University of Ibadan
Journal of Science and Logics in ICT Research

Fault Prediction in Power Transformer Using Ensemble Models

[1]Igodan, C. E., [2]Katyo, P.

[1]Department of Computer Science, Faculty of Physical Sciences, University of Benin, Edo State
[2]Risk Management Unit, ICT - Modernization Department, Nigeria Customs Service Abuja
charles.igodan@uniben.edu, katyoabu@customs.gov.ng

Abstract

One of the highly important elements of electrical system networks is the power transformer. There is an increasing amount of research being done on early warning systems and faults detection because the failure of these elements can ground economic activities. More so, using dissolved gas analysis (DGA) as one of the mostly used conventional techniques is deficient in locating these incipient faults as this may be caused by a variety of factors which includes but not limited to imbalance problem, inadequate and overlap in the DGA datasets, thereby restricts its capacity to obtain precise diagnosis. Therefore, this paper proposed an ensemble machine learning methods for incipient faults prediction using DGA datasets comprising 166 samples and eight variables. This research compares the accuracies of four ensemble machine learning methods: Bagging, Adaboost, Stacking, and Voting methods using multilayer perceptron and support vector machines respectively. The results obtained ranges from 90.50% to 100% with the Adaboost (MLP) achieving the highest accuracy, whilst the misclassification percentage ranges from 1.62% - 18.06% with Stacking method as the least performing. In the end, our findings highlighted the importance of the use of ensemble methods and has future prospects for further advancement.

Keywords: *Ensemble method, Fault in power transformer, Dissolved gas analysis, Supervised algorithms, Machine learning algorithms*

1. Introduction

Electrical transformers are inert device that uses electromagnetic induction to help move electrical energy between circuits. It is a costly and crucial part of the production, transmission, and distribution of electricity [1]. Power transformers undergo considerable mechanical, thermal, and electrical strain during operations and can eventually malfunction if not early detected. Failures of the power transformer not only stop the energy from flowing continuously, but also put the security and stability of the entire power system in jeopardy. Furthermore, these damages may result in significant losses for society and the economy at large [2]. A non-invasive method used for finding these transformers' faults, known as Dissolved gas analysis (DGA), are used to detect the gases emitted [1, 34, 35] in certain measurable quantity. Certain gases leak out in detectable amounts as a result of the insulation materials degrading due to heat produced by the transformer when in operation. Some of the major gases most frequently used for analysis, but not limited to, are Methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), hydrogen (H_2), acetylene (C_2H_2), carbon monoxide (CO), and carbon dioxide (CO_2) and are expressed and measured in parts per million (ppm) [3, 37]. The Dornenburg approach, Rogers' method, key gas method, Duval Triangle method, and IEC ratios are well-known conventional techniques for analyzing DGA data [4], but are unreliable due to their variability of results.

Furthermore, these techniques mostly depend on the technical expertise of human specialists, and they also run the risk of misidentifying the kind or degree of problems, time consuming and produces variable results [1, 36]. As a result, a variety of researches leveraging artificial intelligent (AI) techniques have introduced different modelling frameworks to effectively deal with some of these challenges [38]. Some of the intelligent machine learning

Igodan C. E. and Katyo P. (2024). Fault Prediction in power transformer Using Ensemble Models. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 11 No. 1, pp. 40-53

algorithms includes Multiplayer perceptron neural networks (MLPNN), support vector machines (SVM), fuzzy theory, Adaptive Fuzzy-Neural Inference System, and the combination of other machine learning (ML) algorithms [5], e.g., hybrid, ensemble and cascaded models. The effectiveness of these algorithms have been the subject of numerous investigations which has led to the continued development for new effective intelligent techniques for better maintenance of the power distribution systems and are still ongoing [38]. However, these models are not robust nor stable due to lack of diversity.

Generally, machine learning algorithms are categorized into supervised, unsupervised and reinforcement used mainly to resolve complex classification and regression problems due to their automated and intelligent decision-making capabilities. However, the challenges faced are not limited to the following: increased complexities, increased error rates and false positives, inability to deal with big data problems and large feature sizes which may reduce performance. To the best of the authors' knowledge, only very few articles applied the ensemble machine learning methods – a method of combining different ML algorithms – to address these problems and create diversity capable of yielding robust and scalable model for increased predictive accuracy as evident in the literature [61]. Motivated by these limitations, this paper proposed an intelligent ensemble of machine learning methods for power transformer incipient faults prediction using SVM and MLP as suggested by [49] and [40].

2. Related Works

The important role played by transformer in the power industries has motivated more researchers into conducting studies geared towards the development of different artificial intelligence and machine learning techniques for DGA interpretation. This section presents related works.

In Tran *et al.*, [38] a literature review of the health status of distribution transformers with regards to current methodologies was conducted. Tran *et al.*, posited that in recent years, AI techniques have shown to be indispensable in developing powerful models towards making real-time monitoring next-generation system equipped with high levels of reliability, intelligence, sensitivity, and cheapness. Zhang *et al.*, [56] conducted a systematic review on the application of artificial intelligence techniques for DGA-based diagnostics for solving intractable problems in early transformer fault diagnosis with regard to the use of neural, clustering, support vector machine, etc. Zhang *et al.*, observed that the DGA data does not fully reflect the transformer status, and needs to be combined with new monitoring data for more effective fault diagnostics.

Li [57] did a comprehensive review of the current state of the art in transformer fault diagnosis (TFD) providing valuable insights into the importance of TFD and the role of AI in ensuring the reliability operation of power systems. The study suggested that to improve the accuracy and reliability of transformer fault diagnosis, effective maintenance and reduce downtime, the integration of different diagnostic tools and techniques, including AI and DGA holds good prospect in the future.

Esteves *et al.*, [40] presented a comprehensive review pertaining to the usage of ML systems for fault data analytics to ascertain the most relevant methods. Their study shows that the concatenation of two distinct fields of knowledge (e.g. AI) can provide useful insights for a wide range of scientific experts in any field. The authors study served as a guideline to expand the usability of ML for power transformers and motivation for further AI applications.

Zhang *et al.*, [42] did a comprehensive review of the development of AI technology represented by expert systems, ML, uncertainty, reasoning, intelligent optimization techniques, etc. In the study, Zhang *et al.*, posited that in the context of the era of big data in electric power transformer, data mining, and analysis of power transformer state, data can provide operation and maintenance management of equipment for important decision support. However, the authors suggested focusing on strengthening the management and maintenance of status data, and exploration of effective new monitoring methods so as to further promote the intelligent development of power equipment status maintenance. Furthermore, a multinomial DGA classifier for incipient fault detection in oil-impregnated power transformers was proposed in Odong *et al.,* [1]. Apte *et al.*, [13] combined Roger's ratio and IEC ratio, and fuzzy inference system (FIS) in their work. While FIS system showed improved efficiency for diagnosing power transformer faults, the models are not adaptive. Whereas the application of KosaNET, an ensemble model based on decision trees used for the classification of the multinomial data obtained was 99.98% accuracy.

