Transformer fault diagnosis using dissolved gas analysis: a hybrid ensemble model with data preprocessing

Fatima Zohra Boudjella[1], Souhila Boudjella[2], Nasiru Yahaya Ahmed[3], Hazlee Azil Illias[3], Smail Latifa[4], Reriballah Hafidha[5]

[1]Engineering and Sustainable Development Laboratory, Department of Electrical Engineering, University of Ain Temouchent, Ain Temouchent, Algeria
[2]LGPCS Laboratory, Department of Electrical Engineering, University of Mascara, Mascara, Algeria
[3]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[4]LSTER Laboratory, Department of Electrical Engineering, University Center of El Bayadh, El Bayadh, Algeria
[5]GIDD Laboratory, Department of Electrical Engineering and Automation, University Ahmed Zabana of Relizane, Relizane, Algeria

Article Info

Article history:

Received Mar 13, 2026
Revised Jul 14, 2026
Accepted Jul 21, 2026

Keywords:

Data preprocessing
Dissolved gas analysis
Gas ratio methods
Hybrid ensemble models
Logarithmic transformation
Machine learning classifiers
Square root transformation

ABSTRACT

Failures and guaranteed dependability of the electrical grid, early fault diagnosis in power transformers is essential. By examining gas ratios suggestive of faults, dissolved gas analysis (DGA) continues to be a vital component for transformer health monitoring. Using four preprocessing techniques raw data, min-max normalization, logarithmic transformation, and square root transformation; this study suggests a machine learning method for fault detection using DGA gas ratios (such as CH_4/H_2, C_2H_2/C_2H_4). Random forest (RF), support vector machines (SVM), gradient boosted trees (GBT), and a hybrid RF-GBT model that uses prediction fusion were the four supervised classifiers assessed. Performance was assessed using accuracy, precision, recall, f1-score, and Cohen's kappa. Experimental results show that the hybrid RF-GBT model with logarithmic transformation achieves the highest performance, with 94.93% accuracy and 92.37% Cohen's kappa, significantly outperforming individual classifiers. Data preprocessing, particularly logarithmic and square root transformations, enhances diagnostic robustness by mitigating feature skewness. This study underscores the importance of tailored preprocessing and ensemble methods for reliable transformer fault diagnosis.

Corresponding Author:

Fatima Zohra Boudjella
Engineering and Sustainable Development Laboratory, Department of Electrical Engineering
University of Ain Temouchent
Ain Temouchent, Algeria
Email: fatima.boudjella@univ-temouchent.edu.dz

1. INTRODUCTION

Power transformers are vital components of electrical grids, enabling efficient energy conversion and transfer across voltage levels. Their failure can lead to prolonged outages, significant financial losses, and safety risks for personnel and equipment [1], [2]. For predictive maintenance to reduce these risks, effective condition monitoring is therefore crucial. By examining gases dissolved in insulating oil, dissolved gas analysis (DGA), a commonly used non-invasive method, can identify transformer faults like partial discharges, thermal overheating, and electrical arcing [3], [4]. The Duval Triangle, Rogers Ratios, and Dornenburg approach are examples of traditional DGA interpretation techniques that use empirical thresholds to determine fault types [5]. However, these techniques frequently produce inconsistent diagnoses due to their lack of precision in ambiguous cases or non-standard conditions. New opportunities for improving DGA reliability have been made possible by the development of artificial intelligence (AI) and machine learning (ML) [6], [7]. Numerous studies have investigated the application of intelligent systems to improve transformer fault diagnosis. For example, a multistage support vector machine (SVM) trained on IEC TC10 DGA data achieved a recognition rate of about 90%, demonstrating the potential of machine learning techniques for accurate transformer fault classification [8]. The combination of machine learning methods and conventional diagnostic indicators is still being investigated in recent research. Nițu *et al.* [9] proposed a method combining the three ratio technique with a random forest (RF) classifier to identify transformer faults using DGA gas ratio indicators. Another study proposed a diagnostic framework based on advanced feature extraction and a LightGBM ensemble model, where data preprocessing and feature selection improved classification accuracy from 86.84% to 93.42% [10]. Notwithstanding these developments, DGA-based diagnosis still faces difficulties. Class imbalance, extreme values, and operational variability are examples of data quality problems that can impair model performance [11]. Furthermore, standards such as IEC 60599 recommend the use of gas ratios for more reliable fault diagnosis [12]. In addition, the choice of data preprocessing techniques, including normalization and data transformation, can significantly influence classifier performance, yet this aspect remains insufficiently investigated in the literature.

To address these limitations, this study goes beyond conventional approaches by explicitly investigating the interaction between data preprocessing techniques and ensemble learning models for DGA-based transformer fault diagnosis. While previous works typically focus on either classifier optimization or single preprocessing strategies, the combined effect of multiple data transformations on ensemble model performance remains insufficiently explored. In this context, gas ratios are employed as diagnostic features instead of raw gas concentrations to ensure robustness and alignment with IEC standards. Multiple preprocessing techniques, including raw data, min–max normalization, logarithmic transformation, and square-root transformation, are systematically evaluated to assess their influence on feature distribution and classification performance. Furthermore, this study introduces a hybrid RF–GBT model based on prediction fusion, designed to exploit the complementary strengths of ensemble learning techniques. Unlike traditional benchmarking studies, the proposed approach emphasizes the joint optimization of feature representation and model architecture, providing a structured framework for improving diagnostic reliability. The models are evaluated through multi-class fault classification (electrical faults, thermal faults, and normal conditions) using comprehensive performance metrics such as accuracy, precision, recall, F1-score, and Cohen's kappa. Overall, the main contribution of this work lies in providing a systematic and application-oriented framework that bridges preprocessing strategies and ensemble learning, thereby offering improved robustness and reliability in DGA-based transformer fault diagnosis.

2. METHOD

2.1. Dataset

The dataset was derived from DGA reports of operational power transformers located in various provinces across Algeria, comprising 648 samples. Instead of absolute gas concentrations (ppm), gas ratios (CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/CH_4, C_2H_4/C_2H_6) were used, as recommended by the Modified Rogers' Ratio Method [13], to enhance robustness against operational variability. The samples were evenly distributed across three fault classes:

i) Electrical faults (E): partial discharges and arcing (216 samples).
ii) Thermal faults (T): cellulose or conductor overheating (216 samples).
iii) Normal state (N): no fault detected (216 samples).

Each sample includes four gas ratios selected for their diagnostic relevance. Table 1 presents representative examples of dissolved gas concentrations for each transformer condition (electrical, thermal, and normal), illustrating the typical diagnostic patterns observed for each type of fault. These four ratios CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/CH_4, and C_2H_4/C_2H_6 were selected based on their diagnostic relevance, as they are commonly used in industry standards and reflect distinct fault mechanisms.

