Volume 19, Issue 2, 192543 (1-12)

Majlesi Journal of Electrical Engineering (MJEE)

https://doi.org/10.57647/j.mjee.2025.1902.43

Accuracy enhancement of fault diagnosis for power transformers with a hybrid approach integrating robust and tree-based algorithms

Elahe Moradi*

Department of Electrical Engineering, YI.C., Islamic Azad University, Tehran, Iran.

*Corresponding author: elahe.moradi@iau.ac.ir

Original Research

Received:
10 December 2024
Revised:
23 March 2025
Accepted:
19 April 2025
Published online:
1 June 2025

Abstract:

Power transformers (PTs) are a significant component of power grids that transmit and distribute electricity generated by renewable energy sources. Nevertheless, PTs are susceptible to faults that can cause costly outages and disruptions. Over the past decades, the technique of dissolved gas analysis (DGA) has been extensively employed in oil-immersed transformer fault diagnosis. There are various methods to identify faults using DGA. Due to its superior accuracy compared to other techniques, the dual pentagon method (DPM) is utilized for fault diagnosis of PTs in this research. On the other hand, implementing DPM on large amounts of DGA data can be challenging. To address this issue, we proposed several data-driven, tree-based algorithms, including Decision Tree Classifier (DTC), Random Forest Classifier (RFC), eXtreme Gradient Boosting Classifier (XGBC), LightGBM (LGBM) Classifier, Adaptive Boosting (AdaBoost) Classifier, and Categorical Boosting (CatBoost) Classifier. Furthermore, four data scaling techniques have been used for more effectiveness because the dataset contains outliers. The outcomes of the data analysis and Python simulation demonstrate that the suggested approach performs better than the previous methods. From the simulation analysis, the robust Light-GBM method has achieved an accuracy of 96.08%, and MCC of 95.41%, which is higher compared to the existing techniques.

Keywords: Robust technique; Power transformer; Fault diagnosis; Random forest; Machine learning; Light gradient boosting machine ;robust technique

1. Introduction

In industrial environments, high-voltage power transformers (PT) are mostly needed for heavy-duty and powerful applications. These transformers use specific insulation systems that are generally dependent primarily on the voltage levels. Thus, the lifetime and reliability of the PT are more affected by the higher voltage [1]. Power transformers are essential yet expensive components in power systems, pivotal in energy generation and transmission. Ensuring their reliability is vital, and various diagnostic methods, such as Dissolved Gas Analysis (DGA) and Frequency Response Analysis (FRA), have been developed to detect and classify faults. The application of advanced intelligent classifiers has proven to be highly effective in improving the fault diagnosis process, significantly enhancing the accuracy and reliability of these methods [2]. On the other hand, when exposed to various flaws resulting from overheating, arcing, paper carbonization, and low or high-energy discharges, the PT insulation systems might deteriorate [3]. To ensure that these PTs are an effective service, early-stage detection of faults must be conducted [4, 5].

For this objective, numerous techniques were suggested in the literature. Among the various methods, DGA is one of the most economical, fastest, and most widely used for the early diagnosis of faults PT [6–8]. Methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), and hydrogen (H_2) are the main gases that are measured and analyzed for DGA, and the carbon–oxide gases carbon monoxide (CO) and carbon dioxide (CO_2). Though nitrogen (N_2) and dissolved oxygen (O_2) are measured and evaluated, they are not results of insulation deterioration [9].

Though the measurement of DGA in the PT has been significantly advanced in recent years, interpretation of the outcomes is a challenging issue. The interpretation methods like the Doernenburg Ratio Method (DRM), IEC ratio method (IRM), Duval triangle method (DTM), and Roger ratio method (RRM) include limitations and disadvantages in terms of accuracy [10–13]. In recent years, the dual pentagonal method (DPM) has been utilized to diagnose PT faults, which has achieved higher accuracy and consistency [10].

Many researchers have employed artificial intelligence (AI) methods to improve the accuracy of fault diagnostic predictions for PT based on DGA. Support vector machine (SVM) algorithms were combined with particle swarm optimization (PSO) method in [14] to diagnose fault types in PT and the accuracy of this approach was 83,56%. In [15], an iterative nearest neighbor interpolation was combined with ensemble learning algorithm for improving fault diagnosis of PT. This algorithm has achieved an accuracy rate exceeding 91% for the PT diagnosis. Artificial neural network (ANN) model with an optimization technique was employed to improve the diagnostic precision of identifying PT faults in [16] and after implementing the optimization methods, the proposed ANN showed a high accuracy, reaching 90.7%. In [17], the PT fault diagnosis method was recommended that combined the residual backpropagation (BP) neural network with SVM and the suggested method achieved a high accuracy rate of 92%. A DGA diagnosis method has been used with a clustering approach combined with a modified KNN and the results demonstrated a high degree of accuracy, reaching 93% in [18]. In [19], a new asset management method for mineral oil-immersed PT by the online DGA approach using convolutional neural network (CNN) has been investigated and the result displayed the accuracy of fault diagnostics PT achieved an accuracy 87%. A novel approach combining oversampling and cost-sensitive learning has been proposed to enhance the accuracy of diagnosis of all fault types of PT, and the accuracy of the proposed method could reach over 90% [20]. A multi-layer perceptron (MLP) neural network and an ensemble model have been presented in [21] and the outcomes showed a high degree of accuracy, reaching 91.87%. In [22], a power transformer fault diagnosis model was proposed based on optimization of hybrid kernel extreme learning machine and the proposed algorithm achieved a high accuracy of 94.8%. In [23], random forest (RF) and optimal kernel Extreme Learning Machine (ELM) was presented for PT fault diagnosis and the average accuracy of this approach achieved 94.5%. A transformer fault diagnosis method based on multi-classification AdaBoost algorithms was suggested in [24] and its diagnostic accuracy was up to 92.5%. A new transformer fault diagnosis algorithm based on combining tree-structured probability density estimator (TPE) and XG-Boost model was proposed in [25] and the average accuracy of the proposed method was 89.5%. A hybrid model including naive bayes and decision tree algorithms for identification of PT faults was proposed ibn [26] and this model achieved an accuracy of 86.25%. In [27], a decision tree model was improved the KNN classifier to increase the accuracy rate of fault diagnosis in PT and the proposed technique found an accuracy rate of 93% for the PT diagnosis. The proposed algorithms for fault diagnosis of power transformers in [23–27] were tree-based. Tree-based methods are powerful and versatile tools for machine learning, offering a balance between interpretability, accuracy, and robustness.

Two of the main advantages of tree-based methods are robustness to outliers and non-parametric nature. In previous studies, normalization and standardization methods have been used, which are sensitive to the presence of outliers in the dataset. Due to the presence of outliers in the data, an outlier-resistant method has been applied to the dataset. The main contribution of this research is the development of a robust classifier with more accuracy than the above-mentioned algorithms. Also, in most previous research, conventional criteria such as precision, accuracy, recall, and F-measure have been used. In this paper, the MCC criterion is employed to evaluate the performance of the proposed technique. Although the previously mentioned methods suggest a straightforward implementation in practical contexts, their accuracy and robustness can be further enhanced by employing hybrid algorithms.

