01/2021	422	416	418.833	2.04124	418.8375	0.185073	-1.3889	1.5
07/01/2021	439	420	428.667	7.76316	428.7252	0.066101	-1.9978	6.6667
08/01/2021	440	417	428.667	9.88418	428.7711	0.014249	-2.0706	9.3333
09/01/2021	441	430	436	4.28952	436.0176	-0.15204	-1.9268	3.6667
10/01/2021	442	430	436.667	4.08248	436.6826	-0.32932	-1.3049	3
11/01/2021	429	421	426.25	3.59398	426.2614	-0.63615	-1.7614	2.625
12/01/2021	431	422	425.5	4.1833	425.5171	0.293685	-2.0338	3.5
04/02/2021	427	421	423.333	2.65832	423.3403	0.181385	-2.0485	2.3333

Table 2: Extracted Statistical Features for the Battery Voltage Attribute

	Max	Min	Mean	Std	RMS	Skew.	Kurt.	MAD
04/01/2021	255	251	253.50	1.9748	253.506	-0.4544	-1.977	1.6667
05/01/2021	255	254	254.67	0.5164	254.667	-0.5379	-1.958	0.4444
06/01/2021	255	250	253.83	1.9408	253.840	-1.1754	-0.398	1.2778
07/01/2021	254	252	253.50	0.8367	253.501	-0.8537	-1.172	0.6667
08/01/2021	255	235	253	5.6889	253.059	-2.6125	5.3613	3
09/01/2021	255	250	253.17	2.1370	253.174	-0.4858	-1.834	1.7778
10/01/2021	256	254	254.33	0.8165	254.334	1.36083	-0.083	0.5556
11/01/2021	254	251	252.5	1.2910	252.503	0	-2.078	1
12/01/2021	254	250	252.33	1.8619	252.339	-0.0918	-2.180	1.6667
04/02/2021	255	251	253.17	1.7224	253.172	-0.4059	-1.924	1.4444

Tables 1 and 2 show the first ten statistical features extracted from daily input and battery voltage values recorded by UPS, which are Maximum, Minimum, Mean, Standard Deviation, Root Mean Square (RMS), Skewness and Kurtosis, and Mean Absolute Deviation (MAD). Figure 1 shows a graphical representation of the features generated for each variable.

Retrieval Number: 100.1/ijitee.F987113060524
DOI: 10.35940/ijitee.F9871.13060524
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