Table 1. Representative gas concentrations (ppm) for each transformer condition

H2	CH4	C2H2	C2H4	C2H6	State	number
523.00	44.00	1.50	1.50	17.00	E	216
2526.30	130.55	1.00	1.53	15.00		
291	1260	232	6	9	T	216
3911	4291	627	6	1231		
0.1	1.1	2.02	3.1	4.02	N	216
2	2.5	0.5	0	5		

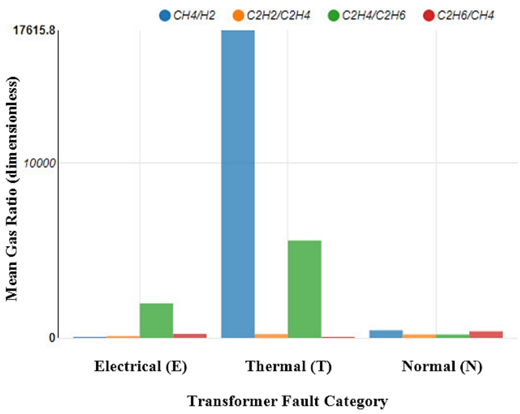
DGA data for transformer diagnostics often exhibit non-linear distributions, varying scales, and high variability. Several factors contribute to this variability, such as operating conditions, fault duration, measurement inaccuracies, and environmental influences. Direct processing of these raw data can lead to a significant degradation in the performance of machine learning models, especially in algorithms sensitive to scales. Prior to transformation, all samples containing missing values were removed to ensure dataset completeness and to avoid introducing bias during model training. Data preprocessing and normalization aim to mitigate issues related to scale, distribution skewness, and noise. This step is particularly important for algorithms sensitive to input distributions, such as SVM:

i) Reduce the impact of large amplitudes: Some variables, such as hydrogen or ethylene, can reach very high values compared to others, thereby skewing distance calculations or gradients.
ii) Ensure fairness between variables: By scaling all variables to a common range, we prevent any single gas from dominating the model's decision.
iii) Improve algorithm convergence: For models like SVM, normalization enables faster and more stable convergence.
iv) Reduce the effect of extreme or outlier values: Certain transformations, such as logarithms or square roots, compress high values, making distributions more symmetric.

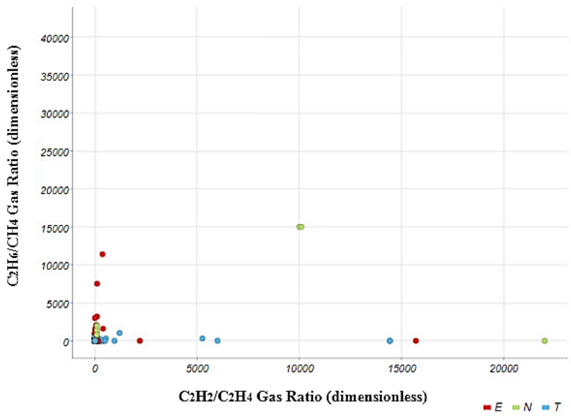
Although the dataset is derived from operational transformers located in various provinces across Algeria, it includes a diverse range of operating conditions and fault scenarios representative of real-world power systems. The dataset covers three major classes of transformer conditions (electrical faults, thermal faults, and normal operation), which correspond to the most common fault categories reported in the literature and industrial standards. The use of gas ratios instead of absolute gas concentrations further enhances the robustness and generalizability of the proposed approach. Unlike ppm values, gas ratios are less sensitive to transformer size, oil volume, and operational conditions, making them more suitable for cross-system applications. This aligns with IEC 60599 recommendations, where ratio-based interpretation is preferred for consistent fault diagnosis. However, it is acknowledged that the dataset originates from a specific geographical and operational context, which may limit its full generalization. Future work will focus on validating the proposed methodology using larger, multi-source datasets from different regions and transformer types to further assess its scalability and robustness. In this work, four main data transformation methods were used.

2.1.1. Raw data

Figure 1 illustrates the distribution of the original (non-transformed) dataset used as a baseline to evaluate the actual effect of preprocessing techniques. In Figure 1(a), the bar chart shows the variation of key gas ratios across different fault types (electrical faults (E), thermal faults (T), and normal condition (N)). The CH_4/H_2 ratio is clearly dominant for thermal faults (T), indicating significant methane production due to oil overheating, making it an excellent indicator of thermal faults. The C_2H_4/C_2H_6 ratio is also high for thermal faults, reflecting moderate system overheating, while it remains relatively low for electrical faults (E) and the normal state (N). Figure 1(b) presents a scatter plot of C_2H_2/C_2H_4 versus C_2H_6/CH_4 ratios. The distribution reveals a high degree of overlap between fault classes, with no clear linear or separable boundary between electrical, thermal, and normal conditions. This dispersion indicates strong variability and nonlinearity in the raw data, which makes direct classification challenging for standard machine learning models.



(a)



(b)

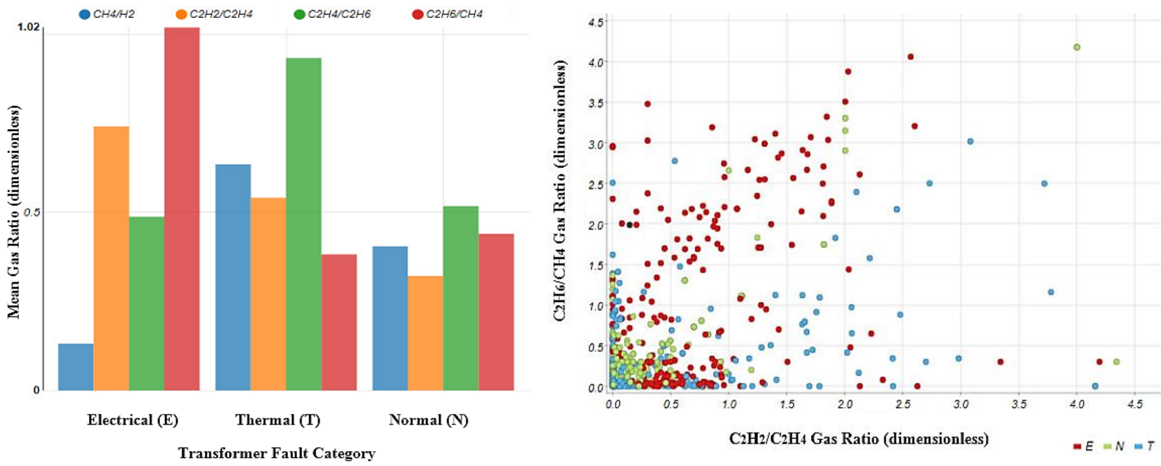
Figure 1. Dataset distribution (raw data): (a) average gas ratio values for each transformer fault category and (b) scatter plot illustrating the relationship between the C_2H_2/C_2H_4 and C_2H_6/CH_4 gas ratios

2.1.2. Logarithmic transformation

Useful for linearizing the data and reducing the influence of extreme values.

$$X' = \log(x + 1) \tag{1}$$

Logarithmic transformation significantly enhances data distribution and readability, as shown in Figure 2. In the bar chart (Figure 2(a)), the C_2H_4/C_2H_6 ratio is elevated for thermal faults (T), reflecting overheating, while the C_2H_2/C_2H_4 ratio is higher for electrical faults (E), indicative of arcing or partial discharges. The CH_4/H_2 ratio effectively discriminates thermal faults, and C_2H_6/CH_4 is prominent for electrical faults, with the normal state (N) showing balanced ratios. The scatter plot (Figure 2(b)) (C_2H_2/C_2H_4 vs. C_2H_6/CH_4) reveals improved class separation for E, T, and N, as logarithmic transformation reduces the impact of extreme values, resulting in more uniform distributions. This preprocessing enhances the performance of classification models.