Therefore, the main aims of this study can be delineated as follows:

- Robust against the outliers: This study suggests a robust technique in which data is resistance to outliers.
- Enhance the accuracy of power transformer fault diagnosis: Robust tree-based algorithms such as DTC (Decision Tree Classifier), RFC (Random Forest Classifier), XGBC (Extreme Gradient Boosting Classifier), AdaBoost (Adaptive Boosting Classifier), LightGBM (Light Gradient Boosting Machine), and CatBoost (Categorical Boosting) are utilized for the data with outliers and improved the accuracy of prediction.
- Employing the MCC criterion: In this research, additional to Accuracy, Precision, AUC, F1-Measure criteria, the MCC criterion is used to evaluate the performance of the proposed methods with four data scaling techniques.

This work is structured as follows. The initial section provides an introduction to the research topic. Subsequently, the second section delves into the application of the DPM approach to the dissolved gas analysis (DGA) of power transformers. The third section outlines the materials and the proposed hybrid method for fault diagnosis of power transformers. The fourth section presents the findings and discusses the results obtained from power transformer fault diagnosis. Finally, the fifth section draws conclusions on the research conducted.

2. Power transformer dissolved gas analysis

Dissolved Gas Analysis (DGA) is widely used for power transformer fault diagnosis. Analyzing the concentration and kinds of dissolved gases in a power transformer's insulating oil is part of the procedure. Various gases have particular quantities and ratios that show to different kinds of faults [28]. These gases are mainly hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), carbon monoxide (CO), and carbon dioxide (CO_2). These gasses can be produced by partial discharge, thermal decomposition, or sustained arcing faults. Particularly, partial discharge produces a smaller amount of CH_4 and a greater amount of H_2 [10]. Arcing generates a noticeable amount of C_2H_2, while overheating produces CH_4, C_2H_6, and C_2H_4. The existence of CO_2, and CO in the PT can be a sign of research degradation [1, 29].

2.1 DGA interpretation techniques

There are several techniques for DGA interpretation that the methodology of the four DGA are summarized in Table 1 [1, 13, 28–30].

Several recent research [21, 28, 31], the DPM success rate outperforms the DTM, because it can detect the transformer insulation's normal aging. In this paper, the Duval Pentagon technique has been used to identify primary faults in the PT. Seven various fault types of the DPM are displayed in figure 1 [12] and described in Table 2.

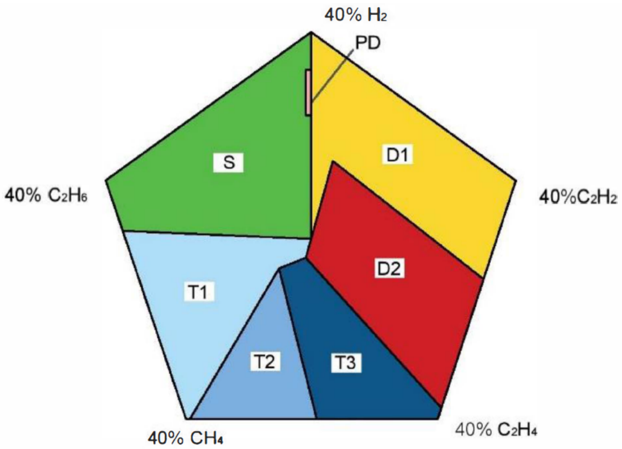

Figure 1. Visual depiction of the DPM.

2.2 DPM equations

DPM uses the following five gases: g_1 (H_2), g_2 (CH_4), g_3 (C_2H_6), g_4 (C_2H_4), and g_5 (C_2H_2) as their input data. The computation for the percentage of gases in DPM is provided in equations (1) to (5) [32].

$$\%\text{Hydrogen, } H_2 = \frac{g_1}{\sum_{i=1}^{5} g_i} \times 100 \tag{1}$$

$$\%\text{Methene, } CH_4 = \frac{g_2}{\sum_{i=1}^{5} g_i} \times 100 \tag{2}$$

$$\%\text{Ethane, } C_2H_6 = \frac{g_3}{\sum_{i=1}^{5} g_i} \times 100 \tag{3}$$

$$\%\text{Ethylene, } C_2H_4 = \frac{g_4}{\sum_{i=1}^{5} g_i} \times 100 \tag{4}$$

$$\%\text{Acetylene, } C_2H_2 = \frac{g_5}{\sum_{i=1}^{5} g_i} \times 100 \tag{5}$$

Afterward, equations (6) and (7) were used to determine the (x, y) coordinates of the centroid of these five points, where C_x and C_{xy} denote the centroid's (x, y) coordinates and x_i and y_i are the coordinates of the five points. The fault could be specified by which fault zone the centroid falls. Equation (8) computes the irregular pentagon's surface, S, using the five points demonstrated in DPM [32, 33].

$$C_x = \frac{1}{6S}\sum_{i=0}^{n-1}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \tag{6}$$

$$C_y = \frac{1}{6S}\sum_{i=0}^{n-1}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \tag{7}$$

$$S = \frac{1}{2}\sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i) \tag{8}$$

3. Materials and methodologies

3.1 Dataset

In this study, we used the DGA dataset initially utilized in [21]. The dataset has 1658 samples of seven different types of faults, which are described in Table 2. The features of the dataset are five different types of gases in power transformers including hydrogen, Methene, Ethane, Ethylene, and Acetylene in ppm. They operate as diagnostic indicators for various power transformer problems. The distribution of samples in fault types on DPM are as follows: D1 (233 data), D2 (237 data), T1 (239 data), T2 (241 data), T3 (240 data), PD (241 data), and S (227 data). Figure 2 displays samples percentage in each fault type on DPM.

During the preprocessing stage of the DGA dataset, both missing values and outliers were carefully analyzed to ensure data integrity. While no missing values were detected

Table 1. Comparison of the four DGA interpretation methods.

Technique	Faults identification	Used gases
Doernenburg Ratio Method (DRM)	Thermal Decomposition (TD), arcing, Partial Discharge (PD)	H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6
IEC Ratio Method (IRM)	PD, Low/High energy discharge, Thermal faults at several temperatures.	H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6
Duval Triangle Method (DTM)	PD, Low/High energy discharge, Thermal faults at several temperatures.	CH_4, $_2H_2$, C_2H_4
Duval Pentagon Method (DPM)	PD, Low/High energy discharge, Normal aging, Thermal faults at several temperatures.	H_2, C_2H_2, CH_4, C_2H_6, and C_2H_4

Table 2. Description of Seven various fault types of the DPM.

Type of faults	Description of faults
Type 1	PD: Corona Partial Discharges
Type 2	T1: Low Thermal fault
Type 3	T2: Medium Thermal fault
Type 4	T3: High thermal fault
Type 5	D1: Discharge of low-energy
Type 6	D2: Discharge of high-energy
Type 7	S: Stray gassing at low temperatures

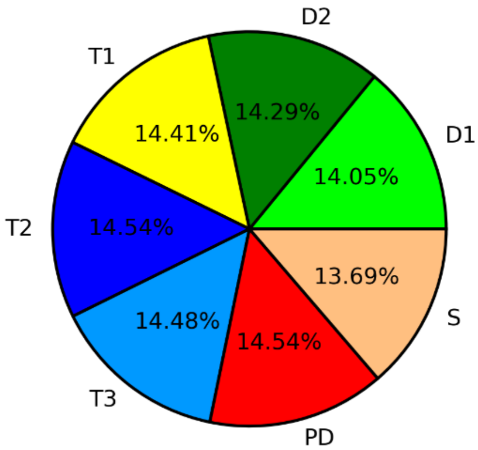

Figure 2. Samples percentage in each fault type on DPM.

in the dataset, a significant number of outliers were identified. These outliers, often caused by measurement errors or rare events, were addressed using the Interquartile Range (IQR) method. This method detects data points that fall outside the range defined by 1.5 times the interquartile range above the third quartile or below the first quartile. By handling outliers appropriately, the reliability and robustness of the dataset were enhanced.
The flowchart of the fault diagnosis model of power transformer is displayed in figure 3.