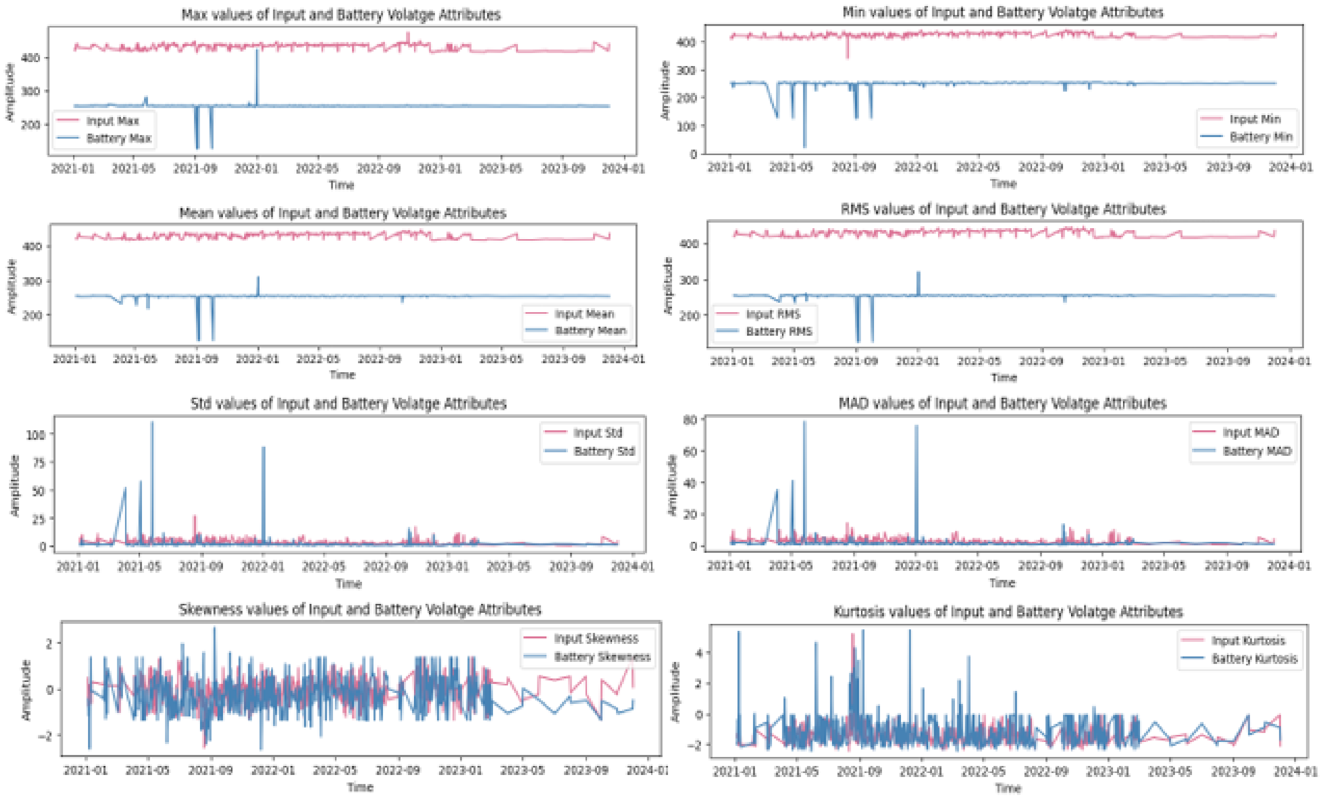

Figure 1. All features for both the Input and Battery Variables

IV ML MODELS DEVELOPED FOR FAULT PREDICTION

Machine learning methods were employed to process the statistical features of input voltage and battery voltage. For the classification task, including fault identification, the most popular and appropriate classifiers were used. In the modelling section, the machine learning models will be trained on the training subset and their performance will be tested against the unknown testing subset, resulting in confusion matrices and learning curves.

A. SVM Model for UPS Fault Prediction

The SVM algorithms were designed with the following features: Mean, RMS, Maximum, and Minimum for input voltage data, and Minimum, Mean, RMS, and Kurtosis for battery voltage data. To boost model performance, the model's hyperparameters were fine-tuned. The radial basis function kernel was employed since it possessed the lowest margin of error and the most iterations (1,000,000). The model's performance is evaluated using multiple metrics, including Accuracy, F1-score, Recall, and Precision.

a. Model Evaluation

Tables 3 and 4 illustrate the performance measures derived from the obtained SVM fault classification for the input voltage and battery voltage features, respectively. Tables 3 and 4 also demonstrate the performance for the training dataset and the validation or testing dataset, respectively.

Table 3: SVM Evaluation Metrics for the Input Voltage Attribute

Model	Training			
	Accuracy	**F1-Score**	**Recall**	**Precision**
SVM - Radial Kernel	0.9577	0.9211	0.8611	1.0000
	Testing			
	0.9624	0.9315	0.8718	1.0000

Table 4: SVM Evaluation Metrics for the Battery Voltage Attribute

Model	Training			
	Accuracy	**F1-Score**	**Recall**	**Precision**
SVM - Radial Kernel	0.9675	0.9140	0.8738	0.9732
	Testing			
	0.9474	0.8679	0.8214	0.9200

Table 3 shows that the radial basis function (RBF) kernel was calibrated and produced the smallest error margins, with Accuracy, F1-Score, Recall, and Precision values of 0.9624, 0.9315, 0.8718, and 1.0000 for the input voltage dataset. Furthermore, the testing data Accuracy, F1-score, Recall, and Precision scores for the battery voltage dataset presented in Table 4 were slightly lower than those for the input voltage dataset.

b. Confusion Matrix

The confusion matrices for the input voltage and battery voltage datasets are shown in Figures 2 and 3. The matrices

Retrieval Number: 100.1/ijitee.F987113060524
DOI: 10.35940/ijitee.F9871.13060524
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

display the number of correct and incorrect predictions made by the model for each class. The absolute numbers of classification success rates are recorded on the diagonals of the confusion matrices, while the misclassified samples are in the parts of the matrix based on the distribution of classification errors. For each confusion matrix, the predicted label represents the expected value of the provided sample by the trained SVM algorithm. In contrast, the accurate label represents the desired value of that sample.

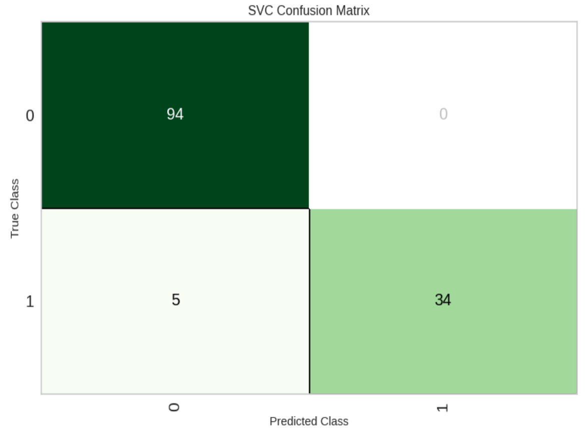

Figure 2: SVM Confusion Matrix for the Input Voltage Testing Data

Figure 2 shows that the SVM model accurately identified all 94 failure-free data points (True Positive). Similarly, 34 True Negatives means that the model accurately identified 34 faulty data points while misclassifying 5 of them as no fault (False positives).