(a) (b)

Figure 2. Dataset distribution (LOG): (a) average gas ratio values for each transformer fault category, and (b) scatter plot illustrating the relationship between the C_2H_2/C_2H_4 and C_2H_6/CH_4 gas ratios

2.1.3 Min-max normalization

Scale features to the [0, 1] range using (2) [14].

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

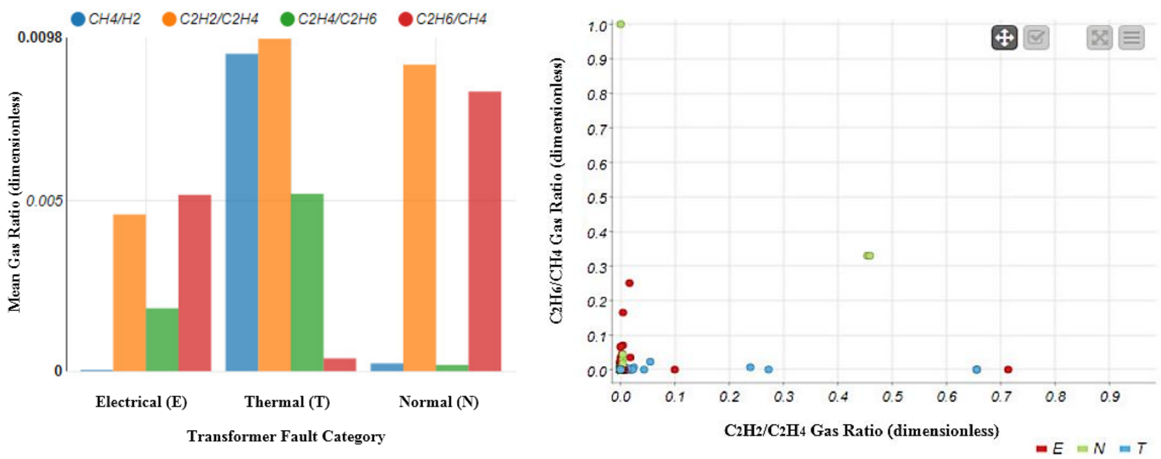
Where X is the original value, x_{min} and x_{max} are the minimum and maximum values of the feature, respectively, and X_{norm} is the normalized value. This technique standardizes the scales of gas ratios (e.g., CH_4/H_2, C_2H_2/C_2H_4), ensuring their comparability. Figure 3 presents the dataset distribution after applying min–max normalization. Min–max normalization rescales gas ratios (e.g., CH_4/H_2, C_2H_6/CH_4) to the [0, 1] range, enhancing their comparability, as illustrated in Figure 3(a). However, it does not significantly reduce dispersion compared to raw data, as observed in Figure 3(b), where the scatter plot (C_2H_2/C_2H_4 vs. C_2H_6/CH_4) still shows considerable class overlap between electrical faults (E), thermal faults (T), and normal conditions (N). Figure 3(a) presents the histogram, which reveals more distinguishable average ratios, with C_2H_6/CH_4 dominant for electrical faults (E) and CH_4/H_2 for thermal faults (T); nevertheless, class separation remains limited compared to the logarithmic transformation (Figure 2).

2.1.4. Square root transformation

The square root transformation highlights small variations while attenuating large amplitudes, according to (3).

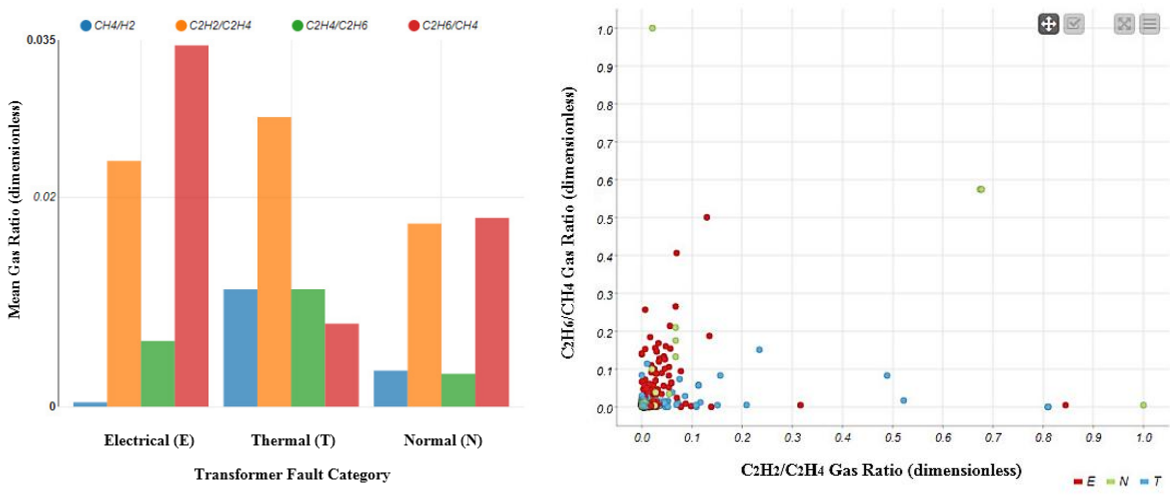
$$X' = \frac{\sqrt{X}}{\sqrt{X_{max}}} \tag{3}$$

Figure 4 presents the dataset distribution after applying the square root transformation. The square root transformation reduces dispersion in gas ratios (e.g., CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/CH_4, C_2H_4/C_2H_6), enhancing comparability across fault types, as shown in Figure 4. In the bar chart (Figure 4(a)), average ratios are more distinguishable, with C_2H_6/CH_4 prominent for electrical faults (E) and C_2H_2/C_2H_4 elevated for thermal faults (T), while normal states (N) show balanced ratios. The scatter plot (Figure 4(b)) (C_2H_2/C_2H_4 vs. C_2H_6/CH_4) reveals improved separation of classes E, T, and N due to reduced skewness, similar to logarithmic transformation, thereby enhancing classifier performance.