3.2 Data scaling

A technique for data preparation called feature scaling enables the independent variables in a dataset to be standardized within a specified range of values. There exist distinct approaches for the dataset's data scaling. The four approaches are presented in equations (9) to (12), and are called Standard Scaling (SS), Min-Max Scaling (MMS), Max Absolute Scaling (MAS), and Robust Scaling (RS), respectively. It is worth noting that most previous studies have used the SS, and the MMS techniques for data scaling of the datasets [34, 35].
The Standard Scaling (SS) technique is described as follows:

$$x_{sik} = \frac{x_{ik} - x_{kmean}}{x_{kstd}} \tag{9}$$

The Min-Max Scaling (MMS) methos can depicted by the following Eq. (10):

$$x_{sik} = \frac{x_{ik} - x_{k\min}}{x_{k\max} - x_{k\min}} \tag{10}$$

The Max Absolute Scaling (MAS) method can be presented in Eq. (11):

$$x_{sik} = \frac{x_{ik}}{|x_{k\max}|} \tag{11}$$

The Robust Scaling (RS) technique can be showed by the following Eq. (12):

$$x_{sik} = \frac{x_{ik} - x_{kmedian}}{\text{IQR}} \tag{12}$$

where x_{ik} and x_{sik} are the values of the i^{th} sample for the k^{th} feature before and after scaling. Also, $x_{kmedian}$ is the mean, and x_{kstd} is the standard deviation of samples. Additional, $x_{k\min}$ and $x_{k\max}$ are minimum, and maximum feature value, respectively. Lastly, $x_{kmedian}$ presents middle value of data and the IQR denotes the distance between the 25th and the 50th percentile points, respectively [35].

3.3 Splitting of data

The original dataset is split into training and test sets. The proportion of the training and the test sets is 80:20. Hence the number of samples in the training set is 1326 and the test set contains 332 samples. The Scikit-learn library and the "train_test_split" command were utilized in this study to achieve the objective [21]. The k-fold Cross-Validation technique is not dependent on the number of training samples. The dataset is divided into k subsets to prevent overfitting. Accordingly, for each iteration, a single subset is held out of the training and the model is trained on the other $k-1$ subsets before being tested on that different single subset.
A five-fold cross-validation is employed to randomly split the samples into a training set and a testing set. This method effectively assesses the performance of the training model while enhancing its stability and generalization capabilities. Using a Jupyter notebook and the Python programming language, all the experiments required for comparing the ML algorithms was conducted. The first pre-processing steps were conducted out using standard Python frameworks, including NumPy, Pandas, Matplotlib, and Seaborn. The Sklearn framework was employed for training most ensemble ML algorithms as well as classical methods [35].

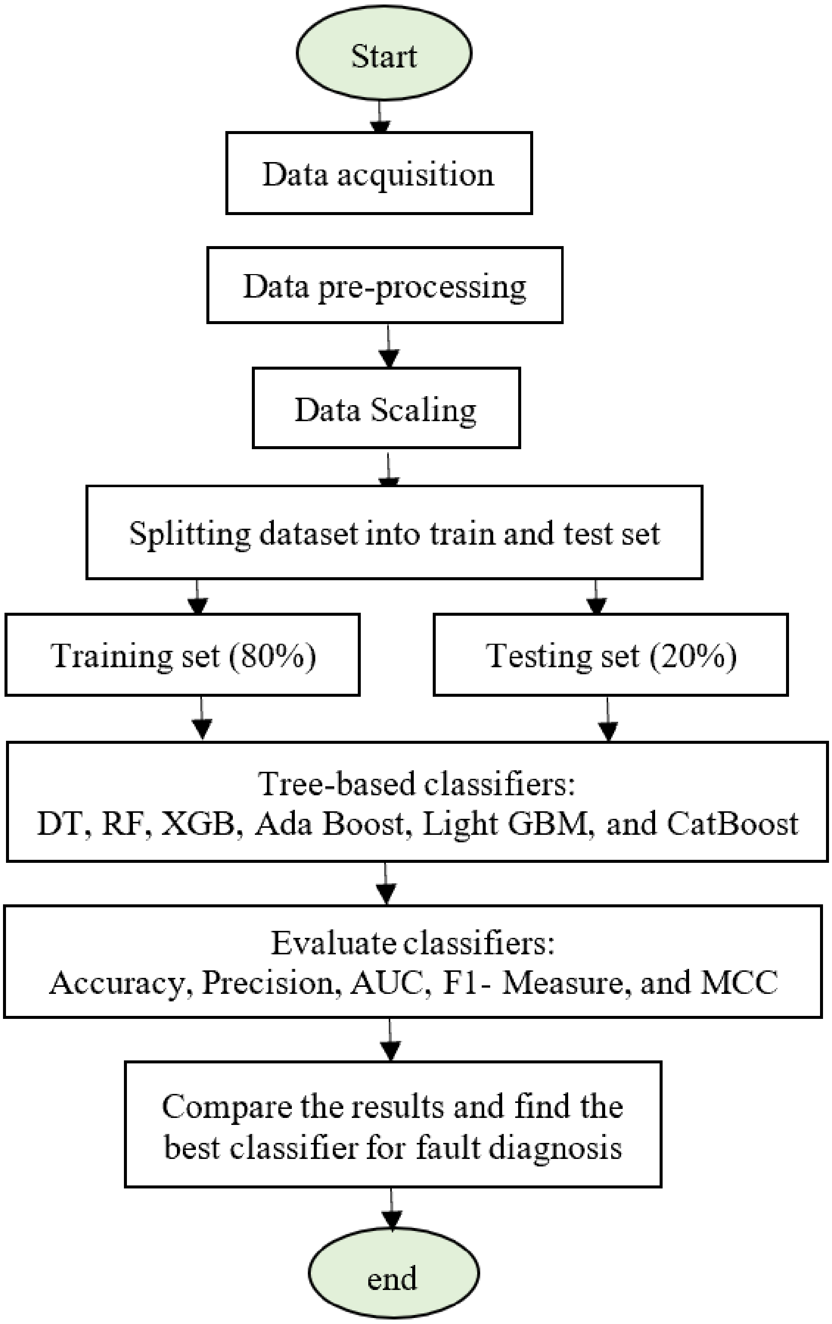

Figure 3. Fault diagnosis flowchart of the tree-based classifiers.

3.4 Tree-based classifiers

In this research, the six ML algorithms-based tree classifiers and the five data scaling approaches are employed together to find the best match for fault diagnosis of the power transformer. After data pre-processing, we selected six tree-based classifiers including Decision Tree Classifier (DTC), Random Forest Classifier (RFC), Extreme Gradient Boosting Classifier (XGBC), AdaBoost Classifier (ABC), LightGBM Classifier (LGC), and CatBoost Classifier (CBC).

- Decision Tree Classifier (DTC)

A decision tree (DT) is a type of recursive data partitioning method, specifically for the classification of datasets. The structure consists of nodes, which are assembled sequentially into a tree. A group of nodes that have one incoming edge each will be directly following the root node which, in turn, is the single node with no incoming edges. Any other node, aside from the root, is referred to as a leaf or a decision node. Nodes with outgoing edges are called internal nodes. Based on specific requirements or specifications of the input variables, each internal node splits the dataset into two or more sub-datasets [36]. This method has several benefits, including the ability to deal with both continuous and discrete attributes, the simplicity with which humans can make small decision trees as decision rules, and the fact that it is a nonparametric technique, meaning that it doesn't need any additional tunings to enhance accuracy. In spite of these profits, DT methods are expensive to train and may lead to error associations [37].

