Available *online* at : http://jnte.ft.unand.ac.id/

Jurnal Nasional Teknik Elektro

| ISSN (Print) 2302-2949 | ISSN (Online) 2407-7267 |

Prediction of Transformer Oil Insulation Strength Based on Physical Quantities Using Ensemble Machine Learning Model

M.G. Indabawa [1], N. Bashir [2], R. A. Abdulkadir [2]

[1] *Technical Department, Kano Elertricity Distribution Company, Kano, 700001, Nigeria*
[2] *Department of Electrical Engineering, Kano University of Science and Technology, PMB 3244, Wudil, Kano, Nigeria*

ARTICLE INFORMATION

Received: May 25, 2026
Revised: June 23, 2026
Accepted: June 23, 2026
Available online: July 31, 2026

KEYWORDS

Transformer oil, Breakdown voltage, ensemble machine learning, long short-term memory neural network, preidction

CORRESPONDENCE

Phone: +2348144441912

E-mail: nour@kustwudil.edu.ng

A B S T R A C T

Existing machine learning models for predicting transformer oil breakdown voltage (BDV) often depend on chemical properties, which require laboratory tests, thus making them time-consuming and costly. Furthermore, these models are typically developed using standalone methods, which may fail to achieve satisfactory prediction accuracy. This study introduces an interpretable ensemble prediction model for transformer oil BDV, leveraging physical properties instead of chemical properties. The ensemble model integrates four standalone models: multilinear regression (MLR), multilayer perceptron (MLP) neural network, support vector machine (SVM), and long short-term memory (LSTM) neural network, to enhance prediction accuracy. A Local Interpretable Model-Agnostic Explanation framework was used to provide insights into the model's predictions. The model's performance was assessed using key metrices, including the correlation coefficient (R), determination coefficient (R^2), root mean squared error (RMSE), and mean absolute error (MAE). Results revealed that the ensemble model outperformed individual models, achieving R = 0.9796, R^2 = 0.9596 during training and R = 0.9781, R^2 = 0.9567 during testing, along with RMSE = 1.7678, MAE = 1.2972 during training, and RMSE = 1.7347, MAE = 1.2985 during testing. The proposed model reliably predicted transformer oil BDV using physical properties, eliminating complex experiments while offering transparency to enhance applicability in the power industry

INTRODUCTION

The transformer is the most important and expensive component in power system networks, due to its ability to use electromagnetic induction and connect circuits with different voltage levels in the most efficient way. Transformer oil is vital in electrical power transformers, functioning as both an insulator to endure high voltages and a coolant to dissipate operational heat [1]. However, over time the ability of insulating oils to serve as effective insulators and coolants are adversely affected by electrical stress and environmental condition; thus, affecting its breakdown voltage (BDV). Insulation plays a major role in the proper functioning of High Voltage equipment [2]. The BDV is the minimum voltage required for a small portion of the insulator to start conducting, thus losing its electrical insulating property and becoming a conductor. It is used to measure the integrity of the insulating medium; a low BDV indicates the existence of contaminants such as air bubbles, moisture, and suspended solid particles.

Temperature is one of the main factors affecting transformer oil BDV; as the temperature of the transformer oil increases, its breakdown voltage decreases which can subsequently leading to transformer failure [3]. This is due to the fact that as the temperature increases, the viscosity of the oil decreases, and the mobility of the charged particles in the oil increases. This can lead to an increased likelihood of electrical breakdown and a decrease in the breakdown voltage. Transformer loading also affects the its oil insulating properties. The total amount of power a transformer delivers is termed as “loading”. As a transformer's loading increases, the temperature of the transformer also increases. This can cause the insulating properties of the transformer oil to deteriorate, which can lead to a decrease in the breakdown voltage of the oil [4]. Moreover, the loading of the transformer is not only related to temperature but also to many other factors affecting breakdown voltage such as acid number, ageing, flash point and moisture. Literature has shown that as transformer loading rises, breakdown voltage reduction intensity increases [5]. Furthermore, the breakdown voltage of aged transformer oil decreases with ageing time/service period [6-8]. This is due to the fact that ageing processes can lead to the formation of internal partial discharge and contaminants such as acids, sludge, and carbonaceous particles, which in turn increase the likelihood of electrical breakdown. Similarly, the ageing process is accelerated by factors such as temperature hot spots, highly localized electric

https://doi.org/10.25077/jnte.v15n2.1605.2026

fields, moisture in paper, and dissolved oxygen in oil. The decrease in breakdown voltage is caused by an increase in particulate impurities, which increases the moisture and makes the oil non-homogeneous. As a result, the transformer oil resistance decreases, leading to a decrease in the maximum value of the breakdown voltage of the transformer oil. Additionally, ageing processes can also cause chemical changes in the oil that reduce its dielectric strength.