(a) (b)

Figure 3. Dataset distribution (min-max): (a) average gas ratio values for each transformer fault category, and (b) scatter plot illustrating the relationship between the C_2H_2/C_2H_4 and C_2H_6/CH_4 gas ratios



(a) (b)

Figure 4. Dataset distribution (Square Root): (a) average gas ratio values for each transformer fault category, and (b) scatter plot illustrating the relationship between the C_2H_2/C_2H_4 and C_2H_6/CH_4 gas ratios

2.2. Machine learning models

This study employs four supervised machine learning models to classify transformer faults based on DGA gas ratios (e.g., CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/CH_4, C_2H_4/C_2H_6): RF [15], [16], SVM [17], gradient-boosted trees (GBT) [18], and a hybrid RF-GBT model. These models were selected for their robustness, ability to handle complex decision boundaries, and effectiveness with noisy or preprocessed data (raw, min-max, logarithmic, and square root transformations). SVM with an RBF kernel is well-suited for capturing complex nonlinear decision boundaries in relatively small datasets. RF is known for its robustness to noise and its ability to reduce overfitting through ensemble averaging. GBT, on the other hand, provides high predictive accuracy by iteratively correcting previous errors, making it effective for capturing subtle patterns in the data. The hybrid RF–GBT model is proposed to leverage the stability of RF and the high accuracy of GBT. The hyperparameters of each model were selected based on a combination of empirical tuning and commonly adopted configurations reported in the literature. For the SVM model, the RBF kernel was chosen due to its ability to model nonlinear relationships, while the regularization parameter C = 1 ensures a balance between margin maximization and classification error. The kernel width parameter σ = 2.2 was selected empirically to control the influence of individual training samples. For the RF model, 100 trees were used to ensure sufficient diversity while maintaining computational efficiency. The maximum tree depth was limited to 10 to prevent overfitting, and a minimum node size of 2 was selected to allow adequate model flexibility. The Gini index was used as the split criterion due to its effectiveness in classification tasks. For the GBT model, a large number of trees (10,000) combined with a learning rate of 0.1 was used to ensure gradual learning and improved generalization. The tree depth was restricted to 2 to reduce model complexity and avoid overfitting, which is a common strategy in boosting algorithms. Each model was trained on 80% of the dataset and tested on 20%. The machine learning models were implemented using the KNIME Analytics Platform with the hyperparameters summarized in Table 2.

Table 2. Hyperparameter settings for the classification models

Model	Hyperparameters
SVM	Kernel = RBF (Gaussian), overlapping penalty (c) = 1, and sigma (σ) = 2.2
RF	Number of trees = 100, split criterion = Gini index, maximum tree depth = 10, and minimum node size = 2
GBDT	Number of trees = 10000, learning rate = 0.1, and maximum tree depth = 2

2.2.1. RF-GBT fusion

The outputs of the RF and GBT classifiers were combined in KNIME using the "Prediction Fusion" node. This fusion approach makes use of RF's robustness and interpretability as well as GBT's high predictive accuracy and global optimization capabilities. A weighted probability averaging technique was used in place of equal weighting, which reflected GBT's better individual performance during model evaluation. Specifically, the fused prediction was computed as (4).

$$P_{fusion}(y,x) = \frac{2}{3}\ P_{RF}(y,x) + \frac{1}{3}P_{GBT}(y,x) \tag{4}$$

The final class label was determined by (5).

$$\hat{y} = arg(max(P_{fusion}(y,x))) \tag{5}$$

The weighting coefficients defined in (4), namely ($\frac{2}{3}$) for the RF model and ($\frac{1}{3}$) for the GBT model, were determined through an empirical grid-search procedure during the validation phase by systematically evaluating multiple weighting combinations. Several weighting combinations were tested in order to assess their impact on the overall performance of the hybrid model. The selected configuration provided the best trade-off between accuracy and robustness. The higher weight assigned to RF is justified by its ability to produce stable predictions and reduce variance, which improves model generalization. In contrast, GBT, although slightly less stable, contributes fine-grained discriminative power through its sequential learning process. With the goal of lowering classification errors and boosting the robustness of the diagnostic system, this weighted fusion technique highlights the more accurate GBT predictions while preserving the ensemble advantages of RF.

2.3. Experimental approach

The entire experimental procedure for transformer fault diagnosis using DGA is illustrated in Figure 5. Gas data are first extracted from DGA reports, and diagnostic ratios, particularly those recommended by the Modified Rogers' Ratio Method, are computed. The preprocessing phase includes the removal of missing values and the application of four data transformation techniques: raw data (used as a baseline), logarithmic transformation, square root transformation, and min-max normalization.

The KNIME environment is then used to develop machine learning models, such as SVM, RF, GBT, and a hybrid RF-GBT fusion model. Eighty percent of the dataset is used to train each model, and the remaining twenty percent is used for testing. Five common metrics accuracy, precision, recall, F1-score, and Cohen's kappa are used to assess the model's performance. This process makes it possible to compare the models in great detail and shows how various preprocessing techniques affect the accuracy and dependability of the final diagnosis.

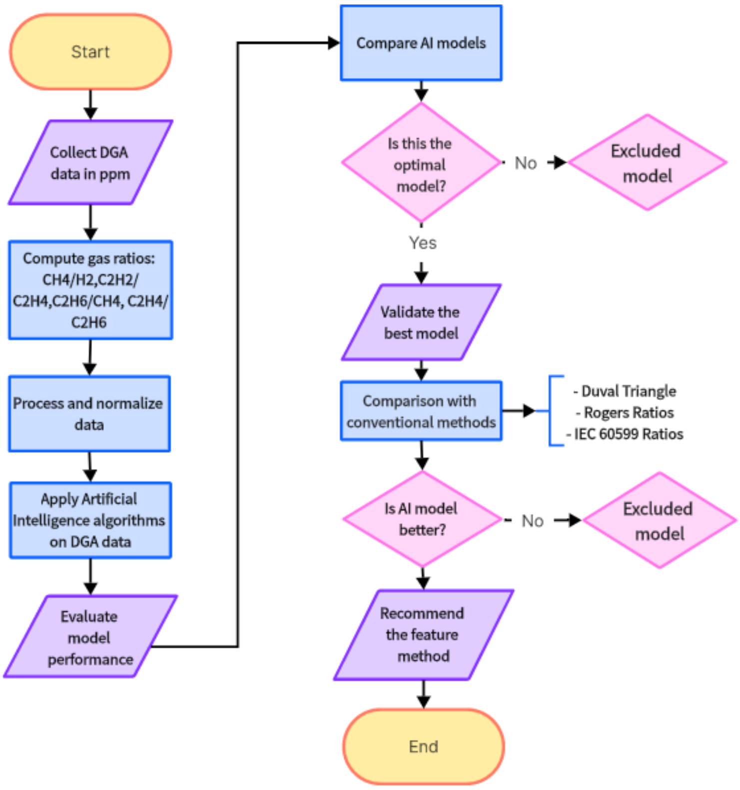


Figure 5. Experimental workflow for transformer fault diagnosis

3. RESULTS AND DISCUSSION

The results of the classification models applied to DGA-based gas ratio data under the four preprocessing strategies of raw, logarithmic, square root, and min-max normalization are shown in this section. SVM with RBF kernel, RF, GBT, and a hybrid RF-GBT fusion model are among the models that have been evaluated. Eighty percent of the dataset was used to train each model, and the remaining twenty percent was used for testing. Model performance was evaluated using five key metrics: accuracy, precision, recall, F1-score, and Cohen's kappa coefficient. The results are summarized in Table 3.