- Random Forest Classifier (RFC)

A parallel type of ensemble algorithm is the random forest classifier (RFC). A combination of decision trees with the same distribution for every tree that only evaluates a random subset of the data is called a random forest. To decrease variance, the vital idea is to take the average of numerous noisy nevertheless roughly unbiased models [38]. RFC is certainly more accurate than an individual DTC because it easily reduces overfitting due to using the advantages of bagging and randomness. Also, one of the advantages of RF is its usability, both for classifier and regressor problems [37].

- Extreme Gradient Boosting Classifier (XGBC)

XGBoost, another DT-based boosting algorithm, is called. "eXtreme Gradient Boosting," and it also differs from previous boosting approaches. XGB is currently used in ML models, and it can be applied to both regressor and classifier problems. It does this by skillfully applying the principles of Parallelization and Gradient Descent (GD) to the Boosting Ensemble method, which allows for the simultaneous achievement of the optimal possible combination of hardware and software. Gradient boosting (GB) builds trees sequentially, but XGB builds the trees sequentially using parallelized implementation; Therefore, weights in GB method are not derived from the misclassifications of the former model, but rather from the weights optimized by the GD to minimalize the cost function. The foundation of XGB is an optimized distributed GB foundation [39].

- AdaBoost Classifier (ABC)

One of the initially efficient boosting-based algorithms for classifications, AdaBoost (AB), further referred to as AB created a foundation for scientists to comprehend boosting ensembles. Though versions of AB for multi-class problems have been introduced in recent years, this method was initially created for binary classification. Utilizing the original boosting issues, AB sequentially adds base models instead of in a combined manner, giving greater weight to the dataset's misclassified cases with each subsequent base model. This procedure keeps going until the training dataset can further be enhanced. Any technique for ML algorithms can be made to perform better by using the AB model [40].

- LightGBM Classifier (LGC)

In this work, the LightGBM (LG) algorithm is employed. The fundamental idea is to utilize the output from the previous training round as input for the subsequent learning round. To split the DT, the LG algorithm uses a superior optimization model, basically through the histogram, then the classic GB methods, which must traverse the whole data set several times [41].

- CatBoost Classifier (CBC)

CatBoost (CB) is a ML framework that supports categorical variables and observes the gradient boosting (GB) process framework. It can efficiently address numerous data migration issues inherent in the original GB framework, and contribution advantages such as reduced parameter requirements, high accuracy, and robust performance. Ensemble learning involves creating multiple ML algorithms, training them to produce several weak learners, and then integrating these weak learners through multiple combination strategies to advance a strong learner [42].

3.5 Evaluation criteria

To enhance the accuracy of DGA fault diagnosis, various criteria are applied to assess the machine learning algorithms. True Negative (TN) is an instance where a prediction of negative data is correct, while True Positive (TP) indicates the prediction of positive data. A False negative (FN) is a case in which positive data is indeed predicted as negative. In contrast, False Positive (FP) is a situation when negative data is incorrectly predicted as positive. There are various measures to assess the machine learning models performance based on confusion matrix (CM) like accuracy, AUC (Area Under the Curve), F1-Measure, precision, and MCC [43, 44].

Accuracy (ACC) is defined the ratio of the true prediction (positive and negative) to the total data as presented by Eq. (13):

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{FN} + \text{TN} + \text{FP} + \text{TP}} \tag{13}$$

Precision (PR) is the ratio of a positive true prediction to a total positive predicted result as denoted by Eq. (14):

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \tag{14}$$

The AUC is a performance metric derived from the Receiver Operating Characteristic (ROC) curve. It quantifies the ability of a model to distinguish between different classes. AUC values range from 0 to 1, where higher values indicate better classification performance.

F1-Measure (F1) is weighted mean comparison of recall and precision as presented by Eq. (15):

$$\text{F1-Measure} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{15}$$

In most previous papers, evaluation criteria such as precision, accuracy, recall and F1-measure have been employed. In this research, in addition to these metrics, the Matthews Correlation Coefficient (MCC criterion) has been utilized by Eq. (16):

$$\text{MCC} = \frac{\text{TN} \times \text{TP} - \text{FN} \times \text{FP}}{(\text{FP} + \text{TP}) \times (\text{FN} + \text{TP}) \times (\text{FP} + \text{TN}) \times (\text{FN} + \text{TN})} \tag{16}$$

4. Results and discussion

In this section, the performance of the six proposed methods combined with four data scaling techniques for fault diagnosis in power transformers is evaluated using various metrics. Tables 3, 4, 5, and 6 present the performance of

Table 3. Performance of the proposed methods using the SS technique.

Methods	ACC (%)	PR (%)	AUC (%)	F1 (%)	MCC (%)
DTC	84.59	85.58	98.91	84.26	82.24
RFC	90.03	90.54	99.68	90.07	88.41
XGBC	87.01	87.36	99.50	86.91	84.90
ABC	87.01	87.26	99.67	87.05	84.84
LGC	92.45	92.59	99.65	92.42	91.20
CBC	91.547 91.89	99.71	91.54	90.16	

Table 4. Performance of the proposed methods using the MMS technique.

Methods	ACC (%)	PR (%)	AUC (%)	F1 (%)	MCC (%)
DTC	83.13	83.57	98.47	82.96	80.38
RFC	91.87	92.31	99.53	91.94	90.55
XGBC	87.35	88.17	96.54	87.51	85.31
ABC	89.16	90.72	99.53	89.48	87.50
LGC	94.28	94.59	99.36	94.30	93.36
CBC	93.07	93.19	98.45	93.10	91.90

hybrid methods with SS, MMS, MAS, and RS data scaling techniques, respectively, using metrics such as ACC, PR, AUC, F1-Measure, and MCC.
Additionally, bar charts including figures 4, 5, 6, 7, and 8 are presented for the Accuracy, Precision, AUC, F1-Measure, and MCC metrics on a percentage basis to compare the proposed methods. Based on the aforementioned tables and figures, it is observed that the three proposed methods RFC, LGC, and CBC outperform the other methods.
To further evaluate the four data scaling techniques, the

Table 5. Performance of the proposed methods using the MAS technique.

Methods	ACC (%)	PR (%)	AUC (%)	F1 (%)	MCC (%)
DTC	82.83	85.4	97.99	83.00	80.38
RFC	91.27	91.84	99.76	91.32	89.87
XGBC	90.66	91.43	97.74	90.86	89.15
ABC	88.55	89.86	99.43	88.77	86.77
LGC	93.07	93.43	98.84	94.04	91.98
CBC	92.77	92.89	98.49	92.80	91.55

Table 6. Performance of the proposed methods using the RS technique.