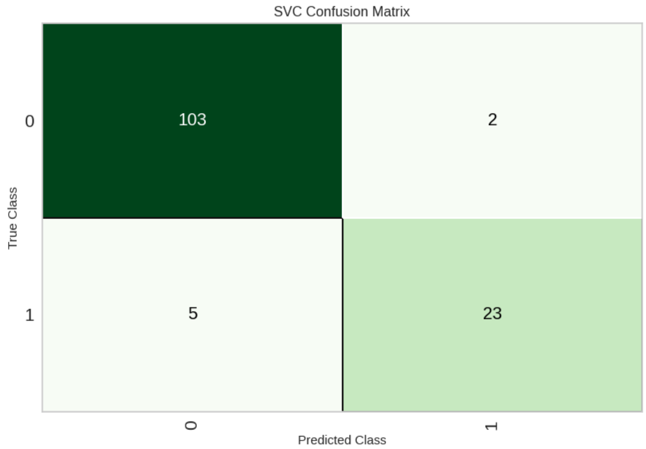

Figure 3: SVM Confusion Matrix for the Battery Voltage Testing Data

Additionally, Figure 3 illustrates that the SVM model accurately identified 103 fault-free data points (True Positive) and misclassified two fault-free data points as faults (False Negative) for the Battery Voltage dataset. Similarly, 23 True Negatives means that the model accurately identified 23 fault data points while incorrectly classifying 5 fault data points as no-fault data points (False Positives). Overall, the model appears to be performing reasonably well; however, it generates a significant number of false negatives for various classes.

c. Cross Validation

Cross-validation is a technique for evaluating a machine learning model's performance by training it on multiple subsets of data and then evaluating it on the remaining data. The models were validated and tested using the k-fold cross-validation approach. In this investigation, a ten-fold cross-validation or a k value of ten was used. Using this cross-validation approach, the dataset is randomly divided into test and training data and then divided into k groups. The model is validated on one of the groups, then training is done on the remaining groups.

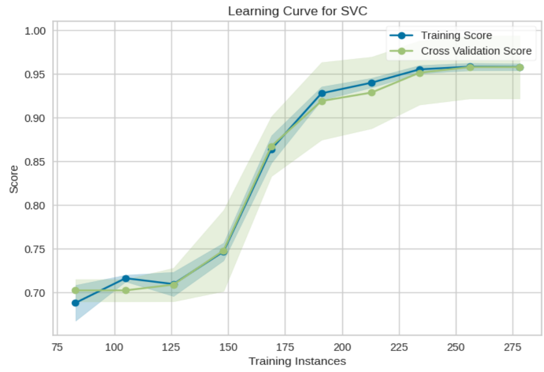

Figure 4: Plot of SVM Learning Curve for the Input Voltage Attribute

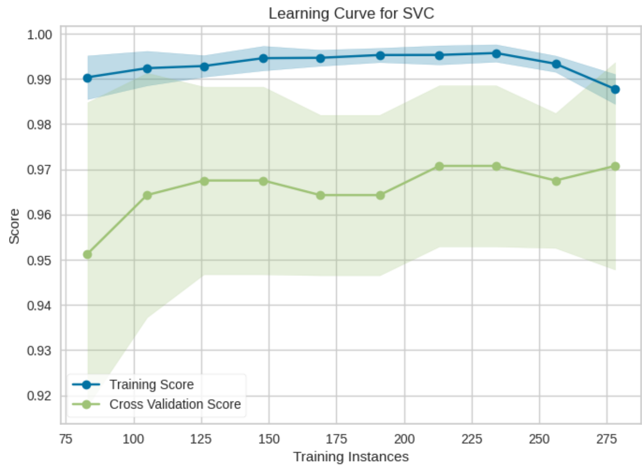

Figure 5: Plot of SVM Learning Curve for the Battery Voltage Attribute

Figures 4 and 5 compare a model's performance on training and testing data over a range of training instances. Figure 4 shows that the model is well-fitting, as indicated by the training and validation scores, which increase to a point of stability with a slight difference between the two final score values. Figure 5 shows that the training score remains exceptionally high, regardless of the number of training instances, and the cross-validation score increases over time. There is also a considerable variance between the training and testing scores, indicating that the SVM model is overfitting the data.

B. Decision Tree Model for UPS Fault Prediction

The DT classification was performed to categorise the data into binary targets and to construct a classification model capable of correctly distinguishing between fault-free and faulty data points.