In Bouchaoui *et al.,* [6], the traditional IEC and Roger's ration techniques combined with machine learning algorithms - neural network increased efficiency from 20% to 70% and 40% to 70% for the IEC and Roger's ratios method respectively. However, the system was prone to produce misleading results. The Parzen window estimation technique was applied in Islam *et al.,* [4] to increase performance and compared with the traditional methods. Their approach obtained a 95% accuracy. However, a small sample size was used.

Wagh and Deshpande [7], used a back-propagation radial basis function and an adaptive ANFIS and obtained a 98.85% accuracy. However, there was no validation and the ANFIS is slow. The study of Muthi *et al.,* [8] adopted the multilayer perceptron (MLP) model, and compared the results with the Doernenburg ratio and Rogers ratio. The results showed that MLP performed far better than others even without tuning the model's parameters. Ghoneim and Taha, [11] developed a new approach for DGA technique based on the gas concentrations to obtain higher agreement accuracy than traditional DGA techniques, and obtained an overall accuracy of 84.70%.

Aburaghiega *et al.,* [12] applied fuzzy logic with an accuracy rate of 99%, however, their limitation is that there was no learning and adaptation. Faiz and Soleimani [15], introduced ANN with fuzzy systems using Duval pentagon method and fuzzy inference system in obtaining improved performance. However, the model lacks explain-ability, interpret-ability whilst their fuzzy model lacks learning and adaptability. Rokani and Kaminaris, [5] applied a fuzzy inference system (FIS), artificial neural network (ANN), and adaptive ANFIS among others.

Different ML techniques investigated in Rajora *et al.*, [39] were used for the effective asset management in power distribution systems. Although, the ML algorithms examined, shared the requirement of needing large input datasets for training these models in obtaining good results, however, they have shown in the study to be potentially significant for optimizing ML processes especially with large dataset which is a disruptive evolving area of research in recent years offering a promising performance in upcoming years.

The works of Senoussaoui *et al.,* [41] used J48 decision tree and random forest (RF) to develop a new, simple, and effective ML approach to monitor the condition of transformer oils based on some aging indicators. The authors used k-means to transform the dataset, afterwards principle component analysis (PCA) was applied before correlation-based feature selection algorithm for filtering purpose, and then modeled with RF. Although, RF performed better than J48 classifier, the work was offline and characterized by poor dataset. Valencia et al., [43] compared four different ML techniques (SVM, ANN, LR, and RF) to help determine electrical faults of a mechanical stressed three-phased power transformers' winding conductors. Their study showed that RF obtained the most accurate results. However, the errors were high as it was due to low values of stress mainly as a result of the big difference between the maximum and minimum stresses limiting the model training.

Alyunov and Vyatkina [44], presented an ANN model for identifying the parameters of mathematical models of power transformers and obtained improved accuracy. Chen *et al.*, [45] proposed a transformer oil temperature prediction model based on empirical mode decomposition-bidirectional long short-term memory (EMB-BiLSTM). The EMB-BiLSTM model outperformed the traditional LSTM, BiLSTM, EMD-BP, and Wavelet Transform Bidirectional Long Short-Term Memory (WT-BiLSTM) methods, demonstrating an effective and accurate predictive model. However, the transformer's functioning condition was not included in their predictive study. Xing *et al.,* [46] observed low data quality, binary classification effect, and small sample learning as critical limitation to fault prediction or power transformers. The study proposed a predictive model based on DGA chromatography data using Mish-SN Temporal Convolutional Network (MSTCN) obtaining better results and providing new insights for fault prediction. However, the study was characterized by longer training time and model complexity problems.

Asad and Fareeh [47], also observed that common single models used for the predictive maintenance of transformer include DGA, ANN, SVM, and multi class least square support vector machine. The proposed work is based on linear regression and principle component analysis (PCA) using two different approaches: phase A voltage's magnitude and phase B current's magnitude. The results obtained showed improvement, however, more advanced techniques were suggested so as to further improve accuracy. Ofori *et al.*, [50] proposed a predictive ML model for the health monitoring of power transformer using SVM, K nearest neighbour (kNN), Logistic Regression (LR) and RF. The RF obtained a 94.4% accuracy. However, the work observed that the life span of a power transformer is affected by the hot-spot temperature and suggested the use of other advanced ML tools for improvement.

Aizpurua *et al.,* [48] observed that the existing analytic models for Hot-Spot Temperature (HST), which determine the remaining useful life (RUL) of transformer insulation paper, calculation are not always accurate due to the inability to generalize the properties of transformers operating in different contexts. The authors presented a novel transformer condition assessment approach integrating uncertainty modelling, data-driven forecasting models and model-based experimental models to increase the prediction accuracy and handle uncertainty. The forecasting results showed that extreme gradient boosting (XGB) algorithm best captures the non-linearities of the thermal model and improves the accuracy among other approaches. Mateus *et al.,* [49] observed that it is possible to predict the health status or fault types in transformers by analyzing oil sample. The study combined Fuzzy Logic (FL) and Neural Network techniques in making informed decision in the early detection of failures in transformers. Mateus *et al.,* obtained a 95% accuracy and highlights the importance of the predictive maintenance model. However, the authors suggested the exploration of other classification models and the introduction of larger datasets to build confidence in the results obtained. Bjelic et al., [51] assessed the health of power transformer based on sweep frequency response analysis using experimental data obtained from power transformers with interval short-circuit faults.

The work classified and divided the examined power transformer state into groups with similar state and probability of failure using k-means cluster methods and applied ANN and ANFIS to detect fault severity of power transformers of different lifetime. The ANFIS and ANN results obtained were better than k-means and cluster methods when compared. However, the authors suggested focusing on testing new algorithms on a greater number of power transformers of different ages and conditions towards achieving predictive maintenance based on automatic and timely decision-making model for the proper time-based maintenance planning system (TBM).

Balasubramanian *et al.,* [52] observed that despite the predictive maintenance practices through the implementation of various condition-based maintenance activities in monitoring and maintaining power transformers, some transformer defects are still left undetected especially at the early stages. The study developed a predictive system as part of the condition-based maintenance to prioritize transformers that are undergoing more severe deterioration before permanent irreversible damage occurs. Five ML algorithms were evaluated using LR, DT, RF, adaptive Boosting, and XGB respectively. RF obtained and maintained the highest accuracy performance.