The BDV of transformer oil, which reflects its dielectric strength, is crucial for evaluating the condition and reliability of power transformers [9, 10]. Therefore, monitoring transformer oil BDV is critical to ensuring the safe, reliable, and effective operation of the power system. In addition, its accurate prediction is essential for effective condition monitoring, fault detection, timely maintenance, and prevention of costly failures thus ensuring an uninterrupted power supply [11].

Traditionally, the prediction of transformer oil BDV has been tackled using empirical and statistical methods. However, due to the growing complexity of power systems and the demand for more precise and automated prediction models, Machine Learning (ML) techniques have gained attention [12-14]. Consequently, numerous researchers have developed data-driven models to enhance accuracy and robustness in these prediction tasks. Multiple Linear Regression (MLR), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), Long Short-Term Memory (LSTM) and Ensemble are the most widely used algorithms in the field of machine learning and deep-learning each offering unique capabilities for regression and time-series analysis tasks [15, 16]. Although, several other algorithms and factors affecting transformer oil breakdown voltage have been used for the prediction of the breakdown voltage.

In the literature, most of the recent researches focused on the prediction of transformer oil breakdown voltage based on either chemical or geometrical factors obtained from oil sample analyses in the laboratory such as moisture, acidity, dielectric loss, stress volume, interfacial tension, dissolve gases etc., but little research has considered physical factors. For example, in a study by [17], Artificial Neural Networks (ANNs) were employed to predict BDV using eight input variables, including gap space, the barrier location relative to the high-voltage electrode %, barrier diameter, barrier thickness, electrode configurations, the presence of contaminating particles, the weight of the contaminating particles and the temperature of the insulating oil. The study reported a high prediction accuracy of 95% but the ANN model needs data obtained via laboratory tests and dimensions that are not easily accessible. [18] employed Weibull distribution and normal distribution and drove a statistical model for prediction breakdown voltage. The accuracy was unexpectedly low despite the study using the complex Taguchi method for design experiments and the response surface methodology. Studies by [19-22] were based on chemical composition of the transformer oil where several factors such as interfacial tension, tan delta , acid number, dissipation factor, permittivity, moisture, and dissolved gases were used for predicting BDV using different AI models. The proposed models achieved more than 90% accuracy, however, the input variables used were difficult to access due to need to conduct series of laboratory tests. Similarly, nonlinearity in the data may not be captured. Furthermore, [23-25] developed hybrid deep neural network and mathematical models that predicts breakdown voltage via chemical anf physical factors . The accuracy of the models were good but were not cost-effective; while saving the cost of breakdown voltage test, the cost of testing chemical parameters added to the maintenance cost. Also, Kernel extreme learning machine (KELM) based on relative transformation (RT) and Kernel principal component analysis (KPCA) to predict BDV and optimized the parameters using differential evaluation algorithm. Many input variables and engineering features used in some of the models not only account for the high accuracy of the model but also for good generalization ability. However, the models are costly and time-consuming. It can be seen most of the transformer oil BDV prediction techniques reported which employed the use of AI in the literature involved the use of variables whose results have to be obtained from laboratory tests making methods costly and complex.

In this work, a novel technique of predicting the transformer oil breakdown voltage using only physical quantities (which do not need laboratory tests) and robustic Ensemble machine learning techniques is presented. The quantities have strong correlation with transformer BDV. Dataset from these quantities were used to train four machine learning models and finally combined using LSTM-Ensemble to predict BDV. Performance of the networks were analysed and compared using a variety of metrices..

METHODS

Figure 1 depicts the methodology employed in this work. The dataset was pre-processed for the ML modelling, which was then split into training and testing datasets. The training dataset was used to train multiple standalone models, including Multiple Linear Regression (MLR), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM). These individual models were then combined into an ensemble model to exploit their collective strengths and improve the prediction performance. The trained models and ensemble were evaluated using the testing dataset to validate their effectiveness and generalization. Finally, model interpretation was performed using Local Interpretable Model-agnostic Explanations (LIME) to provide insights into the decision-making process of the models, enhancing transparency and trust in the predictions.