The criterion is the parameter that determines how the impurity of a split will be measured. The Gini impurity was the criterion parameter

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

used. Also, the "min_samples_split" parameter, which is the minimum number of samples required to split an internal node, was set to 20.

a. Model Evaluation

The classification performance results for the Decision Tree algorithm are summarized in Tables 5 and 6. Table 5 shows that the testing Accuracy, F1-score, Recall, and Precision values for the input voltage dataset are 0.9774, 0.9620, 0.9744, and 0.9500, respectively. Additionally, the battery voltage dataset yields accuracy, F1-Score, Recall, and Precision values of 0.9699, 0.9286, 0.9286, and 0.9286, respectively, as shown in Table 6.

Table 5: DT Evaluation Metrics for the Input Voltage Attribute

Model	Training			
	Accuracy	F1-Score	Recall	Precision
Decision Tree Classifier	0.9706	0.9486	0.9444	0.9633
	Testing			
	0.9774	0.9620	0.9744	0.9500

Table 6: DT Evaluation Metrics for the Battery Voltage Attribute

Model	Training			
	Accuracy	F1-Score	Recall	Precision
Decision Tree Classifier	0.9352	0.8175	0.7476	0.9399
	Testing			
	0.9699	0.9286	0.9286	0.9286

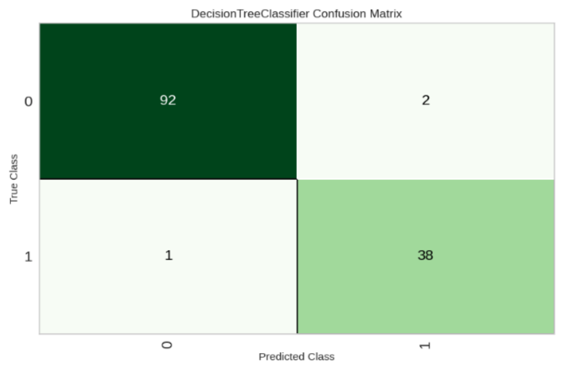

Figure 6: DT Confusion Matrix for the Input Voltage Attribute

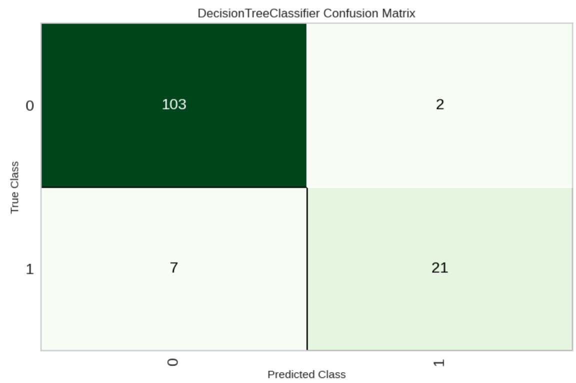

Figure 7: DT Confusion Matrix for the Battery Voltage Attribute

Figure 6 shows that the DT model made 92 correct predictions and 2 wrong predictions for the "No Failure" class.

In addition, the model made 38 correct predictions and 1 incorrect prediction for the "Fault" class. Similarly, for the "No Failure" class, 102 data points were correctly classified, and two were incorrectly classified. In Figure 7, data points in the "Fault" class were accurately classified as faults, while seven were wrongly labelled as "No Fault".

b. Cross Validation

The learning curve in Figure 8 indicates substantial test variability and a high score up to approximately 280 instances; however, after this threshold, the model begins to converge on an F1 score of around 0.98. Since the training and test results have not yet converged, this model may benefit from additional training data. Finally, this model suffers mostly from errors caused by variance (the test data scores are more variable than the training data), suggesting that the model may be overfitting.

Figure 9 also shows that the model initially has a very low training score, which steadily increases as more training examples are added. Both the training and testing scores begin to decline at 250 samples, indicating that adding more examples will enhance the model's convergence and stability. The cross-validation graphs reveal the consistency and stability of the model's performance.

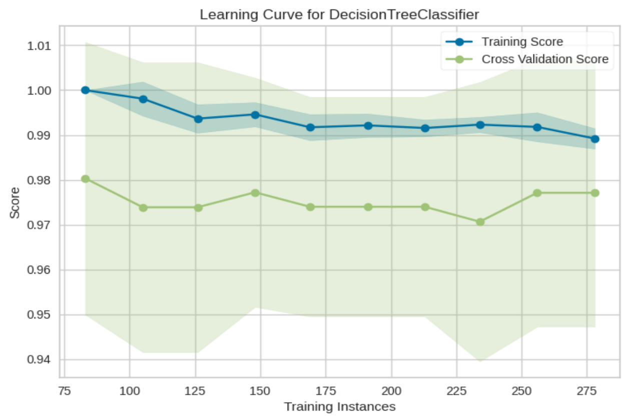

Figure 8: Plot of DT Learning Curve for the Input Voltage Attribute

Retrieval Number: 100.1/ijitee.F987113060524
DOI: 10.35940/ijitee.F9871.13060524
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