Wang et al., [53] observed that traditional operation and maintenance tools lack effective predictive capabilities for potential faults; and the scarcity of existence fault data makes it difficult to apply ML techniques effectively too. As a result, their study proposed a novel approach that leverages the knowledge graph

(KG) technology in combination with gradient boosting trees (GBDT) to efficiently learn from a small set of high-dimensional data, integrating various factors influencing transformer faults and historical operational data. Experimental results showed that their proposed method outperformed other learning approaches in predictive accuracy, such as ANN and LR, offering significant improvements in progressiveness, practicality, and potential for widespread application. However, their proposal lacks generalization ability.

Aning *et al.*, [54] presented a predictive distribution transformer failure detection model before its expected years in service. Aning *et al.,* discovered that the most significant factors that determine the number of years left for a distribution transformer to fail are rate-of-faulty-occurrence, type-of-faults-sustained and Tap-changer-type. The proposed model obtained an RMSE and MAPE of 0.001639 and 0.001321 respectively. However, the study was limited with data scarcity which affected the prediction accuracy of the study. Suwarno et al., [55] observed different conclusions for the same oil sample characterizing the various DGA interpretation techniques: Doernenburg Ratio (DR), Roger Ratio (RR), IEC Ratio (IR), Duval Triangle (DT), and Duval Pentagon Methods (DPM) respectively, and results to out-of-code condition if any of the used gases fall outside the specified limits.

Suwarno et al., proposed a multi-method based on the scoring index and random forest machine learning principles and achieved an average of 96%, 93.4% accuracy and consistency respectively surpassing the conventional methods (DR, RR, IT, DT, and DP). However, the proposed method was based on limited dataset, lacking generalization ability. Moni and Gouri [58], proposed mathematical methods comprising of ML and DL models to predict the degradation of transformer oil. However, their study was limited by the insufficient amount of dataset. Tata and Mansour [59], proposed a novel optimized ML method for power transformer faulty diagnosis. The study employs the DT, discriminant analysis, naïve Bayes, SVMs kNN, and ensemble classification methods with four data transformation techniques which comprised of logarithmic, normalization, standardization, and gas percentage transformations. The proposed model shows superiority over conventional methods with the SVM and ensemble obtaining 90.61% using the gas percentage transformation better than other techniques.

Ngwenyama and Gitau [60], identified incipient faults in oil-immersed transformers (OITs) using ML algorithms and DGA data. The study showed enhanced diagnostic trust with an accuracy of 87.7%, 86.2% and 84.1% for Bagged Trees, Fine kNN, and Quadratic SVM, respectively.

Despite the several ML techniques used in the literature, there has been a few application of ensemble learning algorithms applied in power transformer faults detection [59, 60]. Motivated by these drawbacks the various ensemble learning methods were investigated towards building confidence in accuracy and reliability in fault predictive ensemble models using SVM and MLP as one of the most commonly used classifiers.

3. Methodology

This section outlines the general methodology of our suggested approach as well as the primary research component. Figure 1 depicts the suggested system's architecture.

3.1 Data Discription

The IEC-TC 10 dataset source was obtained from [17, 20] which consists of 166 instances and 8 features. Seven features are independent variables while the last attribute represents the observed fault. From the dataset, six classes were identified as follows: normal state, partial discharge (PD), discharge of low energy (D1), high energy discharge (D2), thermal fault below 300°C and Thermal Fault above 300°C but below 700°C (T1&T2) and thermal Fault above 700°C (T3) [17,18]. The datasets were split into 60% training 40% testing respectively

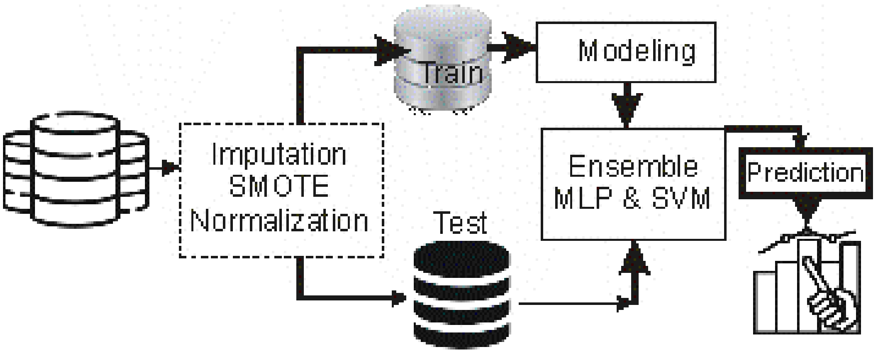


Figure 1: System architecture of the proposed methodology

3.2 Preprocessing the data

The preprocessing of the data was carried in this section to prevent bias, and ensure high-quality data is applied to machine learning. First, the median imputation technique is used to fill-up missing values using Equation 1. Second, the synthetic minority oversampling technique (SMOTE) was applied using Equation 2. After applying SMOTE, our dataset was increased to 300 having 60 samples for each fault types. Finally, the obtained dataset was normalized using Equation 3.

$$d(x, y) = \sqrt{(\sum_{i=1}^{N}(x_i - y_i)^2)} \quad (1)$$

$$x_{syn} = x_i + (x_{knn} - x_i) * t \quad (2)$$

$$X' = \frac{x - min(x)}{max(x) - min(x)} \quad (3)$$

where X' is the normalized value of the original data point $x, min(x)$ is the lowest value of the whole dataset, $max(x)$ is the highest value of the whole dataset .

3.3 Modeling

Figure 2 depicts the machine learning framework used in this paper. The processing is an iterative step leading up to modeling and evaluation. Four different ensemble methods: AdaBoost [23], Boosting [29], Stacking [28], and Soft voting [30], were adopted in this study using two classification algorithms: Multilayer Perceptron (MLP), and Support Vector Machines (SVM) [19, 6]. Bagging and Boosting uses different models to reduce error and optimizes the model. Whilst bagging techniques combines multiple models trained on either same or different subsets of data, the boosting trains the model sequentially, focusing on the error made by previous model. Stacking supplies the predictions of the different base models as input to a meta-model. Logistic regression was adopted as the meta learner. Equations 4 to 7 represents the AdaBoost, Boosting, Stacking and (simple) Soft Voting ensemble methods, whilst the MLP and SVM are depicted in Equations 8, 9, and 10 respectively. The "one-vs-rest" method of multi-class SVM classification was used in this study [23, 24, 28, 33].