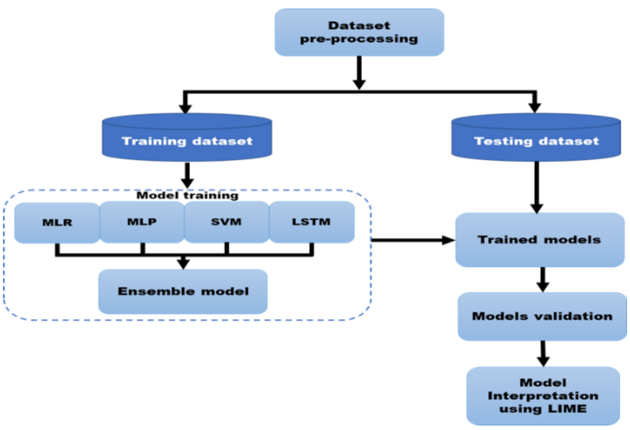

Figure 1. Methodology

A. Data Processing

In this study, 8-years data on the transformer oil BDV tests, service period, oil temperature and loading were used. The data were obtained from the Kano Electricity Distribution Company (KEDCO) 33/11 kV injection sub-stations on 19 medium voltage power transformers. BDV test results conducted on each of 19 transformer units over a period of 96 months (8 years) were obtained. The BDV tests carried out followed the IEC 60156 testing protocol. Each BDV value was averaged from six repeated tests to ensure accuracy. The service period refers to a length of time obtained by subtracting the manufacture year or year of operation of a transformer from the year the transformer oil samples were tested for BDV. The service period of the transformers in this study ranged from 24 to 456 months. The oil temperature (measured in degrees Celsius) was recorded from the transformers' direct mount transformer oil temperature indicator (OTIs). The transformer loading data (in megawatts) were collected from the wattmeter measurements on the 11 kV incomer panel of the nineteen (19) 33/11 kV power transformers within the injection substations. Hourly data of the transformer oil temperature and loading data were obtained from the transformers' respective sub-station logbooks/reading sheets. The data for 8 years was collected. Close to 72,000 data points were collected for the four quantities (transformer BDV, service period, oil temperature and loading)

The collected data underwent a thorough pre-processing and cleaning phase, which involved replacing outliers using winsorization 5th and 95th percentiles, correcting errors and inconsistencies, addressing missing data, and formatting the dataset for analysis. The oil temperature, transformer load, and service periods were used as input features for the AI models, while oil BDV served as the output. The pre-processed data were split into training and testing subsets in a 70:30 ratio. The training data (70%) were used to determine the model parameters, while the testing data (30%) were used to evaluate the models' accuracy. The AI models were developed, simulated, and assessed for performance using MATLAB version 2019b on a personal computer with a CORE i7 processor.

B. Standalone Models

1) Multi-Linear Regression (MLR)

In this work, the MLR model developed consisted of a dependent variable (Y) as the transformer oil BDV while the independent variables were the temperature (x_1), load (x_2) and service period (x_3).

2) Multi-Layer Perceptron (MLP)

In this work, the MLP consisted of an input layer having three inputs viz. temperature, loading and service period, four hidden layers (each containing 20 neurons) and a single output layer was selected after experimenting with different configurations. The output layer was the transformer oil BDV. The model was trained using the Levenberg-Marquardt algorithm; chosen for its fast convergence and ease of implementation [26]. The weights and biases were initialized randomly.

3) Support Vector Machine (SVM)

The SVM model was constructed using a Gaussian kernel function, which effectively captures the non-linear relationships between input features and the target variable. This kernel is highly flexible and adept at modelling both linear and non-linear decision boundaries. The regularization parameter (C) and gamma (γ) hyperparameter were fine-tuned to balance overfitting and underfitting. The optimal values for C and γ were determined to be 43 and 8, respectively.

4) Long Short-Term Memory (LSTM)

The LSTM network was designed using a layer-by-layer approach. Its architecture consisted of a sequence input layer, an LSTM layer, a fully connected layer, and a regression output layer. A single hidden layer with 40 LSTM units was employed, along with a maximum of 1000 epochs, a batch size of 100, and a validation frequency of 10 for faster convergence.

C. Ensemble Machine Learning

The LSTM-ensemble model was developed by integrating the trained SVM, MLP, and LSTM models as inputs to a new LSTM model. The LSTM-ENS ensemble architecture featured a hidden layer with 50 LSTM units, a maximum of 1000 epochs, and a batch size of 30.

Ensemble techniques had been widely applied in several fields such as prediction, classification, time series, and regression problems. In this study, LSTM-ensemble techniques were used based on the four different stand-alone models (MLR, SVM, LSTM and BPNN) to improve the performance of single models.