Methods	ACC (%)	PR (%)	AUC (%)	F1 (%)	MCC (%)
DTC	86.45	86.95	97.42	86.56	84.18
RFC	92.47	92.69	99.53	92.3	91.23
XGBC	90.36	90.45	98.76	90.31	88.71
ABC	89.76	90.93	99.46	90.03	88.09
LGC	96.08	96.09	99.64	96.06	95.41
CBC	94.28	94.47	99.71	94.22	93.32

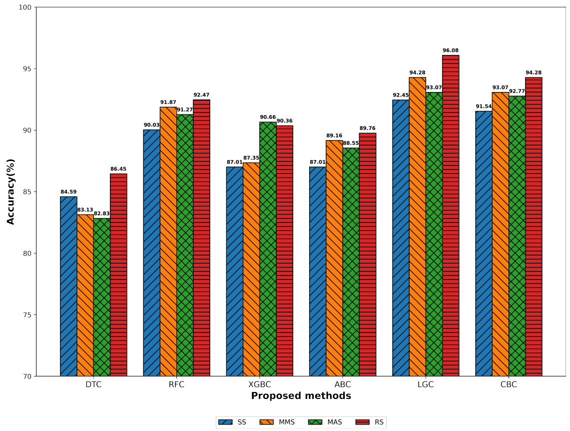

Figure 4. Evaluating the accuracy criterion of proposed methods with data scaling techniques.

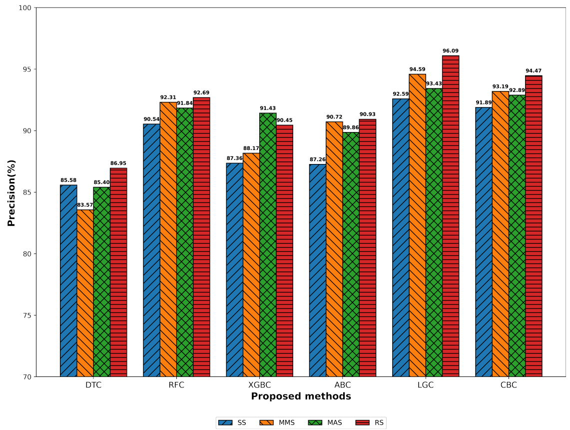

Figure 5. Evaluating the precision criterion of proposed methods with data scaling techniques.

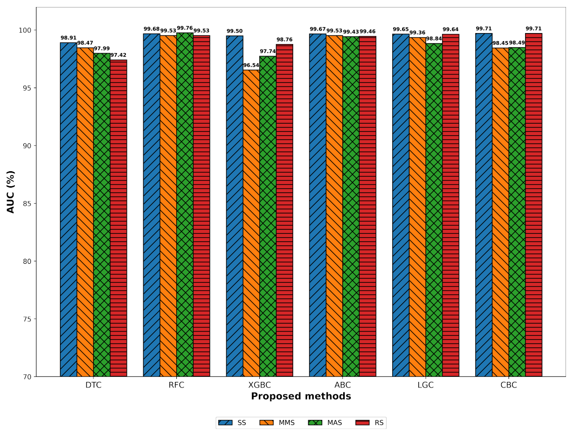

Figure 6. Evaluating the AUC criterion of proposed methods with data scaling techniques.

important MCC metric for fault diagnosis in power transformers is presented for the six hybrid proposed methods using the DPM in figure 9. As can be seen, the Robust Scaling technique has the best performance among the proposed hybrid classifiers.

Figures 10, 11, 12, and 13 survey the ROC curves and AUC metric, which indicate that the hybrid methods RFC, LGC, and CBC have appropriate performance. Furthermore, the confusion matrices in figures 14, 15, and 16 suggest that

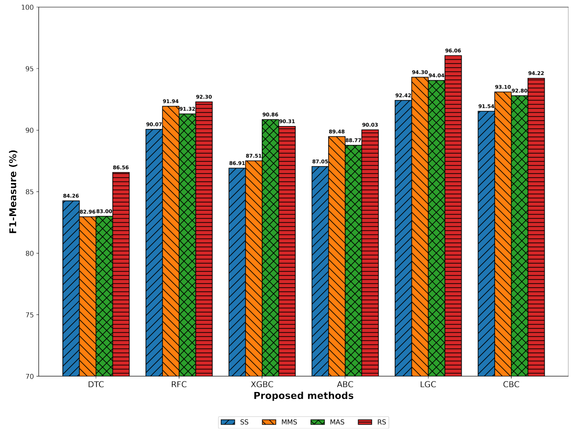

Figure 7. Evaluating the F1-Measure criterion of proposed methods with data scaling techniques.

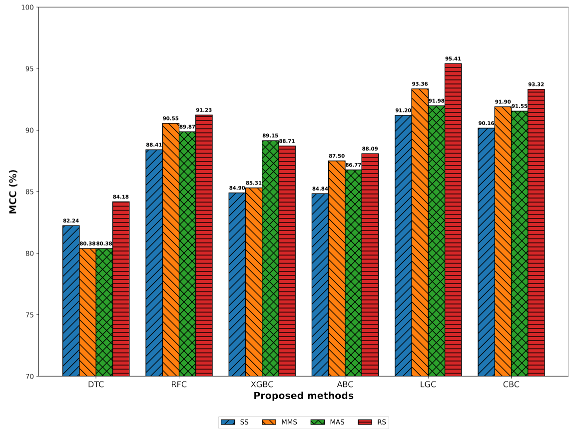

Figure 8. Evaluating the MCC criterion of proposed methods with data scaling techniques.

the hybrid LGC method with Robust Scaler has excellent performance, demonstrating that the training and testing data are well-handled by this classifier.

Therefore, the hybrid methods RFC, LGC, and CBC, when combined with the Robust Scaling technique, present the best performance for fault diagnosis in power transformers. Based on the simulations conducted in Python, the hybrid LightGBM Classifier with Robust Scaling technique achieves the highest performance compared to previous

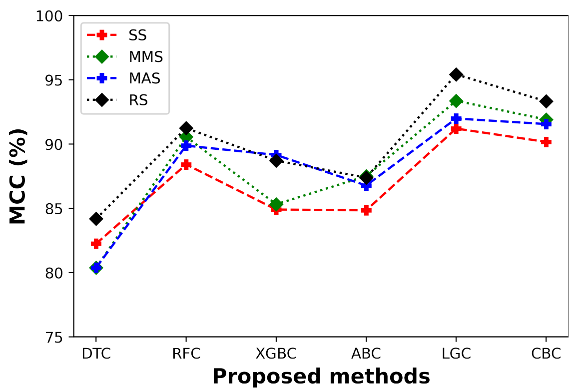

Figure 9. Comparison of the MCC criteria using the four techniques for fault prediction.

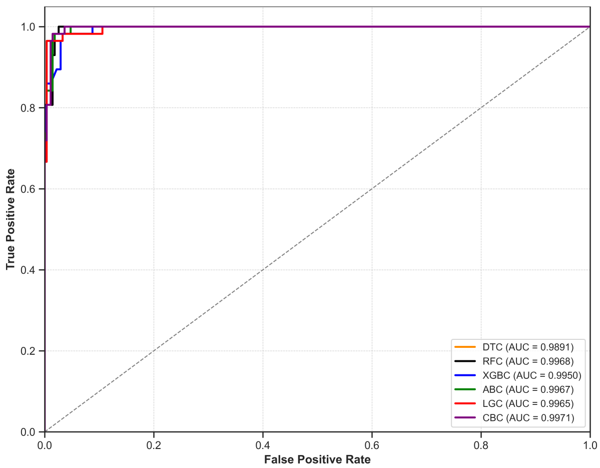

Figure 10. The ROC curve of the proposed methods using the SS technique.