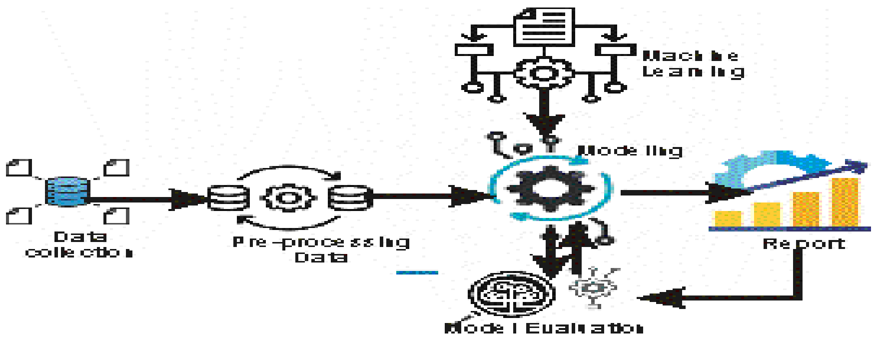


Figure 2: A typical machine learning framework

$$H(x) = Sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \quad (4)$$
$$H(x) = arg\ max_{y \in Y} \sum_{t=1}^{T} \mathbb{I}(h_t(x) = y) \quad (5)$$
$$H(x) = h'(h_i(x), \ldots, h_T(x)) \quad (6)$$
$$H^j(x) = \frac{1}{T} \sum_{i=1}^{T} h_i^j(x) \quad (7)$$

Where h_t is classifier, T is number of learning rounds, x is instance variables, α_t is the weight, $H(\cdot)$is the set of hypothesis, y is class value, $\mathbb{I}(.)$ is the indicator function which takes 1 if dot $(\cdot)$ is true, and 0 otherwise, $Sign(\cdot)$ is the sign function which tales -1,1, and 0 when dot(.)<0, dot(.)>0 and dot $(\cdot)$ = 0, respectively. $h_i^j(x) \in [0,1]$ where individual classifier h_i outputs a 1-dimensional vector

$$(h_i^1(x), \ldots, h_i^l(x))^T.$$
$$y_i[W^T X_i + b] \geq 1 - \xi_i,$$
$$I=1,\ldots,l,\ \xi_i \geq 0 \quad (8)$$
$$g(y_k(x)) = \begin{cases} 1 & y = c \\ 0 & y \neq c \end{cases} \quad (9)$$
$$y_j = f\left(\sum_{i=1}^{n} X_i W_{ij} + \theta_j\right) \quad (10)$$

where X_i are the network inputs; W_{ij} translate the weight-connection between the input neuron i and the neighbouring hidden neuron j; θ_j is the bias of the jth hidden neuron, y_j is the output of the network, and $f(.)$ is the transfer function or also called activation function.

4. Performance Evaluation Methods

This study applied confusion matric (CM) for performance evaluation of our proposed models. Equations 11 to 14 depicts the accuracy, precision, recall, and F-measure respectively. The Receiver Operating Characteristic (ROC) curve was also applied [1,19,31]. The measures are binary classification systems used to classify examples into positive or negative cases on the 1 vs All approach.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (11)$$
$$Precision = \frac{TP}{TP + FP} \quad (12)$$
$$Recall = \frac{TP}{TP + FN} \quad (13)$$
$$F - measure = 2 * \frac{precision \times recall}{precision + recall} \quad (14)$$

5. Results and Discussions

The results from utilizing machine algorithms and ensemble model for classification of power transformer faults are discussed in this section. All experiments were conducted on an 8th-generation Core i5 machine running a Linux operating system. We deployed Jupyter Notebook with the Python programming language to implement our machine learning models. These models' classification performance was experimented using default parameter values in the Jupyter Notebook API environment, and measured.

5.1 Results

The results obtained from the analysis of the study are presented in Tables 1 and 2 respectively. Table 1 shows the results of the different ensemble models adopted in this study with respect to accuracy, precision, recall and F1-scores. The comparative analysis between the proposed model and existing studies are illustrated in Table 2. Whilst Figures 3 and 4 captures the graphical representation of the test results in the tables.

5.2 Discussion

The performance of Adaboost (MLP) showed improved accuracy of 100% as the highest achieving this feat after several runs. This high performance was obtained as a result of the ability of ensemble method to induce high variability among diverse base classifiers and the assumption that the diversity among hypothesis improves overall performance compared to base models. The Adaboost (MLP) was boosted for up to 50 iterations to achieving the high performance. The stacking and voting ensemble models obtained 97.52% and 97.42% accuracy respectively.

Whilst, bagged (SVM), bagged (MLP), and Adaboost (SVM) obtained 90.5%, 90.5%, and 97.5% respectively. Also, while bagging creates an ensemble of training individual classifiers on a bootstrapped sample from the training set generated by random selection with replacement of instances of the training set; in boosting, sampling is proportional to the weight of an instance. The results obtained by the ensemble models shows that diversity was injected homogeneously therefore building stability and confidence of the results as posited in [62, 63, 64].

Our results prove that bagging and boosting ensemble learning methods guarantees good experimental results due to the strong theoretical background. Comparing our proposed model and existing model, ours shows to be the highest performing model and closely followed by the works in [1,12,13] respectively. Other metrics used like the precision, recall, and F1-score values obtained are indicative of the stability and confidence of the ensemble models as shown in the study. While the highest values were 98.05%, 98.88%, 98.89%, 96.45%, 89.90%, and 89.80% for Adaboost (MLP), stacking, Adaboost (SVM), voting, bagged (SVM), and bagged (MLP) follows sequentially.

The class-wise distribution of the predictive performance of the classification models on the test dataset is shown in Figures 5 to 10 depicting their confusion matrix respectively. Figure 5 depicts Bagged (MLP) with an average accuracy of 98.06% and 1.94% misclassification error as the fourth performing model and its ROCAUC curve graphically depicted in Figure 14 respectively. The Bagged (SVM) represented in Figure 6 achieved 92.22% and 7.78% average accuracy and error respectively and its ROCAUC curve shown in Figure 13. Adaboost (MLP) shown in Figure 7 obtained an average accuracy of 100% with no misclassification as the best ensemble model and its ROCAUC graphically depicted in Figure 11.