D. Model Evaluation

The performances of the models during the training and testing were evaluated using four widely used statistical metrices viz. coefficient of determination (R^2), the Root-Mean-Squared-Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The metrics were computed using equations (1) to (4).

$$R^2 = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{x}_i)^2} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum (x_i - \hat{x}_i)^2}{n}} \tag{2}$$

$$MAE = \frac{1}{n} \sum \left| x_i - \hat{x}_i \right| \tag{3}$$

$$MAPE = \frac{1}{n} \sum_1^n \left| \frac{Y_a - Y_p}{Y_a} \right| \times 100\% \tag{4}$$

for $i = 1,2,3\ldots\ldots\ldots\ldots n$.

https://doi.org/10.25077/jnte.v15n2.1605.2026

where x_i, $\hat{x}$, $\bar{x}$ and n stand for the original values, predicted values, an average value of the original data and the total number of data instances [27].

The R^2 assesses how well a regression model fits the data points. It ranges from 0 to 1, indicating the proportion of the variance in the dependent variable that is predictable from the independent(s). An R^2 value of 1 signifies a perfect fit, while an R^2 value of 0 indicates that the model does not explain any of the variation in the dependent variable [28]. The RMSE measures the average deviation of predicted values from actual values. Lower values of RMSE indicate better predictions. Moreover, lower values of MAE indicate smaller errors in prediction. However, compared to RMSE, MAE is less sensitive to outliers, since it only considers the absolute differences between predicted and actual values [29-31].

RESULTS AND DISCUSSION

A. Statistical Analysis

Table 1 shows the correlation and descriptive statistical analysis of the dependent (predictor) and independent variables (response).

Table 1. Correlation Analysis

Parameters	X_{min}	X_{max}	Correlation with BDV
Temperature (°C)	18	114	-0.65869
Load (MW)	0.6	11.4	-0.70667
Period (months)	24	456	-0.72444

It can be seen from Table 1 that all predictors have a good negative correlation coefficient with the response, thus implying a linear relationship between predictors and the response with Service period having the highest correlation with BDV. Consequently, they can be used for modelling the breakdown voltage.

To improve the regression model, a set of four different linear regression models (M_1, M_2, M_3 and M_4) were derived on the basis of significant input variables as shown in Table 2

Table 2. R^2 and SE results of the Linear Regression Models

Model	Regression Equation	R^2	Standard Error (SE)
M_1	$y = 53.01 + 0.15x_1 - 3.62x_2$	0.506	6.642
M_2	$y = 72.46 - 0.26x_1 - 0.05x_3$	0.840	3.772
M_3	$y = 65.64 - 2.07x_2 - 0.05x_3$	0.8504	3.658
M_4	$y = 67.78 - 0.08x_1 - 1.48x_2 - 0.05x_3$	0.852	3.652

It can be seen from Table 2 that Model M_4 had the highest R^2 (0.852) and lowest SE (3.62) values thus signifying the strength of the model. Hence, a combination of oil temperature, transformer loading and service period was used as input combinations for four different machine-learning models.

B. Models Prediction Performance

Table 3 presents the performance of the five models. The training and testing results indicated that the LSTM model outperformed the classical MLR, SVM, and MLP models. The model had the highest R^2, and lowest MAPE, RSME and MAE. Furthermore, the results demonstrated that the proposed LSTM-ENS model significantly outperformed the LSTM model.

Table 3. Comparison of Machine Learning model performances.

TRAINING				
Models	**MAPE**	**R^2**	**RMSE**	**MAE**
MLR	5.4227	0.8572	3.3231	2.5538
SVM	4.3667	0.9072	2.6799	1.9772
MLP	4.3302	0.9133	2.5903	1.9301
LSTM	4.1192	0.9323	2.1686	1.6265
LSTM-ENS	**2.8207**	**0.9596**	**1.7678**	**1.2972**
TESTING				
Models	**MAPE**	**R^2**	**RMSE**	**MAE**
MLR	8.9707	0.7536	4.1363	2.9314
SVM	7.5542	0.8069	3.6620	2.4592
MLP	6.4002	0.8726	2.9746	2.2637
LSTM	4.4540	0.9217	2.4606	1.8452
LSTM-ENS	**3.3350**	**0.9567**	**1.7347**	**1.2985**