The RF classifier delivered strong and consistent performance across all preprocessing strategies. With raw data, it achieved 79.5% accuracy and a 79.63% F1-score. The logarithmic transformation led to the best results for RF, increasing accuracy to 83.3% and Cohen's kappa to 74.5%. Both square root and min-max transformations also provided performance gains, indicating that RF benefits from reduced skewness and feature scaling. This behavior can be explained by the ensemble nature of RF based on bagging, which reduces variance and makes the model less sensitive to noise and data distribution. By reducing skewness, logarithmic and square root transformations allow more balanced tree splits, improving classification boundaries. GBT showed stable but slightly lower performance than RF across all transformations. Accuracy remained around 77.7% for both raw and min-max scaled data. Logarithmic and square root transformations did not improve results and even caused slight drops in accuracy and kappa. This suggests that boosting algorithms may not benefit from external feature transformations. This can be explained by the sequential nature of GBT, which already performs internal error correction. As a result, external transformations may distort patterns that the model is already capable of capturing, especially in highly non-linear and noisy DGA data. SVM was the most sensitive to preprocessing. It performed moderately on raw data (68.5% accuracy), but dramatically underperformed with min-max normalization, reaching only 33.8% accuracy and a kappa of 1%. However, with logarithmic transformation, SVM achieved 70.8% accuracy and a more balanced F1-score of 69.8%. These results highlight the strong sensitivity of SVM to preprocessing techniques. This is mainly due to its reliance on distance-based kernel functions (RBF), where feature scaling directly impacts the geometry of the decision space. Logarithmic transformation improves performance by stabilizing variance and reducing extreme values, making the data more suitable for kernel-based separation.

The hybrid ensemble combining RF and GBT outperformed all individual models. On raw data, it already achieved 90.3% accuracy, with a Cohen's kappa of 81.5%. The best results were obtained with the logarithmic transformation, peaking at 94.93% accuracy and 94.81% F1-score, along with an excellent Cohen's kappa of 92.37%, indicating near-perfect agreement with true labels. The fusion benefits from the complementary strengths of RF (low variance) and GBT (bias correction), resulting in superior generalization and robustness across all preprocessing techniques. This confirms that combining models with different learning mechanisms is particularly effective for handling the complexity and variability of DGA data.

This performance is further illustrated by the confusion matrix presented in Table 4, which provides a detailed view of the classification results for each fault category. Out of 138 test samples, 131 were correctly classified, corresponding to a very low error rate of approximately 5%. As shown in Table 4, the model achieved excellent classification accuracy for the normal (N) and electrical (E) classes, with 48 out of 49 normal samples and 45 out of 48 electrical fault samples correctly identified.

The thermal (T) class also shows strong performance, with 38 correctly classified samples out of 41. However, most misclassifications occur between thermal faults and normal states. Specifically, Table 4 indicates that 2 thermal fault cases were incorrectly predicted as normal, and 1 normal case was misclassified as thermal. This confusion can be attributed to moderate thermal degradation conditions, where gas ratios remain close to those observed in healthy transformers. Additionally, a small number of misclassifications can be observed between electrical and thermal faults (Table 4), suggesting partial overlap in certain borderline gas patterns. This phenomenon reflects the physical reality of transformer behavior, where fault signatures are not always sharply separable, especially in early fault stages. From an application perspective, these results are significant. The low number of misclassifications and the high accuracy for critical fault classes demonstrate that the proposed RF–GBT model is reliable for real-world transformer monitoring. In particular, minimizing confusion between normal and faulty conditions is essential for reducing false alarms and avoiding unnecessary maintenance actions. Overall, Table 4 confirms not only the high global accuracy of the model, but also its strong class-wise performance and practical suitability for DGA-based fault diagnosis.

This methodological choice was made to ensure reliable and balanced learning during model training. The collected DGA reports presented heterogeneous data availability, and detailed subcategories such as T1/T2/T3 or PD/D1/D2 were not systematically reported in industrial records. Including these subclasses would have resulted in significant class imbalance and reduced the statistical robustness of the models. Therefore, the adopted three-class structure provides a practical and generalizable diagnostic framework, capable of supporting condition-based maintenance decisions while maintaining strong predictive stability and interpretability.

Table 3. Comparison of classifier performance according to data transformation techniques

Classifier	Data	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)	Cohen's kappa (%)
RF	Raw data	79.5	79.93	79.53	79.63	71.3
	Min-max	80.5	81.07	80.53	80.7	71.3
	LOG	83.3	83.66	83.3	83.4	74.5
	Square root	82.3	82.67	82.3	82.4	73.5
GBT	Raw data	77.7	78.163	77.63	77.4	66.5
	Min-max	77.7	78.17	77.63	77.4	66.5
	LOG	76.9	77.16	76.86	76.63	65.4
	Square root	76.9	77.17	76.87	76.2	76.63
SVM	Raw data	68.5	72.13	68.4	68.03	52.7
	Min-max	33.8	25.25	34.07	29.85	1
	LOG	70,8	71.43	70.93	69.8	56.2
	Square root	51.5	51.87	51.57	51.27	27.3
RF-GBT	Raw data	90.3	90.66	90.3	90.33	81.5
	Min-max	92.3	92.33	92.3	92.33	83.5
	LOG	94.93	94.88	94.8	94.81	92.37
	Square root	93.8	94.03	93.93	93.93	90.8

Table 4. Confusion matrix for the RF-GBT model with logarithmic transformation

Number of rows: 138	E (Predicted)	N (Predicted)	T (Predicted)
E (Actual)	45	1	2
N (Actual)	0	48	1
T (Actual)	1	2	38

3.1. Comparative analysis of the models

The graph (Figure 6) provides a comparative evaluation of four classifiers: RF, GBT, SVM, and their ensemble combination (RF-GBT), based on five key performance indicators. The classification results based on the final evaluation metrics clearly demonstrate that the RF-GBT ensemble model significantly outperforms all other classifiers. With an accuracy of 92.6%, F1-score of 92.42%, and a Cohen's kappa of 86.8%, it delivers the most reliable and consistent predictions. This hybrid approach leverages the variance control of RF and the bias correction strength of GBT, resulting in superior generalization and fault discrimination. The standalone RF model follows with 81.4% accuracy, 81.53% F1-score, and a Cohen's kappa of 72.7%. These results reflect a robust and stable classifier that handles noisy and variable data well, making it a solid option when model simplicity or interpretability is required.