Figure 8 shows an average accuracy of 98.33 with a 1.67% error rate for bagged (SVM), and its ROCAUC curve shown in Figure 6. The Stacking model shown in Figure 9 achieved an average accuracy of 81.94% with an average error of 18.06% respectively, and the ROCAUC is illustrated in Figure 15. Lastly, the average accuracy and misclassification for the voting ensemble is depicted in Figure 10 as 97.77% and 2.23% while the ROCAUC captured in Figure 16 respectively

Table 1: Evaluation comparison of the models implemented

Model	Accuracy	Precision	Recall	F1_Score
Bag (SVM)	90.50	89.90	85.50	85.70
Bag (MLP)	90.50	89.80	85.54	85.71
Adaboost (SVM)	97.50	96.45	95.40	95.50
Adaboost (MLP)	100	98.05	98.88	98.89
Stacking	97.52	97.21	96.52	96.62
Voting	97.42	96.94	96.10	96.11

Table 2: Comparative Analysis

Models	Proposed Model Adaboost (MLP)	[1]	[5]	[7]	[9]	[12]	[13]
Accuracy (%)	100	99.98	97.5	98.85	85.71	99	99.98

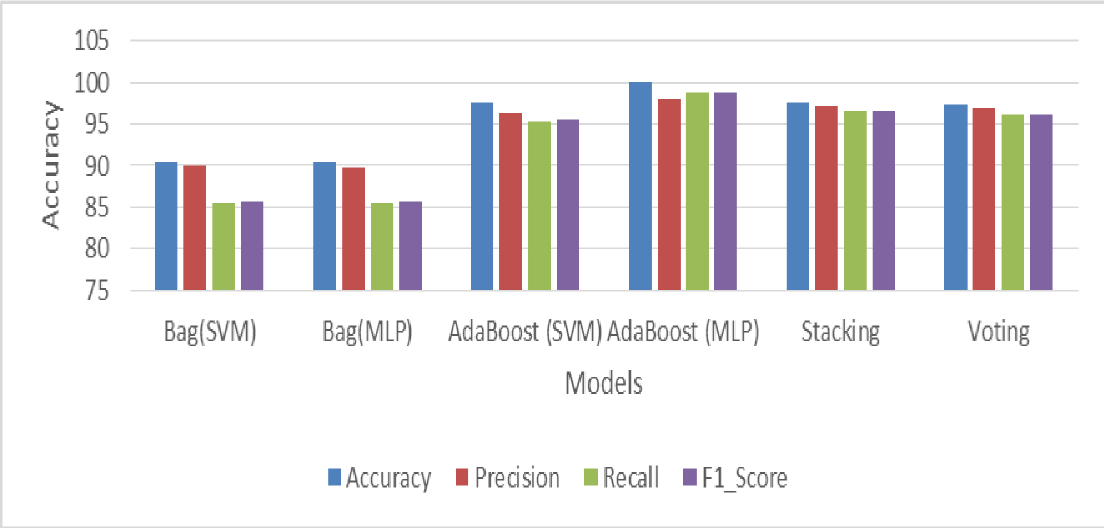
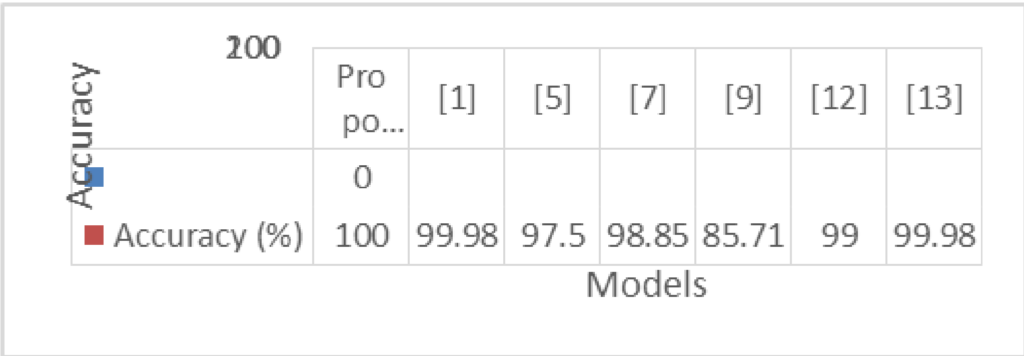
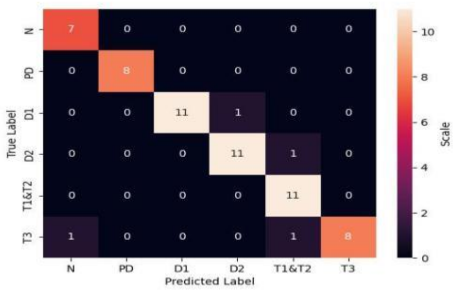


Figure 3: Performance metrics of Proposed models



	Pro po...	[1]	[5]	[7]	[9]	[12]	[13]
Accuracy (%)	100	99.98	97.5	98.85	85.71	99	99.98

Figure 4: Comparison of Proposed and existing models



5: Bagging (MLP)

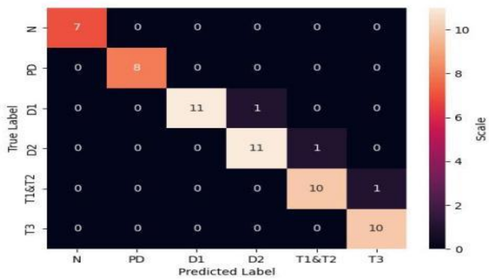


Figure 6: Bagging (SVM)

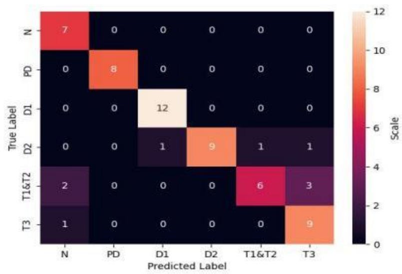


Figure 7: Adaboost (MLP)

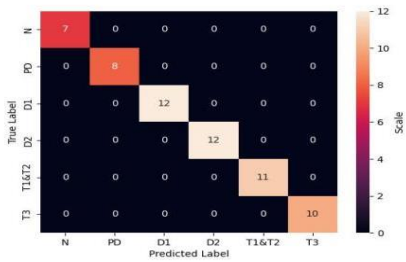


Figure 8: Adaboost (SVM)