Figure 2 presents a radar chart comparing the performance of the models based on their R^2 values, a metric that measures the goodness of fit. Models with R^2 values closer to 1 (100%) demonstrate higher accuracy. Consequently, the models were ranked in descending order of predictive accuracy. In this study, all models achieved over 75% accuracy in both training and testing stages, surpassing the 70% baseline for acceptable performance. The chart clearly shows that the LSTM-ENS model outperformed all other models, demonstrating the best fit for the BDV data. Conversely, the MLR model exhibited the lowest performance, likely due to its inability to capture the data's non-linear characteristics [32]. A significant improvement in performance is observed as the models progress from simple machine learning algorithms (MLP and SVM) to deep learning algorithms (LSTM). Furthermore, the ensemble algorithm (LSTM-ENS) enhanced performance by 20.3% and 10.3% in the training and testing stages, respectively, compared to the standalone algorithms. The models' predictive accuracy can be ranked in descending order (for both training and testing datasets) as follows: LSTM-ENS > LSTM > SVM > MLP > MLR.

https://doi.org/10.25077/jnte.v15n2.1605.2026

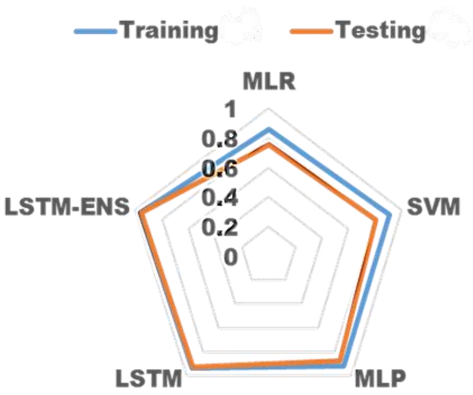

Figure 2. Models' performance comparison based on Determination Coefficient

Figure 3 presents a time series plot comparing the performance of the MLR, MLP, SVM, LSTM, and LSTM-ENS models in predicting the observed BDV data for transformer oil samples. The LSTM-ENS model closely aligned with the observed BDV data, demonstrating superior predictive accuracy and robustness in capturing complex nonlinear patterns. The LSTM model also performed well but is slightly less precise than the LSTM-ENS. In contrast, the MLP and SVM models showed moderate alignment with the experimental data, while the MLR model, being linear, failed to effectively capture the nonlinear trends, particularly at peak values. Figure 3 further emphasized the progression in performance as the models advance from simple machine learning algorithms (MLP and SVM) to deep learning models (LSTM) and, ultimately, to the ensemble approach (LSTM-ENS), which leverages the complementary strengths of the individual models to achieve improved predictions [33]. These results emphasize the importance of using advanced algorithms for complex time-series tasks, Robustness of this model can be seen consistanly from the multiple evaluation matrics,with the LSTM-ENS model proving the most reliable for accurate BDV predictions. This reliability makes the LSTM-ENS model particularly suitable for practical applications such as load forecasting and predictive maintenance in power systems, where precise modelling of nonlinear and temporal data is critical for operational efficiency and system reliability [34].

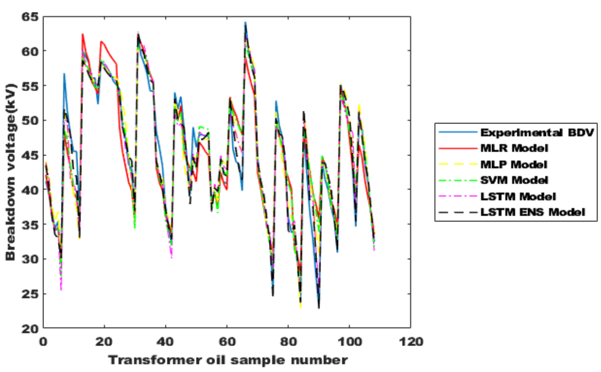

Figure 3. Times series plot of Transformer Oil BDV prediction results for all models

B. Model Prediction Interpretation using LIME

The interpretability of the proposed prediction pipeline was evaluated by generating the LIME explanation for the LSTM-ENS prediction on the testing dataset. This provided insights into how the predictors (period, temperature, and loading) influenced the model's attack predictions. Figure 4 presents the LIME interpretation, demonstrating the interaction of period, load, and temperature on the predicted BDV. The prediction, with a moderate output of 0.49, reflects the complex interaction of these variables, highlighting their roles in determining the oil's dielectric strength.

The period (0.39 < Period ≤ 0.72) had a significant negative impact on BDV, with a contribution of 0.67. This aligns with the fact that transformer oil degrades over time due to heat, moisture, and chemical reactions, leading to impurities that reduce dielectric properties. Regular oil monitoring, filtration, and regeneration are essential to mitigate aging effects.

The load (0.20 < Load ≤ 0.34) positively influenced BDV, contributing 0.24. Moderate, consistent loading helps maintain the oil's properties by minimizing thermal and mechanical stresses, which could otherwise accelerate degradation. Optimizing load distribution and operational stability can enhance transformer performance.