The GBT perform moderately, with 77.3% accuracy, 77.16% F1-score, and 68.7% on Cohen's kappa. While effective, the model falls short of RF in both precision and reliability, possibly due to overfitting or sensitivity to specific data patterns. The SVM model delivers the weakest results, with accuracy limited to 56.15%, F1-score at 54.24%, and a low Cohen's kappa of 34.3%. This highlights its poor alignment with actual fault classes and suggests limited suitability for DGA-based transformer diagnostics in its current configuration.

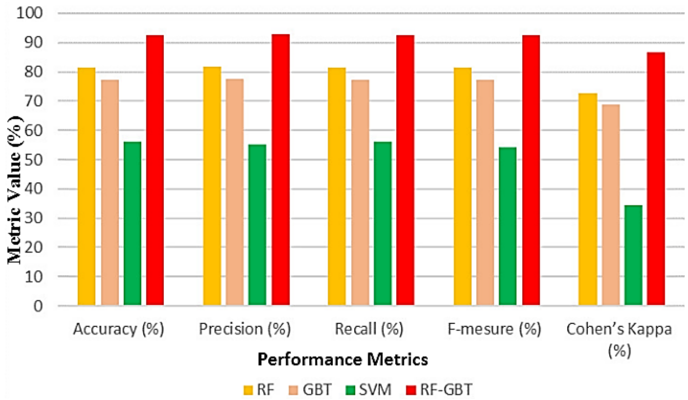


Figure 6. Comparison of precision, recall, F-measure, accuracy, and Cohen scores for each classifier

3.2. Comparison of performance based on data transformations (normalization)

To better isolate the impact of preprocessing methods on the most consistent classifiers, the analysis focused solely on RF, GBT, and their hybrid ensemble (RF-GBT), excluding SVM due to its instability (see Figure 7). Based on average performance across these models, logarithmic transformation yielded the best overall results, with an average accuracy of 84.4%, F1-score of 84.3%, and Cohen's kappa of 80%. It clearly improves the learning process by reducing skewness and stabilizing feature distributions.

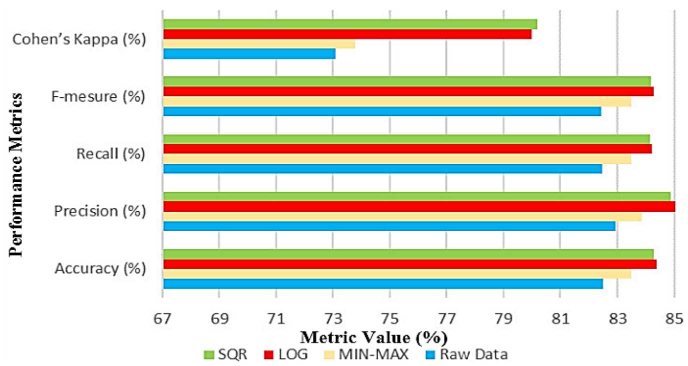


Figure 7. Comparison of average performances according to data transformation techniques

Square root transformation produced almost identical results (accuracy: 84.3%, F1-score: 84.18%, kappa: 80.2%), confirming its suitability for this classification task. The negligible difference suggests both transformations are effective, with square root offering slightly better agreement with true labels. Min-max normalization led to a modest improvement over raw data, with accuracy increasing from 82.5% to 83.5%, and F1-score from 82.45% to 83.48%. However, the impact on Cohen's kappa remains limited (from 73.1% to 73.8%), indicating minor influence on label consistency. Overall, the findings support the use of logarithmic or square root transformation as the most effective preprocessing strategies, while raw and min-max scaled data remain acceptable but less optimal.

3.3. Ablation study: impact of DGA ratios

To evaluate the contribution of different features derived from DGA, an ablation study was conducted. The objective was to analyze how various combinations of gas ratios influence the diagnostic performance of the proposed model. The results of this analysis are summarized in Table 5 and are discussed in detail to highlight the contribution of each feature combination. For this experiment, the fusion model combining GBT and RF was used, as it achieved the best overall performance in the previous experiments. In addition, logarithmic normalization was applied to the input data since it provided the most effective results compared with other normalization techniques. Four feature configurations based on different combinations of DGA ratios were investigated to evaluate their impact on the diagnostic model performance.

The results show that using only the CH_4/H_2 ratio provides a moderate baseline performance with an accuracy of 85.4%. As shown in Table 5, this ratio captures essential information related to thermal faults, but remains insufficient to fully discriminate between different fault types. However, the combination of two ratios slightly decreases the performance, indicating that these two features alone are not sufficient to capture the complexity of transformer fault patterns. This decrease suggests possible redundancy or weak complementarity between the selected ratios, which may introduce noise rather than improving class separability. A significant improvement is observed when three ratios are used, where the accuracy increases to 94.6% and Cohen's kappa reaches 91.9%, demonstrating that the additional ratio contributes meaningful diagnostic information. According to Table 5, this configuration provides a better balance between feature diversity and discriminative power, enabling the model to capture more complex fault signatures. The highest performance is achieved with four ratios, reaching an accuracy of 94.93% and a Cohen's kappa of 92.37%. This result, clearly highlighted in Table 5, confirms that combining multiple complementary gas ratios significantly enhances the model's ability to distinguish between fault types. From a practical perspective, these findings indicate that relying on a single or limited number of gas ratios may lead to incomplete diagnosis, while incorporating multiple ratios improves robustness and reliability in real-world transformer monitoring systems. This confirms that incorporating multiple DGA ratios enhances the model's ability to capture relevant fault characteristics and improves the reliability of transformer fault diagnosis.

Table 5. Ablation study results for different combinations of DGA ratios

Features	Accuracy (%)	Precision (%)	Recall (%)	F-mesure (%)	Cohen's kappa (%)
CH4/H2	85.4	86.13	86.26	85.4	78.2
CH4/H2,C2H2/C2H4	83.1	82.5	83.63	82.66	74.5
CH4/H2,C2H2/C2H4, C2H4/C2H6	94.6	94.2	95.16	94.4	91.9
CH4/H2, C2H2/C2H4, C2H4/C2H6, C2H2/CH4	94.93	94.88	94.8	94.81	92.37

3.4. Comparative evaluation with traditional DGA interpretation methods

To assess the relevance of the proposed hybrid model, a direct comparison was carried out with the most commonly used traditional DGA interpretation methods: the Duval Triangle, the Rogers Ratio Method, and the IEC 60599 ratio method. The defect subcategories were grouped as follows: (T1 + T2 + T3 → T) and (PD + D1 + D2 → E). The comparative results presented in Table 6 show that conventional approaches achieve accuracies between 82% and 86%, while the proposed hybrid RF–GBT model reaches 94.93% overall accuracy and 92.37% Cohen's kappa. As illustrated in Table 6, the performance gap clearly highlights the limitations of traditional ratio-based methods when dealing with complex or borderline fault conditions. These superior performances can be explained by the ability of ensemble models to capture nonlinear and cross-dependencies between dissolved gases, which ratio-based methods cannot represent. In contrast, conventional methods rely on fixed threshold rules and predefined zones, which limits their flexibility and leads to ambiguity when gas ratios fall near decision boundaries. However, it is important to emphasize that traditional methods remain robust, interpretable, and standardized, whereas AI-based models require a carefully calibrated and validated dataset to ensure reliable industrial application.