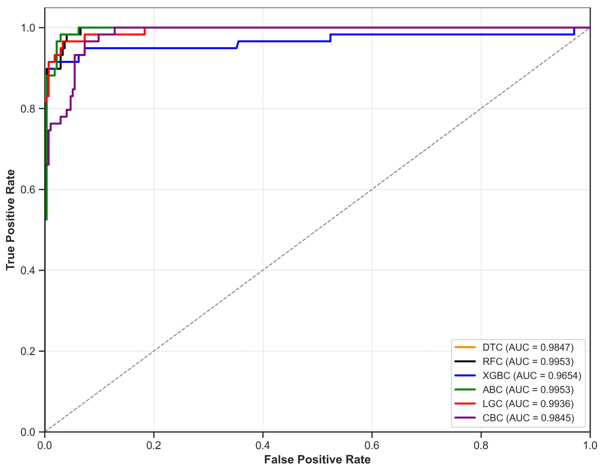

Figure 11. The ROC curve of the proposed methods using the MMS technique.

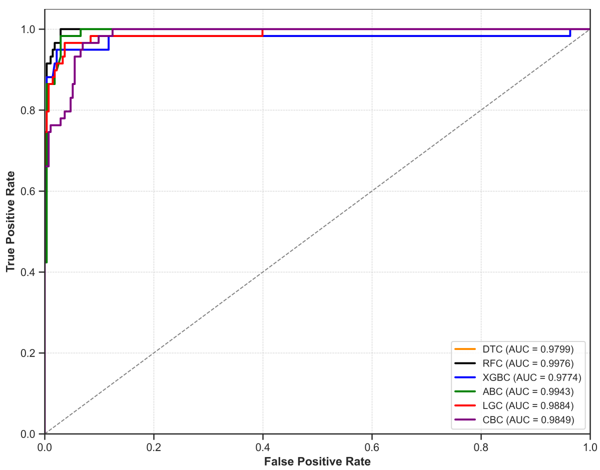

Figure 12. The ROC curve of the proposed methods using the MAS technique.

studies. This is attributed to its enhanced accuracy and robustness, reflected in the criteria of Accuracy (96.08%), Precision (96.09%), AUC (99.64%), F1-Measure (96.06%), and MCC (95.41%).
It is noteworthy that among the six proposed classifiers for

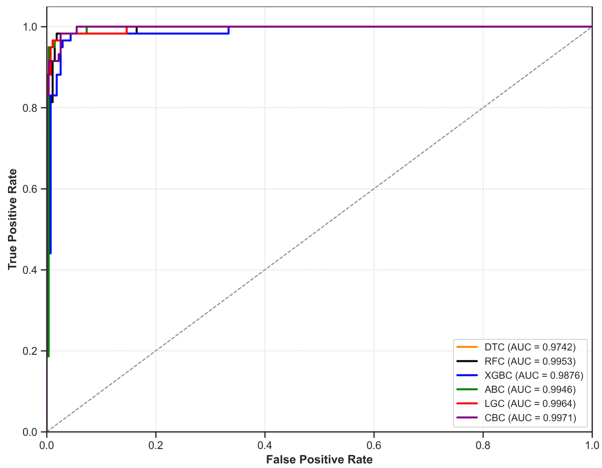

Figure 13. The ROC curve of the proposed methods using the RS technique.

error detection, the LightGBM Classifier (LGC) exhibited the best performance across all four scaling techniques employed in the study. Consequently, the computational times for this classifier across the four techniques are as follows: MAS, SS, RS, and MMS took 675.25 seconds, 676.42 seconds, 399.95 seconds, and 425.64 seconds, respectively.

5. Conclusion

Ensuring a consistent and efficient power supply relies on the dependable functioning of power transformers, which play a vital role in the power grid. Nevertheless, transformers can encounter various problems that may lead to costly disruptions and power outages. Therefore, timely maintenance and prevention depend on accurate fault detection. In this work, we suggested hybrid approaches to improve the power transformer fault diagnosis robustness and accuracy. The suggested classifiers successfully categorize fault-related features from the dissolved gas analysis (DGA) dataset by combining robust and tree-based methods such as Decision Tree, Random Forest, XGBoost, LightGBM, AdaBoost, and CatBoost. Based on benchmark dataset experiments, the hybrid suggested classifiers perform much better in diagnosis accuracy than the state-of-the-art techniques. Particularly, the LightGBM classifier, combined with a robust scaling technique, achieves an accuracy rate of over 96%. This research can be extended by investing the application of Internet of Things (IoT) in real-time sample testing.

Authors contributions
Authors have contributed equally in preparing and writing the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

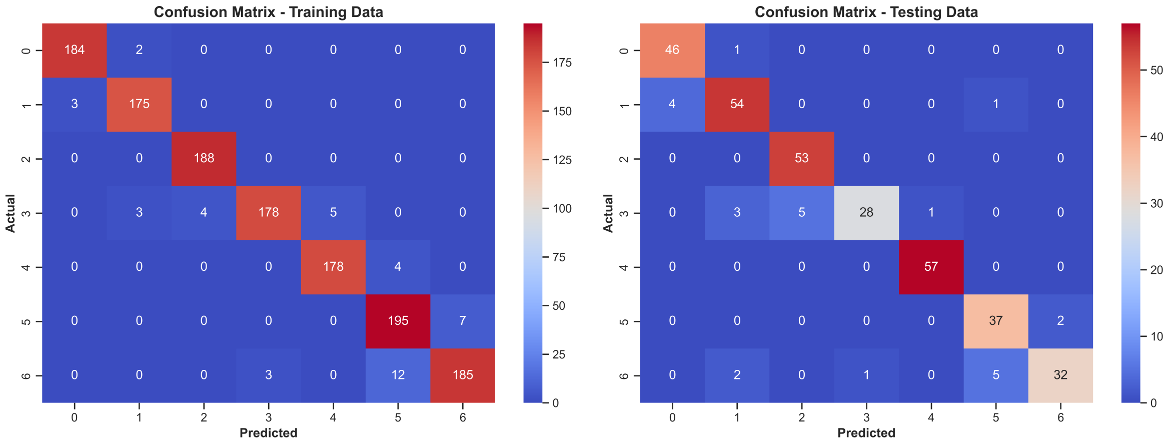

Figure 14. Confusion Matrix of RFC with the RS technique.

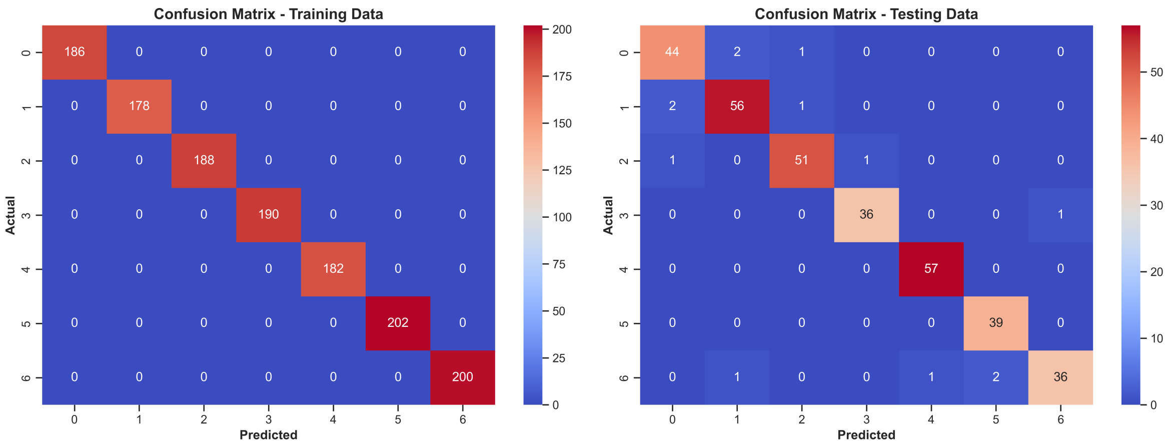

Figure 15. Confusion Matrix of LGC with the RS technique.