Temperature (0.30 < Temperature ≤ 0.45) positively affects BDV, contributing 0.38. Controlled temperatures improve oil performance by reducing viscosity and allowing impurities to settle, supporting better insulation properties. However, excessive temperatures can cause oxidation and degradation, reducing BDV. Proper cooling and temperature monitoring are essential to balance this effect.

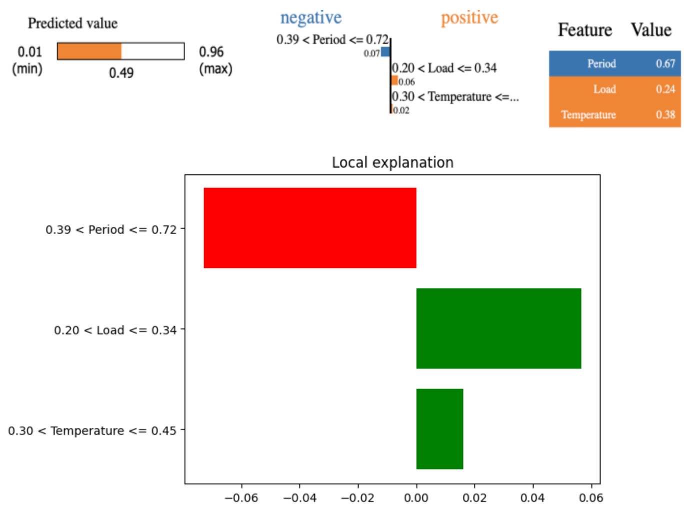

Figure 4. LIME interpretation of LSTM-ENS prediction scenario

https://doi.org/10.25077/jnte.v15n2.1605.2026

CONCLUSIONS

This study presented a novel technique of predicting the transformer BDV using in-situ physical quantities that have strong correlation with transformer oil ageing thereby eliminating the need for laboratory analyses using deep-learning method. Transformer oil breakdown voltage (BDV) was predicted using three physical factors viz. temperature, loading, and service period. Four different machine learning models and an LSTM-ensemble machine learning model were developed and used to predict the BDV. The study found strong correlations between these physical quantities and BDV, making them suitable for BDV prediction. During the testing phase models developed has the following accuracy, Multiple Linear Regression (MLR): 75.4%, Support Vector Machine (SVM): 80.7%, Multilayer Perceptron (MLP): 87.3% and Long Short-Term Memory (LSTM): 92.2%. The LSTM-ensemble model, which combined predictions from individual LSTM models, achieved 95.7% accuracy, demonstrating improved prediction accuracy and robustness by leveraging the diversity of individual models. This ensemble approach increased model performance by 20.3% in the training phase and 10.3% in the testing phase compared to the MLR model, indicating a promising method for BDV predictions. The use of these quantities eliminates need of data from laboratory analyses of the oil for prediction of the BDV which hitherto was the practice; thus, saving time. The application of machine learning models, particularly the LSTM-ensemble approach, presents an effective and non-destructive method for predicting transformer oil breakdown voltage (BDV). The LIME interpretation provided actionable insights into the impact of operational factors, making this approach highly valuable for real-time transformer maintenance and reducing reliance on costly laboratory tests. This combination of predictive accuracy and interpretability holds significant potential for improving the reliability and efficiency of transformer operations in power systems. For future work, larger data sets and deep machine learning models may improve the accuracy and reliability of the proposed models.

REFERENCES

[1] H. Alhelou, M. Hamedani-Golshan, T. Njenda, and P. Siano, "A survey on power system blackout and cascading events: Research motivations and challenges. Energies, 12 (4), 682," 2019.

[2] N. Bashir, H. Ahmad, and M. S. Suddin, “Ageing studies on transmission line glass insulators using dielectric dissipation factor test,” 2010 9th International Power Energy Conference (IPEC), Singapore, 27 – 29 October 2010: IEEE, pp. 1062–1066.

[3] H. L. Willis, Electrical Power Equipment Maintenance and Testing, 2nd ed. London: CRC Press, 2009.

[4] X. Wang, C. Tang, B. Huang, J. Hao, and G. Chen, "Review of research progress on the electrical properties and modification of mineral insulating oils used in power transformers," Energies, vol. 11, no. 3, p. 487, 2018.

[5] I. S. Chairul, S. A. Ghani, N. A. Bakar, N. H. Rahim, and N. H. Ab Aziz, "Effect of acids formation on the breakdown voltage properties of transformer insulating oil," in 1st International Symposium on Engineering and Technology (ISETech), Perlis, Malaysia, 23 December 2019, vol. 767, no. 1: IOP Publishing, pp. 1-6.