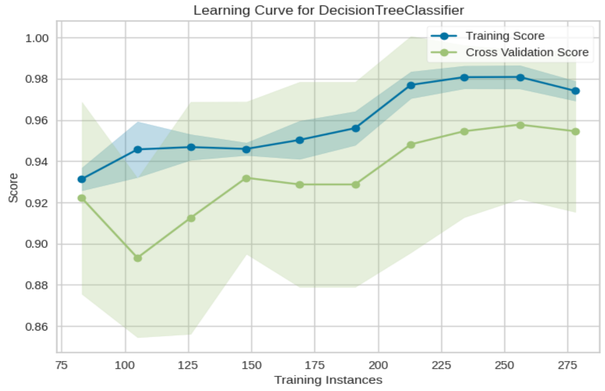

Figure 9: Plot of DT Learning Curve for the Battery Voltage Attribute

The dependability of any developed model is determined by its capacity to generalize adequately based on test outcomes. As a result, a quantitative comparison of the proposed ML models was performed, as shown in Table 7. Individual models were found to be capable of predicting faults in the UPS. DT outperformed the SVM model in terms of Accuracy, F1-score, Recall, and Precision.

In Table 7, the DT-based prediction model achieved an Accuracy of 0.9774, F1-score of 0.9620, Recall of 0.9744, and Precision of 0.9500 for the Input Voltage data, compared to SVM, which yielded an Accuracy of 0.9624, F1-score of 0.9315, Recall of 0.8718, and Precision of 0.9500. The DT model's accuracy rating (0.9774) indicates that the model is highly dependable in forecasting failures in the TDS UPS system. Similarly, for the battery voltage data, the DT model demonstrated superior performance compared to the SVM model. The importance of input voltage and battery voltage-based variables in fault prediction was also discovered, with input voltage features proving to be better predictors of faults.

Table 7a: Summary of comparison between ML models

Models	Input data			
	Accuracy	F1-Score	Recall	Precision
SVM	0.9624	0.9315	0.8718	1.0000
DT	0.9774	0.9620	0.9744	0.9500
	Battery Data			
SVM	0.9474	0.8679	0.8214	0.9200
DT	0.9699	0.9286	0.9286	0.9286

Table 7b Error Metrics between SVM and DT Models

Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE
Input Data						
SVM	0.42%	1.18%	4.45%	99.97%	0.04%	0.21%
DT	0.55%	2.25%	12.25%	98.23%	0.03%	0.31%
Battery Data						
SVM	0.52%	2.28%	5.26%	90.57%	0.04%	0.51%
DT	1.24%	1.84%	6.07%	90.22%	0.13%	0.63%

V CONCLUSION

This study introduced the concept of fault prediction and demonstrated how businesses can utilise it to improve their maintenance cycles. The test result demonstrates that, with an accuracy of 97.74% for input voltage features and 96.99% for battery voltage features, the proposed decision tree's fault classification accuracy is superior to that of SVM. SVM and Decision Tree were employed as trained classification models, and the recorded UPS data from Ghana Gas were used in the ML modelling to create the prediction models. Two variables (input voltage and battery voltage) were evaluated using data obtained from the UPS during a three-year and three-month period. Furthermore, the eight statistical features were derived from both the input voltage and the battery voltage to further characterise the data. In pattern classification, the capabilities of machine learning algorithms were used. In addition, the primary input parameters used by the models were the Mean, Min, Max, RMS, and Skewness. After modelling, the performance of each algorithm for fault classification was examined. To assess the efficacy and capacities of the constructed models, four performance metrics were used: Accuracy, F-1 Score, Recall, and Precision. While the model performed well overall, it was discovered that it was unable to predict all classes with the same level of accuracy using battery voltage data. This implies that there might be opportunities to enhance the model's functionality.

RECOMMENDATIONS

The proposed method's robustness and accuracy demonstrate its potential for protecting UPS systems in major power companies. It provides companies with additional support for equipment reliability decision-making, enabling them to remain more competitive in the market. It is necessary to conduct further studies to forecast equipment failure times, as well as to perform real-time online calibration monitoring as data is being gathered. Additionally, the model facilitates the analysis of equipment data records, enabling the detection of faults without prior knowledge of the equipment's status. Furthermore, the model is generalizable to any number of UPS systems. However, to improve the model in the future, it is advised to implement an online algorithm to diagnose and prognose the equipment during operation. An artificial neural network class, such as the Multi-Layer Perceptron (MLP), may be utilised to enhance the performance of the HMM.