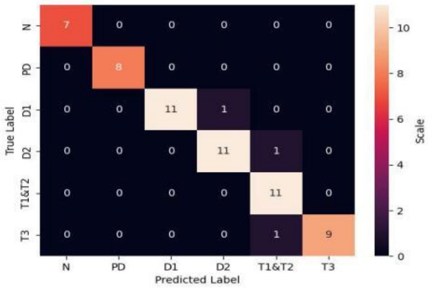


Figure 9: Stacking of MLP & SVM

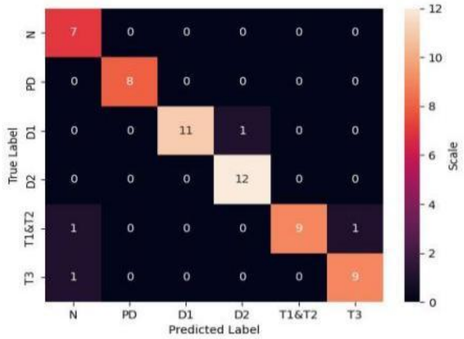


Figure 10: Voting Ensemble

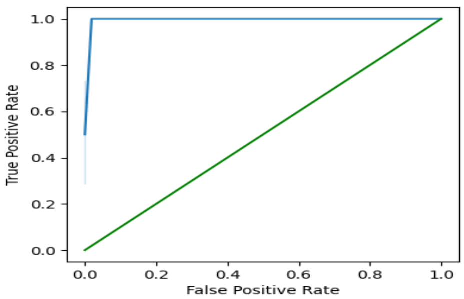


Figure 11: Adaboost (SVM) ROC curve

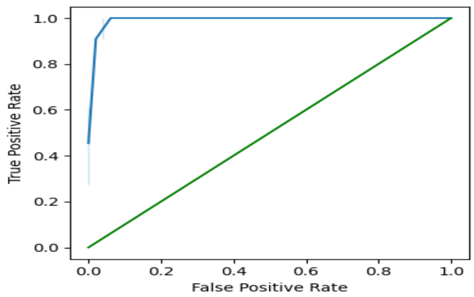


Figure 12: Adaboost (MLP) ROC curve

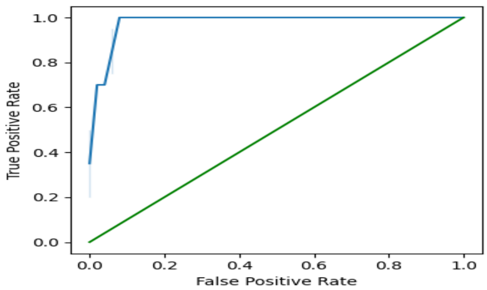


Figure 13: Bagging (SVM) ROC curve

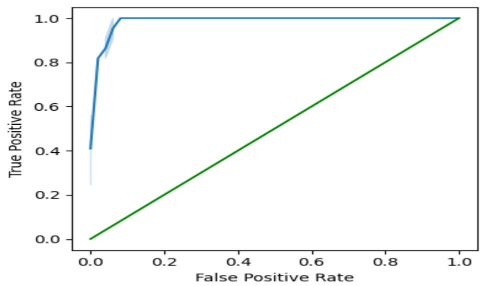


Figure 14: Bagging (MLP) ROC curve

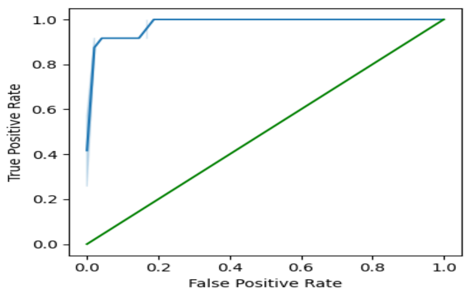


Figure 15: Stacking ROC curve

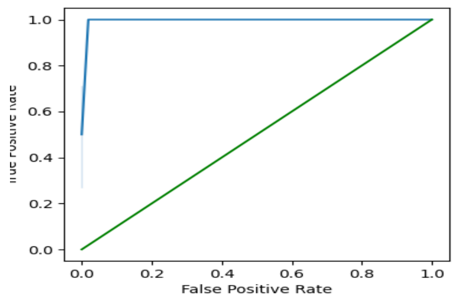


Figure 16: Stacking ROC curve

6. Conclusion

The present study showcases the progress made in classifying power transformer faults by employing ensemble techniques in using dissolved gas analysis datasets preprocessed with SMOTE and normalization techniques. To find the best performing model for the classification task, we developed and tested the ensemble of SVM and MLP models for Bagging, AdaBoost, Stacking, and Voting ensemble methods respectively. The results clearly showed that, the AdaBoost ensemble of the MLP obtained 100% accuracy after several iterations outperforming other existing methods from the literature. In future studies, we would like to introduce other ML techniques as well as computational algorithms for parameter tuning on reliable real-world datasets and compare the ensemble approaches to ascertain the best performing model. Also, we would like to curate the data obtained, apply cross validation methods and focus on the addressing the optimization of model's parameters for improved performance.

References

[1] Odongo, G., Musabe, R., & Hanyurwimfura, D. (2021). A Multinomial DGA Classifier for Incipient Fault Detection in Oil-Impregnated Power Transformers. Algorithms. Vol. 14(4), 128. https://doi.org/10.3390/a14040128

[2] Kari, T., Gao, W., Zhao, D., Abiderexiti, K., Mo, W., Wang, Y., & Luan, L. (2018). Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm. IET Generation, Transmission & Distribution, 12(21), 5672-5680. https://doi.org/10.1049/iet-gtd.2018.5482

[3] Kirkbas, A., Demircali, A., Koroglu, S., & Kizilkaya, A. (2020). Fault diagnosis of oil-immersed power transformers using common vector approac. Electric Power Systems Research. 184. https://doi.org/10.1016/j.epsr.2020.106346

[4] Islam, M.M., Lee, G., & Hettiwatte, S.N. (2018). Application of Parzen Window estimation for incipient fault diagnosis in power transformers. High Voltage, 3(4), 303-309. https://doi.org/10.1049/hve.2018.5061

[5] Rokani, V., & Kaminaris, S.D. (2020). Power transformers fault diagnosis using AI techniques. AIP Conference Proceedings, 2307(1), 020056. https://doi.org/10.1063/5.0032820

[6] Bouchaoui, L., Hemsas, K.E., & Mellah, H. (2021). Power transformer faults diagnosis using undestructive methods (Roger and IEC) and artificial neural network for dissolved gas analysis applied on the functional transformer in the Algerian north-eastern: a comparative study," Electrical Engineering & Electromechanics, (4), 3–11. https://doi.org/10.20998/2074-272X.2021.4.01

[7] Wagh, N., & Deshpande, D.M. (2014). Investigations on Incipient Fault Diagnosis of Power Transformer Using Neural Networks and Adaptive Neurofuzzy Inference System," Applied Computational Intelligence and Soft Computing, 2014(845815), 1-9. https://doi.org/10.1155/2014/845815

[8] Muthi, A., Sumarto, S., & Saputra, W.S. (2018). Power Transformer Interruption Analysis Based on Dissolved Gas Analysis (DGA) using Artificial Neural Network," IOP Conference Series: Materials Science and Engineering, 384(012073), 1-5. doi.org/10.1088/1757-899X/384/1/012073

[9] Ma, H., Zhang, W., Wu, R., & Yang, C. (2018). A Power Transformers Fault Diagnosis Model Based on Three DGA Ratios and PSO Optimization SVM," IOP Conference Series: Materials Science and Engineering, 339(012001), 1-6. https://doi.org/10.1088/1757-899X/339/1/012001

[10] Zhang, W., Yang, X., Deng, Y., & Li, A. (2020). An Inspired Machine-Learning Algorithm with a Hybrid Whale Optimization for Power Transformer PHM. Energies, 13(12), 3143. https://doi.org/10.3390/en13123143

[11] Ghoneim, S.S., & Taha, I.B. (2016). A new approach of DGA interpretation technique for transformer fault diagnosis. International Journal of Electrical Power & Energy Systems, 81, 265-274. https://doi.org/10.1016/j.ijepes.2016.02.018