[6] Y. Chen, X. Wang, J. Shi, J. Zhao, and J. Zhang, "Transformer oil breakdown voltage prediction using an artificial neural network model with optimized input variables," Applied Sciences, vol. 11, no. 2552, pp. 346-372, 2021.

[7] J. Liu, Y. Zhang, and J. Ma, "Power transformer fault detection and diagnosis using gas-in-oil analysis and machine learning algorithms," Journal of Electrical Engineering and Technology, vol. 12, no. 12, pp. 2245-2253, 2017.

[8] O. E. Gouda and A. Z. El Dein, "Prediction of aged transformer oil and paper insulation," Electric Power Components and Systems, vol. 47, no. 4-5, pp. 406-419, 2019.

[9] H. Nazar, A. Rashid, M. Shahzad, F. Khan, and B. Khan, "Optimization of the factors affecting the breakdown voltage of the transformer oil using genetic algorithm," in 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 3-4 March 2018: IEEE, pp. 1-6.

[10] S. V. Kulkani, Transformer Engineering: Design, Technology, and Diagnostics, 2nd ed. London: CRC Press, 2013.

[11] J. Sun and H. H. Hu, "An improved artificial neural network model for predicting breakdown voltage of transformer oil," Measurement, vol. 148, pp. 298-305, 2020.

[12] M. Senoussaoui, M. Brahami, and I. Fofana, "Transformer Oil Quality Assessment Using Random Forest with Feature Engineering. Energies, 14, ," Energies, vol. 14, p. 1809, 2021.

[13] A. Christina, M. A. Salam, Q. Rahman, F. Wen, S. Ang, and W. Voon, "Causes of transformer failures and diagnostic methods–A review," Renewable and Sustainable Energy Reviews, vol. 82, pp. 1442-1456, 2018.

[14] R. Murugan and R. Ramasamy, "Understanding the power transformer component failures for health index-based maintenance planning in electric utilities," Engineering Failure Analysis, vol. 96, pp. 274-288, 2019.

[15] B. W. Tuinema, J. L. Rueda-Torres, A. I. Stefanov, F. M. Gonzalez-Longatt, and M. A. M. M. V. D. Meihden, Probabilistic Reliability Analysis of Power Systems: A Student’s Introduction. Switzerland: Springer, 2020.

[16] Y. Su, J. Zhang, Y. He, Y. Shu, M. Liu, and Q. Zhou, "Evaluating BDV in Transformer Oil Combined MFU and GRNN," in 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Beijing, China, 6-10 September 2020: IEEE, pp. 1-4.

[17] S. S. Ghoneim, S. S. Dessouky, A. A. Elfaraskoury, and A. B. A. Sharaf, "Prediction of insulating transformer oils breakdown voltage considering barrier effect based on artificial neural networks," Electrical Engineering, vol. 100, no. 4, pp. 2231-2242, 2018.

[18] S. A. Khan, M. Tariq, M. Farhan, and Z. A. Khan, "Statistical techniques for prediction of breakdown voltage of liquid dielectrics used in power transformers," in 2023 International Conference on Power, Instrumentation, Energy and Control (PIECON), Aligarh, India, 10-12 Feb 2023: IEEE, pp. 1-5.

[19] S. Zandbaaf, M. R. K. Khorrami, and M. G. Afshar, "Genetic algorithm based artificial neural network and partial least squares regression methods to predict of breakdown voltage for transformer oils samples in power industry using ATR-FTIR spectroscopy," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 273, p. 120999, 2022.

[20] C. Y. Yang and D. Wu, "A hybrid model combining random forest algorithm and gradient boosting decision tree algorithm for transformer oil breakdown voltage prediction," Journal of Energy Engineering, vol. 147, pp. 102-112, 2021.

[21] H. Chen and R. Liao, "Transformer oil breakdown voltage prediction based on improved adaptive particle swarm optimization algorithm and support vector machine," Journal of Intelligent and Fuzzy Systems, vol. 41, pp. 2023-2033, 2021.

[22] I. Ahmad and U. Maqsood, "Prediction of transformer oil breakdown voltage using multiple linear regression analysis," in 3rd International Conference on Computer, Communication and Signal Processing, Sanya, China, 20-22 December 2020.

[23] M. A. Wahab, M. Hamada, A. Zeitoun, and G. Ismail, "Novel modeling for the prediction of aged transformer oil characteristics," Electric Power Systems Research, vol. 51, no. 1, pp. 61-70, 1999.

[24] Y. Xiong, "Prediction Model for Breakdown Voltage of Transformer Oil Based on Relative Transformation and Kernel Principal Component Analysis," Journal of System Simulation, vol. 30, no. 5, pp. 1657-1664, 2019.