DECLARATION STATEMENT

Authors declare that there is no conflict of interest whatsoever. The authors are solely responsible for this research work, and it was not funded by any industry or other external sources.

Funding	No, I did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval or consent to participate, as it presents evidence.
Availability of Data and Materials	Yes, Data can be accessed upon request via email: bodoi@umat.edu.gh or isaacmawulidoe@gmail.com.
Authors Contributions	Each author has made an independent contribution to the article. The individual contributions of each author are presented below for clarity and transparency. Isaac M. Doe formulated the research problem, analysed, and discussed it. John K. Annan and Benjamin Odoi supervised the entire research work.

REFERENCES

1. Nan, B., Chen, L., Rodrigo, N.D., Borodin, O., Piao, N., Xia, J., Pollard, T., Hou, S., Zhang, J., Ji, X. and Xu, J., 2022. Enhancing Li+ transport in NMC811|| graphite lithium-ion batteries at low temperatures by using low-polarity-solvent electrolytes. *Angewandte Chemie International Edition*, *61*(35), p.e202205967. https://doi.org/10.1002/anie.202205967
2. Lukovic, M., Lukovic, V., Belca, I., Kasalica, B., Stanimirovic, I. and Vicic, M., 2016. LED-based Vis-NIR spectrally tunable light source: the optimisation algorithm. *Journal of the European Optical Society-Rapid Publications*, *12*, pp. 1-12. https://doi.org/10.1186/s41476-016-0021-9
3. Guerrero, J.M., Matas, J., de Vicuna, L.G., Castilla, M. and Miret, J., 2007. Decentralized control for parallel operation of distributed generation inverters using resistive output impedance. *IEEE Transactions on Industrial Electronics*, *54*(2), pp. 994-1004. https://doi.org/10.1109/TIE.2007.892621
4. Low, K.S. and Cao, R., 2008. Model predictive control of parallel-connected inverters for uninterruptible power supplies. *IEEE Transactions on Industrial Electronics*, *55*(8), pp. 2884-2893. https://doi.org/10.1109/TIE.2008.918474
5. Martin, F. and Aguado, J.A., 2003. Wavelet-Based ANN Approach for Transmission Line Protection *IEEE Transactions on Power Delivery*, *18*(4), pp.1572-1574. https://doi.org/10.1109/TPWRD.2003.817523
6. Sharafi, A., Jafarian, P. and Sanaye-Pasand, M., 2010, May. A combined algorithm for high-speed transmission-line protection based on travelling waves. In *2010, 9th International Conference on Environment and Electrical Engineering*, pp. 132-135). https://doi.org/10.1109/EEEIC.2010.5489973
7. Park T, Efros AA, Zhang R, Zhu JY. Contrastive learning for unpaired image-to-image translation. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Springer International Publishing, Part IX 16 2020, pp. 319-345. https://doi.org/10.1007/978-3-030-58545-7_19
8. Katipamula, S. and Brambley, M.R., 2005. Methods for fault detection, diagnostics, and prognostics for building systems—A review, part II. *Hvac&R Research*, *11*(2), pp.169-187. https://doi.org/10.1080/10789669.2005.10391133
9. Awadallah, M.A. and Morcos, M.M., 2003. Application of AI Tools in Fault Diagnosis of Electrical Machines and Drives: An Overview. *IEEE Transactions on Energy Conversion*, *18*(2), pp. 245-251. https://doi.org/10.1109/TEC.2003.811739
10. Giuntini, L. and Brioschi, M., 2015, March. Markov Chain Analysis for Failure Prediction of Power Converters. In *2015 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1336-1341. IEEE. https://doi.org/10.1109/ICIT.2015.7125283
11. Eusgeld I, Fraikin F, Rohr M, Salfner F, Wappler U. Software reliability. Dependability Metrics: Advanced Lectures. 2008:104-25. https://doi.org/10.1007/978-3-540-68947-8_10
12. Salfner, F. and Tschirpke, S., 2008. Error Log Processing for Accurate Failure Prediction. *WASL*, *8*, pp.4.
13. Zhang, X., Zhou, X., Lin, M. and Sun, J., 2018. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6848-6856). https://doi.org/10.1109/CVPR.2018.00716
14. Tao L, Cinquanta E, Chiappe D, Grazianetti C, Fanciulli M, Dubey M, Molle A, Akinwande D. Silicene field-effect transistors operating at room temperature. Nature nanotechnology. 2015 Mar;10(3): pp. 227-31. https://doi.org/10.1038/nnano.2014.325
15. Guardado, J.L., Naredo, J.L., Moreno, P. and Fuerte, C.R., 2001. A Comparative Study of Neural Network Efficiency in Power Transformer Diagnosis Using Dissolved Gas Analysis. *IEEE Transactions on Power Delivery*, *16*(4), pp. 643-647. https://doi.org/10.1109/61.956751