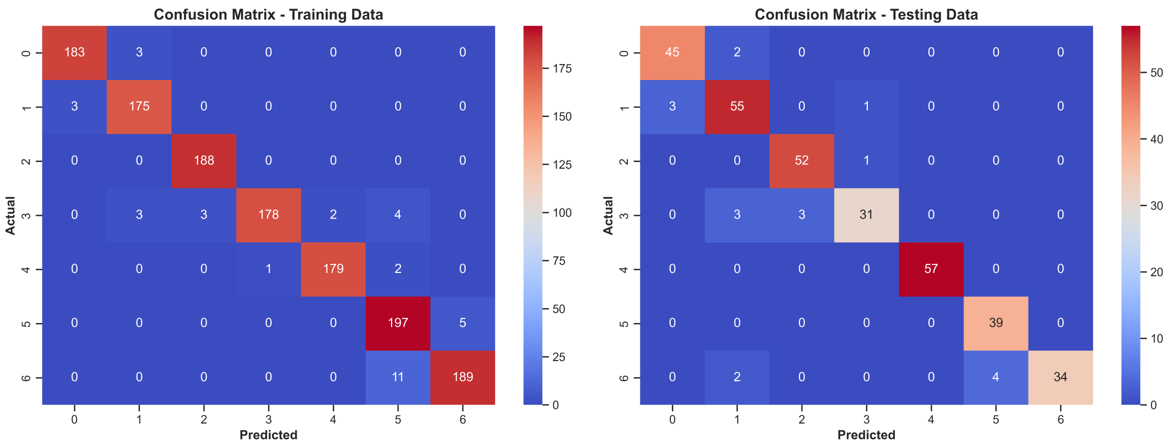

Figure 16. Confusion Matrix of CBC with the RS technique.

References

[1] IEEE Guide for the Interpretation of Gases Generated in ;Mineral Oil-Immersed Transformers. *IEEE Std C57.104-2019*, pages 1–98, 2019.
DOI: https://doi.org/10.1109/IEEESTD.2019.8890040.

[2] M. Bigdeli, P. Siano, and H. H. Alhelou. "**Intelligent Classifiers in Distinguishing Transformer Faults Using Frequency Response Analysis**.". *IEEE Access*, 9:13981–13991, 2021.
DOI: https://doi.org/10.1109/access.2021.3052144.

[3] M. Duval. "**Dissolved gas analysis: It can save your transformer**.". *IEEE Electrical Insulation Magazine*, 5(6):22–27, 1989.
DOI: https://doi.org/10.1109/57.44605.

[4] E. H S. S. M. Ghoneim. "**Intelligent prediction of transformer faults and severities based on dissolved gas analysis integrated with thermodynamics theory**.". *IET Science, Measurement & Technology*, 12(3):388–394, 2018.
DOI: https://doi.org/10.1049/iet-smt.2017.0450.

[5] I. B. M. Taha, A. Hoballah, and S. S. M. Ghoneim. "**Optimal ratio limits of rogers' four-ratios and IEC 60599 code methods using particle swarm optimization fuzzy-logic approach**.". *IEEE Transactions on Dielectrics and Electrical Insulation*, 27(1):222–230, 2020.
DOI: https://doi.org/10.1109/tdei.2019.008395.

[6] P. R. M. Giri, I. M. Y. Negara, and D. A. Asfani. "**Recent Development in DGA Diagnosis Using Graphical Analysis Method**.". *International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2021.
DOI: https://doi.org/10.1109/ISITIA52817.2021.9502223.

[7] L. Cheng and T. Yu. "**Dissolved Gas Analysis Principle-Based Intelligent Approaches to Fault Diagnosis and Decision Making for Large Oil-Immersed Power Transformers: A Survey**.". *Energies*, 11(4):913, 2018.
DOI: https://doi.org/10.3390/en11040913.

[8] A. Nanfak, S. Eke, C. H. Kom, R. Mouangue, and I. Fofana. "**Interpreting dissolved gases in transformer oil: A new method based on the analysis of labelled fault data**.". *IET Generation, Transmission & Distribution*, 15(21):3032–3047, 2021.
DOI: https://doi.org/10.1049/gtd2.12239.

[9] J. Dukarm and F. Jakob. "**Thermodynamic estimation of transformer fault severity**.". *Case Studies in Thermal Engineering*, 2016.
DOI: https://doi.org/10.1109/TDC.2016.7519962.

[10] H. Suwarno, R. A. Sutikno, Prasojo, and A. Abu-Siada. "**Machine learning based multi-method interpretation to enhance dissolved gas analysis for power transformer fault diagnosis**.". *Heliyon*, 10 (4):1–21, 2024.
DOI: https://doi.org/10.1016/j.heliyon.2024.e25975.

[11] N. Pattanadech and W. Wattakapaiboon. "**Application of Duval Pentagon Compared with Other DGA Interpretation Techniques: Case Studies for Actual Transformer Inspections Including Experience from Power Plants in Thailand**.". *Gazi Journal of Engineering Sciences*, 2019.
DOI: https://doi.org/10.1109/iceast.2019.8802523.

[12] W. Wattakapaiboon and N. Pattanadech. "**The state of the art for dissolved gas analysis based on interpretation techniques**.". 2016.
DOI: https://doi.org/10.1109/cmd.2016.7757763.

[13] A. Abu-Siada. "**Improved Consistent Interpretation Approach of Fault Type within Power Transformers Using Dissolved Gas Analysis and Gene Expression Programming**.". *Energies*, 12(4): 730, 2019.
DOI: https://doi.org/10.3390/en12040730.

[14] W. Mo, T. Kari, H. Wang, L. Luan, and W. Gao. "**Power Transformer Fault Diagnosis Using Support Vector Machine and Particle Swarm Optimization**.". 2017.
DOI: https://doi.org/10.1109/iscid.2017.165.

[15] O. Kherif, Y. Benmahamed, M. Teguar, A. Boubakeur, and S. S. M. Ghoneim. "**Accuracy Improvement of Power Transformer Faults Diagnostic Using KNN Classifier with Decision Tree Principle**.". *IEEE Access*, 9:81693–81701, 2021.
DOI: https://doi.org/10.1109/access.2021.3086135.

[16] V. Rokani, S. D. Kaminaris, P. Karaisas, and D. Kaminaris. "**Power Transformer Fault Diagnosis Using Neural Network Optimization Techniques**.". *Mathematics*, 11(22):4693, 2023.
DOI: https://doi.org/10.3390/math11224693.

[17] Y. Jin, H. Wu, J. Zheng, J. Zhang, and Z. Liu. "**Power Transformer Fault Diagnosis Based on Improved BP Neural Network**.". *Electronics*, 12(16):3526, 2023.
DOI: https://doi.org/10.3390/electronics12163526.

[18] M. M. Islam, G. Lee, and S. N. Hettiwatte. "**A nearest neighbor clustering approach for incipient fault diagnosis of power transformers**.". *Electrical Engineering*, 99(3):1109–1119, 2016.
DOI: https://doi.org/10.1007/s00202-016-0481-3.