[12] Aburaghiega, E., Farrag, M.F., Hepburn, D., & Haggag, A. (2018). Enhancement of Power Transformer State of Health Diagnostics Based on Fuzzy Logic System of DGA. 2018 Twentieth International Middle East Power Systems Conference (MEPCON), 400-405. https://doi.org/10.1109/MEPCON.2018.8635154

[13] Apte, S., Somalwar, R., & Wajirabadkar, A. (2018). Incipient Fault Diagnosis of Transformer by DGA Using Fuzzy Logic. 2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), 1-5. https://doi.org/10.1109/PEDES.2018.8707928

[14] Saravanan, D., Hasan, A., Singh, A., Mansoor, H., & Shaw, R.N. (2020). Fault Prediction of Transformer Using Machine Learning and DGA. 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), 1-5. https://doi.org/10.1109/GUCON48875.2020.9231086

[15] Faiz, J., & Soleimani, M. (2018). Assessment of Computational Intelligence and Conventional Dissolved Gas Analysis Methods for Transformer Fault Diagnosis. IEEE Transactions on Dielectrics and Electrical Insulation, 25(5), 1798-1806. https://doi.org/10.1109/TDEI.2018.007191

[16] Venkatesan, R., & Er, M.J. (2014). Multi-Label Classification Method Based on Extreme Learning Machines. 2014 13th International Conference on Control, Automation, Robotics & Vision Marina Bay Sands, Singapore, 10-12th December 2014 (ICARCV 2014), pp. 619-634.

[17] Duval, M., & dePablo, A. (2001). Interpretation of Gas-In-Oil Analysis Using New IEC Publication 60599 and IEC TC 10 Databases. Dielectrics and Electrical Insulation Society, 17(2), 31-41. doi.org/0883-7554/01/$10.00©2001

[18] Dhini, A., Faqih, A., Kusumoputro,B., Surjandari, I., & Kusiak, A. (2020). Data-Driven Fault Diagnosis of Power transformers using Dissolved Gas Analysis (DGA)," International Journal of Technology. 11(2), 388-399. http://ijtech.eng.ui.as.id

[19] Subasi, A. (2020). Practical Machine Learning for Data Analysis Using Python. Elsevier Science. United Kingdom.

[20] Jassim, M.A., & Abdulwahid, S.N. (2021). Data Mining preparation: Process, Techniques and Major Issues in Data Analysis. IOP Conf. Series: Materials Science and Engineering 1090 012053. https://doi:10.1088/1757-899X/1090/1/012053.

[21] Acuna, E., & Rodriguez, C. (2004). The Treatment of Missing Values and Its Effect on Classifier Accuracy. In: Classification, Clustering, and Data Mining Applications. Springer, Berlin, Heidelberg, 639-647. https://doi:/10.1007/978-3-642-17103-1_60.

[22] Chawla, N.V., Bowyer, K.W., Hall, L.O., & Kegelmeyer, W.P. (2002). SMOTE: Synthetic Minority Over-Sampling Technique. Journal Of Artificial Intelligence Research," 16, 321-357. https://doi.org/10.1613/jair.953.

[23] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A. et al., (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12(85), 2825-2830. https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html

[24] Vapnik, C.V. (1995). Support-vector networks. Machine Learning. 20, 3, 273, 1995. https://doi:10.1007/BF00994018.

[25] Nield, T. (2022). Essential Math for Data Science: Take Control of Your Data with Fundamental Linear Algebra, Probability, and Statistics. O'Reilly.

[26] Grus, J. (2019). Data Science from Scratch: First Principles with Python. O'Reilly Media.

[27] Oshiro, T.M., Perez, P.S., & Baranauskas, J.A. (2022). How Many Trees in a Random Forest? Lecture Notes in Computer Science, 7376, 155-168, 2022. https://dx.doi.org/10.1007/978-3-642-31537-4_13

[28] Zhou, Z.-H. (2012). Ensemble Methods: Foundations and Algorithms. CRC Press.

[29] Zhu, J. Rosset, S., Zou, H., & Hastie, T. (2006). Multi-class AdaBoost. Statistics and its Interface. 2(3), 2006. https://dx.doi.org/10.4310/SII.2009.v2.n3.a8

[30] Wolpert, D.H. (1992). Stacked generalization. Neural Networks. 5(2), 241-259. https://doi.org/10.1016/S0893-6080(05)80023-1

[31] Cichosz, P. (2014). "Data mining algorithms: Explained using R. West Sussex, UK: John Wiley & Sons.

[32] Flach, P. (2012). Machine learning: The art and science of algorithms that make sense of data. Cambridge, UK: Cambridge University Press.

[33] Kecman, V. (2001). Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models. The MIT Press. Cambridge Massashusetts, Londan, England.

[34] Santamaria-Bonfil, G., Arroyo-Figueroa, G., Zuniga-Garcia, M.A., Azcarraga Ramos, C.G., & Bassam, A. (2024). Power Transformer Fault Detection: A Comparison of Standard Machine Learning and autoML Approaches. Energies. 17(1):77. https://doi.org/10.3390/en17010077

[35] Wang, T., Li, Q., Yang, J., Xie, T., Wu, P., & Liang, J. (2023). Transformer Fault Diagnosis Method Based on Incomplete Data and TPE-XGBoost. Appl. Sci., 13, 7539. https://doi.org/10.3390/app13137539

[36] Zou, D., Xiang, Y., Peng, Q., Wang, S., Shi, Y., Hong, Z., Dai, W., & Zhou, T. (2022). Power Transformer Fault Diagnosis Method Based on Machine Learning. International Conference on Cyber-Physical Social Intelligence (ICCSI), Nanjing, China, pp. 484-491. https://doi:10.1109/ICCSI55536.2022.9970667

[37] Naderian, A., Piercy, R., Cress, S., Wang, F., & Service, J. (2008). An Approache to Determine the Health Index of Power Transformers. IEEE International Symposium. DOI:10.1109/ELINSL.2008:4570308.

[38] Tran, T.Q., Davies, K., Roose, L., Wiriyakitikum, P., Janjampop, J., Sanseverino, R.E., & Zizzo, G. (2020). A Review of Health Techniques for Distribution Transformers in Smart Distribution Grids. Applied Sciences. 10, 8115, pp.1-20. DOI:10.3390/app10228115

[39] Rojora, G.L., Sanz-Bobi, M.A., & Donungo, C.M. (2020). Application of Machine

Learning Methods for Asset Management on Power Distribution Networks. Engineering Science Journal. Vol. 6, No. 4, pp. 905-920.