16. Aeon, C., Breadsell, J., Morrison, G. and Byrne, J., 2019. Shifting home energy consumption through a holistic understanding of the home system of practice. *Decarbonising the built environment: Charting the transition*, pp. 431-447. https://doi.org/10.1007/978-981-13-7940-6_23
17. Giuntini, L. and Brioschi, M., 2015, March. Markov Chain Analysis for Failure Prediction of Power Converters. In *2015 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1336-1341, IEEE. https://doi.org/10.1109/ICIT.2015.7125283Jan, S.U., Lee, Y.D., Shin, J. and Koo, I., 2017. Sensor fault classification based on support vector machine and statistical time-domain features. *IEEE Access*, *5*, pp. 8682-8690. https://doi.org/10.1109/ACCESS.2017.2705644
18. Bigdeli, M., Vakilian, M. and Rahimpour, E., 2012. Transformer Winding Fault Classification Based on Transfer Function Analysis Using Support Vector Machine. *IET electric power applications*, *6*(5), pp. 268-276. https://doi.org/10.1049/iet-epa.2011.0232
19. Linzen, D., Buller, S., Karden, E. and De Doncker, R.W., 2005. Analysis and evaluation of charge-balancing circuits on performance, reliability, and lifetime of supercapacitor systems. *IEEE transactions on industry applications*, *41*(5), pp. 1135-1141. https://doi.org/10.1109/TIA.2005.853375
20. Litchfield, N.J., Villamor, P., Dissen, R.J.V., Nicol, A., Barnes, P.M., A. Barrell, D.J., Pettinga, J.R., Langridge, R.M., Little, T.A., Mountjoy, J.J. and Ries, W.F., 2018. Surface rupture of multiple crustal faults in the 2016 Mw 7.8 Kaikōura, New Zealand, Earthquake. *Bulletin of the Seismological Society of America*, *108*(3B), pp. 1496-1520.
21. Lind, H. and Muyingo, H., 2012. Building Maintenance Strategies: Planning Under Uncertainty, *Property Management*, *30*(1), pp. 14-28. https://doi.org/10.1108/02637471211198152
22. Aamir, M., Kalwar, K.A. and Mekhilef, S., 2016. Uninterruptible power supply (UPS) system. *Renewable and sustainable energy reviews*, *58*, pp. 1395-1410. https://doi.org/10.1016/j.rser.2015.12.335
23. Cherkassky V, Mulier FM. Learning from data: concepts, theory, and methods. John Wiley & Sons; 2007 Sep 10. https://doi.org/10.1002/9780470140529
24. Suykens JA. Nonlinear modelling and support vector machines. In IMTC 2001. proceedings of the 18th IEEE instrumentation and measurement technology conference. Rediscovering measurement in the age of informatics (Cat. No. 01CH 37188) 2001 May 21 (Vol. 1, pp. 287-294). IEEE.
25. Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison FJ, Antonanzas-Torres F. Review of photovoltaic power forecasting. *Solar Energy*. 2016;136: pp. 78-111. https://doi.org/10.1016/j.solener.2016.06.069
26. Sharma G, Tripathi V, Mahajan M, Srivastava AK. Comparative analysis of supervised models for diamond price prediction. In: *Proceedings of the Conflu 2021 11th Int Conf Cloud Comput Data Sci Eng*. Published online 2021: pp. 1019-1022. https://doi.org/10.1109/Confluence51648.2021.9377183
27. Fontana F A, Mäntylä MV, Zanoni M, Marino A. Comparing and experimenting machine learning techniques for code smell detection. *Empir Softw Eng Vol*. 2016;21: pp. 1143–1191. https://doi.org/10.1007/s10664-015-9378-4
28. Muhammad IA, Fabio P, Lin S, Qing W. Machine learning techniques for code smell detection: A systematic literature review and meta-analysis. *Inf Softw Technol*. 2019;108 (1), pp. 15-138. https://doi.org/10.1016/j.infsof.2018.12.009
29. Ofosu A. R, Odoi B, Asamoah M, "Electricity consumption forecast for Tarkwa using autoregressive integrated moving average and adaptive neuro fuzzy inference system", *Serbian Electr Eng*. 2021;18(1): pp. 75-94. https://doi.org/10.2298/SJEE2101075O
30. Twumasi-Ankrah, S., Odoi, B., Adoma Pels, W. and Gyamfi, E.H., 2019. Efficiency of imputation techniques in univariate time series.
31. Odoi, B., Brew, L. and Attafuah, C., 2021, "The use of Principal Component Regression and Time Series Analysis to Predict Nitrous Oxide Emissions in Ghana", Energy Res. J., 2021, 12(1), pp.1-12 https://doi.org/10.3844/erjsp.2021.1.12
32. Odoi B., Boahen A. D. and Brew L. (2024), Predicting Nitrous Oxide Emissions in Ghana using Long Short-term Memory and Gated Recurrent Neural Network, *eNergetics 2023, 9th Virtual International Conference on Science, Technology and Management in Energy, Conference, 81.*
33. Rathakrishnan V, Bt. Beddu S, Ahmed AN. Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using a boosting machine and learning algorithms. *Sci Rep*. 2022;12(1): pp. 1-16. https://doi.org/10.1038/s41598-022-12890-2
34. Liu Q, Wang X, Huang X, Yin X. Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. *Tunn Undergr Sp Technol*. 2020; 106: pp.