Table 6. Comparison of the performance of the proposed hybrid RF–GBT model with conventional methods

Method	Accuracy (%)	F1-Score (%)	Cohen's kappa (%)
Duval triangle [19]-[21]	82.41	80.35	76.50
Rogers ratios [22], [23]	84.72	83.00	78.90
IEC 60599 ratios [24], [25]	86.11	85.02	81.20
Proposed (RF–GBT)	94.93	94.81	92.37

3.5. External validation on unseen data

To further assess the generalization capability of the proposed diagnostic system, an external validation phase was conducted using 30 new samples that were completely excluded from the initial training and preprocessing steps. The best-performing configuration, namely, the RF-GBT hybrid model combined with logarithmic transformation, was applied to this independent dataset. The external validation dataset was obtained from transformers located outside Algeria [26], further supporting the generalization capability of the proposed model across different geographical contexts.

The results confirmed the model's predictive reliability, achieving an accuracy of 96.67%, an F1-score of 96.6%, and a Cohen's kappa of 95%, with 29 out of 30 samples correctly classified. As detailed in Table 7, the confusion matrix highlights the distribution of correct and incorrect predictions across the three fault classes. The confusion matrix presented in Table 7 showed perfect classification for all thermal faults and normal state cases, with only one misclassification between an electrical and a thermal fault. More specifically, as shown in Table 7, all 10 thermal fault samples and all 10 normal samples were correctly identified, while only one electrical fault was misclassified as a thermal fault. This near-perfect classification demonstrates the strong generalization capability of the proposed model, confirming that it does not suffer from overfitting despite being trained on a limited dataset. These results demonstrate the robustness of the hybrid model and its ability to maintain high performance on unseen data, reinforcing its suitability for real-world transformer fault diagnosis based on DGA.

Table 7. Confusion matrix for validation set using RF-GBT

Rows of number: 30	E (Predicted)	N (Predicted)	T (Predicted)
E (Actual)	9	0	1
N (Actual)	0	10	0
T (Actual)	0	0	10

4. CONCLUSION

This study evaluated four classification approaches RF, GBT, SVM, and their hybrid ensemble RF-GBT under four preprocessing strategies: raw data, logarithmic scaling, min-max normalization, and square-root transformation. The results clearly show that preprocessing plays a critical role in the performance of DGA-based diagnostic models, particularly for algorithms sensitive to feature scale and data distribution. Among all tested configurations, the RF-GBT ensemble achieved the most stable and highest overall performance, confirming the benefit of combining variance-reducing RF and bias-reducing GBT learners within a unified framework. RF consistently provided robust and stable predictions, while GBT contributed complementary discriminative power through sequential error correction. In contrast, SVM showed limited effectiveness, with weak generalization across preprocessing methods and comparatively lower accuracy.

From a practical perspective, these findings have direct implications for power system reliability. Accurate and stable transformer fault diagnosis based on DGA enables earlier detection of incipient failures, reduces the risk of unexpected outages, and supports maintenance planning strategies. The proposed framework can be integrated into online transformer monitoring and SCADA systems to provide real-time fault diagnosis and support predictive maintenance. The superiority of the RF-GBT ensemble suggests that hybrid machine learning models can significantly enhance decision support systems for condition monitoring of power transformers.

Regarding preprocessing, logarithmic and square-root transformations proved to be the most effective in improving classification performance and metric consistency. Min-max normalization showed limited added value, while raw data remained acceptable but less optimal in most cases. These observations highlight the importance of designing appropriate preprocessing pipelines tailored to the statistical nature of DGA data.

Overall, this work contributes to improving the reliability and interpretability of data-driven transformer diagnostics by demonstrating the effectiveness of ensemble learning combined with suitable data transformations. Future research should focus on more advanced hybrid architectures, optimization techniques, and dynamic feature extraction from time-dependent DGA signals. The integration of time-series modeling and real-time monitoring systems is also essential to further enhance predictive accuracy, fault detection speed, and operational robustness in real-world power grid applications.

ACKNOWLEDGMENTS

The authors thank DGRSDT of the Ministry of Higher Education and Scientific Research, Algeria.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Fatima Zohra Boudjella	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Souhila Boudjella	✓	✓	✓	✓	✓	✓			✓	✓	✓			
Nasiru Yahaya Ahmed	✓	✓	✓	✓	✓			✓		✓	✓			
Hazlee Azil Illias	✓	✓		✓	✓					✓		✓	✓	
Smail Latifa		✓			✓	✓				✓				
Reriballah Hafidha		✓			✓	✓				✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

[1] N. Y. Ahmed, H. A. Illias, H. Mokhlis, D. Fahmi, M. T. Bin Mohd Khairi, and N. Abdullah, "Mitigation of VFTO from a power transformer bushing in a 275 kV gas insulated substation," in *Proceedings - International Seminar on Intelligent Technology and its Applications, ISITIA*, Jul. 2024, no. 2024, pp. 244–249, doi: 10.1109/ISITIA63062.2024.10668028.

[2] G. S. Sarma, B. R. Reddy, P. Nirgude, and P. V. Naidu, "A long short-term memory based prediction model for transformer fault diagnosis using dissolved gas analysis with digital twin technology," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 13, no. 2, p. 1266, Jun. 2022, doi: 10.11591/ijpeds.v13.i2.pp1266-1276.

[3] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 2, pp. 1239–1248, 2017, doi: 10.1109/TDEI.2017.005959.

[4] 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*, vol. 12, pp. 27794–27809, 2024, doi: 10.1109/ACCESS.2024.3366905.

[5] H.-C. Sun, Y.-C. Huang, and C.-M. Huang, "A review of dissolved gas analysis in power transformers," *Energy Procedia*, vol. 14, pp. 1220–1225, 2012, doi: 10.1016/j.egypro.2011.12.1079.

[6] A. G. Saleh, A. E. Azzam, G. Attiya, and E. Ibrahim, "Detecting incipient faults in power transformers through hybrid model of DGA and machine learning," *Electric Power Systems Research*, vol. 256, p. 112877, Jul. 2026, doi: 10.1016/j.epsr.2026.112877.

[7] A. M. H. Saeed, D. Wang, H. A. M. Alnedhari, K. Mei, and J. Wang, "A survey of machine learning and deep learning based DGA detection techniques," 2022, pp. 133–143, doi: 10.1007/978-3-030-97774-0_12.

[8] A. Dhini, I. Surjandari, A. Faqih, and B. Kusumoputro, "Intelligent fault diagnosis for power transformer based on DGA data using support vector machine (SVM)," in *2018 3rd International Conference on System Reliability and Safety (ICSRS)*, Nov. 2018, pp. 294–298, doi: 10.1109/ICSRS.2018.8688719.