[19] L. Jin, D. Kim, K. Y. Chan, and A. Abu-Siada. "**Deep Machine Learning-based Asset Management Approach for Oil-Immersed Power transformers using Dissolved Gas Analysis**.". *IEEE Access*, 2024.
DOI: https://doi.org/10.1109/access.2024.3366905.

[20] L. Zhang, G. Sheng, H. Hou, N. Zhou, and X. Jiang. "**An adaptive fault diagnosis method of power transformers based on combining oversampling and cost-sensitive learning**.". *IET Smart Grid*, 4 (6):623–635, 2021.
DOI: https://doi.org/10.1049/stg2.12044.

[21] E. Moradi. "**A Data-Driven based Robust Multilayer Perceptron Approach for Fault Diagnosis of Power Transformers**.". 2024.
DOI: https://doi.org/10.1109/aisp61396.2024.10475257.

[22] J. Li, C. Hai, Z. Feng, and G. Li. "**A Transformer Fault Diagnosis Method Based on Parameters Optimization of Hybrid Kernel Extreme Learning Machine**.". *IEEE Access*, 9:126891–126902, 2021.
DOI: https://doi.org/10.1109/ACCESS.2021.3112478.

[23] T. Kari, Z. He, A. Rouzi, Z. Zhang, X. Ma, and L. Du. "**Power Transformer Fault Diagnosis Using Random Forest and Optimized Kernel Extreme Learning Machine**.". *Intelligent Automation and Soft Computing/Intelligent Automation & Soft Computing*, 37(1): 691–705, 2023.
DOI: https://doi.org/10.32604/iasc.2023.037617.

[24] J. Li, G. Li, C. Hai, and M. Guo. "**Transformer Fault Diagnosis Based on Multi-Class AdaBoost Algorithm**.". *IEEE Access*, 10: 1522–1532, 2021.
DOI: https://doi.org/10.1109/ACCESS.2021.3135467.

[25] T. Wang, Q. Li, J. Yang, T. Xie, P. Wu, and J. Liang. "**Transformer Fault Diagnosis Method Based on Incomplete Data and TPE-XGBoost**.". *Applied Sciences*, 13(13):7539, 2023.
DOI: https://doi.org/10.3390/app13137539.

[26] Y. Mahamdi, A. Boubakeur, A. Mekhaldi, and Y. Benmahamed. "**Power Transformer Fault Prediction using Naive Bayes and Decision tree based on Dissolved Gas Analysis**.". *ENP Engineering Science Journal*, 2(1):1–5, 2022.
DOI: https://doi.org/10.53907/enpesj.v2i1.63.

[27] A. G. C. Menezes, M. M. Araujo, O. M. Almeida, F. R. Barbosa, and A. P. S. Braga. "**Induction of Decision Trees to Diagnose Incipient Faults in Power Transformers**.". *IEEE Transactions on Dielectrics and Electrical Insulation*, 29(1):279–286, 2022.
DOI: https://doi.org/10.1109/TDEI.2022.3148453.

[28] S. Bustamante, M. Manana, A. Arroyo, P. Castro, A. Laso, and R. Martinez. "**Dissolved Gas Analysis Equipment for Online Monitoring of Transformer Oil: A Review**.". *Sensors*, 19(19): 4057, 2019.
DOI: https://doi.org/10.3390/s19194057.

[29] L.-Z. Wang, J.-F. Chi, Y.-Q. Ding, H.-Y. Yao, Q. Guo, and H.-Q. Yang. "**Transformer fault diagnosis method based on SMOTE and NGO-GBDT**.". *Scientific Reports*, 14(1), 2024.
DOI: https://doi.org/10.1007/s00202-023-02084-y.

[30] D. Martin, N. Lelekakis, J. Wijaya, M. Duval, and T. Saha. "**Investigations into the Stray Gassing of Oils in the Fault Diagnosis of Transformers**.". *IEEE Transactions on Power Delivery*, 29(5):2369–2374, 2014.
DOI: https://doi.org/10.1109/TPWRD.2014.2316501.

[31] L. Cheim, M. Duval, and S. Haider. "**Combined Duval Pentagons: A Simplified Approach**.". *Energies*, 13(11):2859, 2020.
DOI: https://doi.org/10.3390/en13112859.

[32] M. S. Ali et al. "**Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review**.". *Electric Power Systems Research*, 2023.
DOI: https://doi.org/10.1016/j.epsr.2022.109064.

[33] M. Duval and L. Lamarre. "**The Duval pentagon-a new complementary tool for the interpretation of dissolved gas analysis in transformers**.". *IEEE Electrical Insulation Magazine*, 30(6):9–12, 2014.
DOI: https://doi.org/10.1109/MEI.2014.6943428.

[34] M. Ahsan, M. Mahmud, P. Saha, K. Gupta, and Z. Siddique. "**Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance**.". *Technologies*, 9(3):1–17, 2021.
DOI: https://doi.org/10.3390/technologies9030052.

[35] G. Santamaria-Bonfil, G. Arroyo-Figueroa, M. A. Zuniga-Garcia, C. G. A. Ramos, and A. Bassam. "**Power Transformer Fault Detection: A Comparison of Standard Machine Learning and autoML Approaches**.". *Energies*, 17(1):1–22, 2023.
DOI: https://doi.org/10.3390/en17010077.

[36] L. Rokach and O. Maimon. "**Data Mining with Decision Trees: Theory and Applications; 2nd Edition**.". 2014.
DOI: https://doi.org/10.1142/9097.

[37] M. Pirizadeh, N. Alemohammad, M. Manthouri, and M. Pirizadeh. "**A new machine learning ensemble model for class imbalance problem of screening enhanced oil recovery methods**.". *Journal of Petroleum Science & Engineering*, 198:1–22, 2021.
DOI: https://doi.org/10.1016/j.petrol.2020.108214.

[38] G. Biau and E. Scornet. "**A random forest guided tour**.". *arXiv.1511.05741*, 2015.
DOI: https://doi.org/10.48550/arXiv.1511.05741.

[39] T. Chen and C. Guestrin. "**Xgboost: a scalable tree boosting system**.". *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
DOI: https://doi.org/10.1145/2939672.2939785.

[40] T. Hastie, S. Rosset, J. Zhu, and H. Zou. "**Multi-class AdaBoost**.". 2(3):349–360, 2009.
DOI: https://doi.org/10.4310/SII.2009.v2.n3.a8.

[41] G. Tao, Z. Wang, and X. Zhang. "**Power Grid Fault Diagnosis Method Based on LightGBM**.". *Journal of Physics. Conference Series*, 2023.
DOI: https://doi.org/10.1088/1742-6596/2479/1/012029.

[42] G. Huang, L. Wu, X. Ma, W. Zhang, J. Fan, X. Yu, W. Zeng, and H. Zhou. "**Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions**.". *J. Hydrol.*, 574: 1029–1041, 2019.
DOI: https://doi.org/10.1016/j.jhydrol.2019.04.085.

[43] A. A. Yaghoubi, P. Karimi, E. Moradi, and R. Gavagsaz-Ghoachani. "**Implementing Engineering Education Based on Posing a Riddle in Field of Instrumentation and Artificial Intelligence**.". *9th International Conference on Control, Instrumentation and Automation (ICCIA)*, 2023.
DOI: https://doi.org/10.1109/ICCIA61416.2023.10506384.

[44] D. Chicco and G. Jurman. "**The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification**.". *BioData Mining*, 16(4):1–23, 2023.
DOI: https://doi.org/10.1186/s13040-023-00322-4.