[40] Estevesa, V., Pinto, J., & Sousa, R. (2020). Machine Learning Algorithms for Fault Data Analytics in Power Transformers: A Literature Review. Elsevier, pp. 2-14. ORCID(s): 0000-0002-2324-5719

[41] Senoussaoui, M.E.A., Brahami, M., & Fofana, I. (2021). Transformer Oil Quality Assessment Using Random Forest with Feature Engineering. Energies, 14, 1809. https://doi.org/10.3390/en14071809

[42] Zhang, X., Zhang, M., & Hu, Q. (2020). Research on the Application of Artificial Intelligence in Operation and Maintenance for Power Equipment. 2020 International Symposium on New Energy and Electrical Technology. IOP Conf. Series: Earth and Environmental Science 617 (2020) 012001. doi:10.1088/1755-1315/617/1/012001

[43] Valencia, F., Arcos, H., & Quilumba, F. (2021). Comparison of Machine Learning Algorithms for the Prediction of Mechanical Stress in Three-Phase Power Transformer Winding Conductors. Journal of Electrical and Computer Engineering Volume, Article ID 4657696, pp1-9 https://doi.org/10.1155/2021/4657696

[44] Alyunov, A.N., & Vyatkina, O.S. (2023). Identification of parameters of power transformer models using artificial intelligence methods. E3S Web of Conferences 411, 01001, pp.1-6. https://doi.org/10.1051/e3sconf/202341101001

[45] Chen, H., Meng, L., Xi1, Y., Xu, C., Wang, Y., Xin, M., Chen, G., & Chen, Y. (2023). Power transformer oil temperature prediction based on empirical mode decomposition-bidirectional long short-term memory. Journal of Algorithms & Computational Technology Volume 17: 1–9. DOI: 10.1177/17483026231176196

[46] Xing, M., Ding, W., Li, H., & Zhang, T. (2022). A Power Transformer Fault Prediction Method through Temporal Convolutional Network on Dissolved Gas Chromatography Data. Security and Communication Networks Volume 2022, Article ID 5357412, pp. 1-11. https://doi.org/10.1155/2022/5357412

[47] Asad, M., & Fareeha. (2021). Machine Learning for Predictive Maintenance of Transformer. International Research Journal of Modernization in Engineering Technology and Science. Volume:03, Issue 04, pp. 414-422. www.irjmets.com

[48] Aizpurua, J.I., & McArthur, D.J., Stewart, B.J., Lambert, B., Cross, J.G., Catterson, V.M. (2018). Adaptive Power Transformer Lifetime Predictions through Machine Learning & Uncertainty Modelling in Nuclear Power Plants. IEEE Transactions on Industrial Electronics.

[49] Mateus, B.C., Farinha, J.T., & Mendes, M. (2023). Fault Detection and Prediction for Power Transformers Using Fuzzy Logic and Neural Networks. Energies 2024, 17, 296. https://doi.org/10.3390/en17020296

[50] Ofori, D., Armah, E., Marfo, B., Otchere, K.I., Quansah, A., & Kyeremeh, K.A. (2023). Predictive Health Monitoring of a Power Transformer Using Machine Learning. Electrical and Electronic Engineering 2023, 13(1): 12-18 DOI: 10.5923/j.eee.20231301.03

[51] Bjelic, M., Brkovic, B., Zarkovic, M., & Miljkovic, T. (2024). Machine learning for Power transformer SFRA based fault detection. International Journal of Electrical Power and Energy Systems 156, pp.1-10, 109779.

[52] Balasubramaniam, Y.A/L., Yaw, C.T., Koh, S.P., Tan, C.K., Wong, J.Y., & Wong, Y.R. (2022). Power Transformer Health Prediction using Machine Learning. Journal of Engineering Research and Education Volume 14, pp. 44-54.

[53] Wang, C., Chen, Z., & Zhang, Z., Li, C. and Song, K. (2024). Power Transformer Fault Prediction Based on Knowledge Graphs, pp. 1-14. https://doi.org/10.48550/arXiv.2402.07283.

[54] Aning, J., George, P.G., & Enoch, N.A. (2020). Predicting the Remaining Lifetime of Distribution Transformers using Machine Learning. American Journal of Engineering and Applied Sciences 2020, 13 (4): 627.638. DOI: 10.3844/ajeassp.2020.627.638

[55] Suwarno, Sutikno, H., Prasojo, R.A., & Abu-Siada, A. (2024). Machine learning based multi-method interpretation to enhance dissolved gas analysis for power transformer fault diagnosis. Science Direct. Heliyon, 10, e25975, pp.1-21. https://doi.org/10.1016/j.heliyon.2024.e25975

[56] Zhang, Y., Tang, Y., Liu, Y., & Liang, Z. (2022). Fault diagnosis of transformer using artificial intelligence: A review. Front. Energy Res. 10:1006474. doi: 10.3389/fenrg.2022.1006474

[57] Li, Y. (2023). The State of the Art in transformer fault diagnosis with artificial intelligence and Dissolved Gas Analysis: A Review of the Literature. https://arxiv.org/abs/2304.11880

[58] Moni, V., & Gouri, S.S.M. (2021). Transformer Oil Degradation Prediction Using Machine Learning. International Journal of Engineering Research & Technology (IJERT). Vol. 10 Issue 10, pp. 105-108. https://www.ijert.org/

[59] Taha, I.B.M., & Mansour, D-E. A. (2021). Novel Power Transformer Fault Diagnosis Using Optimized Machine Learning Methods. Intelligent Automation & Soft Computing (IASC), vol. 28, no.3, pp. 730-752. DOI:10.32604/iasc.2021.017703

[60] Ngwenyama, M. K., & Gitau, M.N. (2024). Discernment of transformer oil stray gassing anomalies using machine learning classification techniques. Scientific Reports, 14:376. https://doi.org/10.1038/s41598-023-50833-7

[61] Seijo-Pardo, S., Porto-Diaz, I., Bolon-Canedo, V., & Alonso-Betanzos, A. (2017). Ensemble Feature Selection: Homogeneous and Heterogeneous Approaches. Knowledge-Based System. 118:124-139

[62] Seijo-Pardo, B., Bolon-Canedo, V., & Porto-Diaz, I. (2018). Ensemble Feature Selection for Rankings of Features. Springer, 2015, pp. 29-42. DOI: 10.1007/978-3-319-19222-2-3.

[63] Colombelli, F., Kowalski, T.W., & Recamonde-Mendoza, M. (2022). A hybrid ensemble feature selection design for candidate biomarkers discovery from transcriptome profiles. Knowledge-Based Systems, Volume 254, 109655, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2022.109655.

[64] Osanaiye, O. Cai, H., Choo, K-K., Dehghantanha, R., Xu, Z., & Dlodlo, M. (2016). Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. EURASIP Journal on Wireless Communications and Networking, pp. 1-10. DOI 10.1186/s13638-016-0623-3.