[25] M. A. A. Wahab, "Artificial neural network-based prediction technique for transformer oil breakdown voltage," Electric Power Systems Research, vol. 71, no. 1, pp. 73-84, 2004.

[26] F. Guerbas et al., "Neural networks and particle swarm for transformer oil diagnosis by dissolved gas analysis," Scientific Reports, vol. 14, no. 1, p. 9271, 2024.

[27] A. O. Osofisan, 1st, Ed. Fundamental concept of Machine learning in Artificial Intelligence with Practical Examples. Abuja: TETFUND, 2022.

[28] J. K. Solano Meza, D. Orjuela Yepes, J. Rodrigo-Ilarri, and M.-E. Rodrigo-Clavero, "Comparative analysis of the implementation of support vector machines and long short-term memory artificial neural networks in municipal solid waste management models in megacities," International Journal of Environmental Research and Public Health, vol. 20, no. 5, p. 4256, 2023.

[29] L. Li and T. Zhang, "Research on Text Generation Based on LSTM " International Core Journal of Engineering, vol. 7, no. 525-535, 2021.

[30] M. J. Jimenez-Navarro, M. Martinez-Ballesteros, F. Martínez-Álvarez, and G. Asencio-Cortés, "A new deep learning architecture with inductive bias balance for transformer oil temperature forecasting," Journal of Big Data, vol. 10, no. 1, p. 80, 2023.

[31] L. Cheng and T. Yu, "A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems," International Journal of Energy Research, vol. 43, no. 6, pp. 1928-1973, 2019.

[32] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, "Linear Regression," in An Introduction to Statistical Learning: with Applications in Python. Cham: Springer International Publishing, 2023, pp. 69-134.

[33] Y. Joo, D. Kim, Y. Noh, J. Choi, and J. Lee, "Performance Comparison of LSTM and ESN Models in Time-Series Prediction of Solar Power Generation," Sustainability, vol. 17, no. 19, p. 8538, 2025, doi: https://doi.org/10.3390/su17198538.

[34] Y. Joo, D. Kim, Y. Noh, J. Choi, and J. Lee, "Performance Comparison of LSTM and ESN Models in Time-Series Prediction of Solar Power Generation," Sustainability, vol. 17, no. 19, p. 8538, 2025, doi: https://doi.org/10.3390/su17198538.

AUTHORS BIOGRAPHY

Musbahu Garba Indabawa was born in Kano, Nigeria. He obtained his BEng (Hons) from Aliko Dangote University of science and technology, Kano, Nigeria; MEng in Electrical Engineering in 2024. Currently, he is a technical engineer all in Kano electricity distribution company (KEDCO). His research interests include insulation condition monitoring of power system equipment, new insulating materials and volt-var control (VVC). He is a registered Engineer with council for regulation of engineering in Nigeria (COREN) and member Nigerian society of engineers (NSE), also Member of the Nigerian Institute of Electrical and Electronics Engineers (NIEEE).

Nouruddeen Bashir was born in Kano, Nigeria. He obtained his BEng (Hons) from Abubakar Tafawa Balewa University, Bauchi, Nigeria; MEng in Electrical Power Engineering and PhD in Electrical Engineering with specialisation in the area of High Voltage Engineering from Universiti Teknologi Malaysia in 2006 and 2009 respectively. Currently, he is a Professor of High Voltage Engineering, Faculty of Electrical Engineering, Kano University of Science and Technology. His research interests include high voltage equipment condition monitoring and diagnosis, high voltage insulation materials, renewable energy and energy efficiency. He is a Chartered Engineer with the Engineering Council UK and a Fellow of Institution of Engineering and Technology (IET), UK.

Rabiu Aliyu Abdulkadir was born in Kano, Nigeria. He obtained his BEng in Electrical Engineering from Kano University of Science and Technology, Wudil, Nigeria. He earned an MTech in Instrumentation and Control, from Sharda University, India, and a PhD in Electrical and Electronics Engineering with a specialization in Intelligent Control Systems, from Near East University, Northern Cyprus, Turkey, in 2020. Currently, he is an Associate Professor in the Department of Electrical Engineering of Kano University of Science and Technology). His research interests include cyber-physical systems, the Internet of Things (IoT), and artificial intelligence (AI). He is a Member of the Institute of Electrical and Electronics Engineers (IEEE) and a Registered Engineer with the Council for the Regulation of Engineering in Nigeria (COREN).

https://doi.org/10.25077/jnte.v15n2.1605.2026