1-15. https://doi.org/10.1016/j.tust.2020.103595
35. Nguyen H, Vu T, Vo TP, Thai HT. Efficient machine learning models for the prediction of concrete strengths. *Constr Build Mater*. 2021; 266: pp.1-17. https://doi.org/10.1016/j.conbuildmat.2020.120950
36. Saucedo-Dorantes J.J, Arellano-Espitia F, Delgado-Prieto M, Osornio-Rios RA. Diagnosis methodology based on deep feature learning for fault identification in metallic, hybrid and ceramic bearings. Sensors, 2021 Aug 30;21(17):5832. https://doi.org/10.3390/s21175832
37. Agyare, S., Odoi, B. and Wiah, E. N. Predicting Petrol and Diesel Prices in Ghana, A Comparison of ARIMA and SARIMA Models. *Asian Journal of Economics, Business and Accounting*, 2024, *24*(5), pp. 594-608. https://doi.org/10.9734/ajeba/2024/v24i51333
38. Twumasi-Ankrah, Sampson, Benjamin Odoi, Wilhemina Adoma Pels, and Eric Herrison Gyamfi, "Efficiency of Imputation Techniques in Univariate Time Series", International Journal of Science, Environment and Technology, 8 (3), 2019, pp. 430 – 453.
39. Ofosu RA, Odoi B, Boateng DF, Muhia AM. Fault Detection and Diagnosis of a 3-Phase Induction Motor Using Kohonen Self-Organising Map. *Jurnal Nasional Teknik Elektro.* 2023 Mar 31. 12(1), pp. 30-41.

AUTHORS PROFILE

Ing. Isaac M. Doe, PE is an MSc Student in the Department of Electrical and Electronic Engineering at the University of Mines and Technology, UMaT, Tarkwa, Ghana. He is a member of the Ghana Institution of Engineering (GhIE). His research interests include SCADA Systems, Control Systems and Networking, Power System Controls, Gas Metering and Data Transmission, and Gas Operations.

Dr. John Kojo Annan is a Senior Lecturer and the Head of Department of Electrical and Electronic Engineering at the University of Mines and Technology, UMaT, Tarkwa, Ghana. He is a member of the International Association of Engineers and the Society of Petroleum Engineers. He has guided the research work of many students in engineering at the PhD, MPhil, MSc, and BSc levels. His research interests include, power generation and utilization, control systems design and analysis, and renewable energy and resources.

Dr. Benjamin Odoi is a Lecturer, contemporary statistician and a prominent data analyst in the Department of Mathematical Sciences at the University of Mines and Technology, UMaT, Tarkwa, Ghana. He holds the degrees of BSc from UMaT, an MSc from the University of Peradeniya (UoP), Sri Lanka, an MPhil, and a PhD from UMaT. He also holds a Certificate in Advanced Study in Data Science and Big Data Analytics from the Ghana Institute of Management and Public Administration (GIMPA), Ghana. He is a member of Laboratory for Interdisciplinary Statistical Analysis (LISA 2020 Global Network), Colorado, United States of America, the Officer for Monitoring, Evaluation, and Membership for LISA 2020 Global Network Ghana Region, Nigerian Statistical Association (NSA), Nigeria, and Institute of Applied Statistics (IASSL), Sri Lanka. He provides technical advice in areas ranging from sample size calculation and data collection design to data management procedures, data analysis, statistical and mathematical modelling, and more. He has taught many statistical and related courses at both the undergraduate and postgraduate levels. He has also guided the research work of many students in applied statistics, mathematical modelling, public health, financial Statistics, logistics, and engineering at the PhD., MPhil, MSc, and BSc levels. His computer programming skills include using SAS, R, Minitab, SPSS, EViews, XLMiner, and PASALADE. His research interests and consultancy work cover big data analytics, applied statistics, climate change, financial statistics, biostatistics, and statistical engineering.