[9] M.-C. Niţu, A.-M. Aciu, M. Nicola, and C.-I. Nicola, "Transformer fault detection using DGA based on three ratio technique and random forest algorithm," in *2024 International Conference on Applied and Theoretical Electricity (ICATE)*, Oct. 2024, pp. 1–6, doi: 10.1109/ICATE62934.2024.10748635.

[10] G. Xu, M. Zhang, W. Chen, and Z. Wang, "Transformer fault diagnosis utilizing feature extraction and ensemble learning model," *Information*, vol. 15, no. 9, p. 561, Sep. 2024, doi: 10.3390/info15090561.

[11] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: experimental evaluation," *Information Sciences*, vol. 513, pp. 429–441, Mar. 2020, doi: 10.1016/j.ins.2019.11.004.

[12] S. S. M. Ghoneim and I. B. M. Taha, "A new approach of DGA interpretation technique for transformer fault diagnosis," *International Journal of Electrical Power & Energy Systems*, vol. 81, pp. 265–274, Oct. 2016, doi: 10.1016/j.ijepes.2016.02.018.

[13] I. B. M. Taha, S. S. M. Ghoneim, and A. S. A. Duaywah, "Refining DGA methods of IEC code and rogers four ratios for transformer fault diagnosis," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Jul. 2016, pp. 1–5, doi: 10.1109/PESGM.2016.7741157.

[14] H. Henderi, "Comparison of min-max normalization and Z-Score normalization in the K-nearest neighbor (kNN) algorithm to test the accuracy of types of breast cancer," *IJIIS: International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 13–20, Mar. 2021, doi: 10.47738/ijiis.v4i1.73.

[15] X. D. Hoang and X. H. Vu, "An improved model for detecting DGA botnets using random forest algorithm," *Information Security Journal: A Global Perspective*, vol. 31, no. 4, pp. 441–450, Jul. 2022, doi: 10.1080/19393555.2021.1934198.

[16] N. Syam and R. Kaul, "Random forest, bagging, and boosting of decision trees," in *Machine Learning and Artificial Intelligence in Marketing and Sales*, Emerald Publishing Limited, 2021, pp. 139–182, doi: 10.1108/978-1-80043-880-420211006.
[17] Z. B. Sahri and R. B. Yusof, "Support vector machine-based fault diagnosis of power transformer using k nearest-neighbor imputed DGA dataset," *Journal of Computer and Communications*, vol. 02, no. 09, pp. 22–31, 2014, doi: 10.4236/jcc.2014.29004.
[18] A. I. Khalyasmaa, M. D. Senyuk, and S. A. Eroshenko, "Analysis of the state of high-voltage current transformers based on gradient boosting on decision trees," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2154–2163, 2021, doi: 10.1109/TPWRD.2020.3021702.
[19] K. A. Nguyen, H. H. Le, B. T. Phung, and H. V. Tran, "Enhancing data preprocessing layer for power transformer fault diagnosis using DGA combining fuzzy logic and duval triangle 1," in *2025 8th International Conference on Circuits, Systems and Simulation (ICCSS)*, May 2025, pp. 94–98, doi: 10.1109/ICCSS65911.2025.11081751.
[20] A. Muhtar and I. A. Umar, "DGA-Duval triangle analysis for early thermal fault diagnosis of transformer oil," E*MITOR: Jurnal Teknik Elektro*, vol. 25, no. 3, pp. 279–286, 2024.
[21] S. An and D. Ahn, "Hybrid AI model for predicting transformer failure types and diagnosing anomalies using Duval characteristics," in *2026 International Conference on Electronics, Information, and Communication (ICEIC)*, Jan. 2026, pp. 1–3, doi: 10.1109/ICEIC69189.2026.11386317.
[22] S. A. M. Abdelwahab, I. B. M. Taha, R. Fahim, and S. S. M. Ghoneim, "Transformer fault diagnose intelligent system based on DGA methods," *Scientific Reports*, vol. 15, no. 1, p. 8263, Mar. 2025, doi: 10.1038/s41598-024-78293-7.
[23] V. Bargate, A. Yadav, and S. Gupta, "Analysis of transformer defect using different dissolved gas analysis techniques: a comparative study," *Journal of The Institution of Engineers (India): Series B*, vol. 106, no. 5, pp. 1385–1399, Oct. 2025, doi: 10.1007/s40031-024-01156-2.
[24] N. Bäcker, S. Selzer, M. Zdrallek, A. Babizki, and K. Lindl, "Extended analysis of power transformer conditions in the DACH region: comparative study of typical gas levels and formation rates," *IET Conference Proceedings*, vol. 2025, no. 14, pp. 483–486, Dec. 2025, doi: 10.1049/icp.2025.1527.
[25] A. Rangel Bessa, J. Farias Fardin, P. Marques Ciarelli, and L. Frizera Encarnação, "Conventional dissolved gases analysis in power transformers: review," *Energies*, vol. 16, no. 21, p. 7219, Oct. 2023, doi: 10.3390/en16217219.
[26] I. Katser, "Power transformers FDD and RUL," *Kaggle*, 2024, [Online]. Available: https://www.kaggle.com/datasets/yuriykatser/power-transformers-fdd-and-rul/data.

BIOGRAPHIES OF AUTHORS

Fatima Zohra Boudjella is a lecturer in Electrical Engineering at Belhadj Bouchaib University of Ain Temouchent, Algeria. She received her B.Eng. and M.Eng. degrees in Electrical Engineering from Djillali Liabes University of Sidi Bel Abbès, Algeria, in 2015 and 2017, respectively, and her Ph.D. degree in Electrical Engineering from the same university in 2021. She earned her Ph.D. in Electrotechnics, specializing in Electrical Networks. Her research focuses on the diagnosis and condition monitoring of power transformers using dissolved gas analysis (DGA), as well as on power electronics and static converters. She is also interested in the integration of renewable energy into electrical networks and in applying artificial intelligence and machine learning techniques for the analysis and optimization of electrical systems. She can be contacted at email: fatima.boudjella@univ-temouchent.edu.dz.

Souhila Boudjella was born in Mohammadia Wilayat of Mascara on 12 April 2000. She graduated with a Bachelor of Engineering and Master of Engineering degrees in Electrical Engineering from Djillali Liabes University of Sidi Bel Abbès, Algeria, in 2020 and 2022, respectively. She is a Ph.D. student in Electrical Engineering specializing in Electrical Networks at Mustapha Stambouli University of Mascara since 2022. Her research focuses on the diagnosis and condition monitoring of power transformers using dissolved gas analysis (DGA). She is also interested in the integration of artificial intelligence and machine learning techniques for the analysis and optimization of electrical systems. She can be contacted at email: souhila.boudjella@univ-mascara.dz.

Nasiru Yahaya Ahmed has obtained his Ph.D. degree in 2026 from the Department of Electrical Engineering, University of Malaya. Nasiru does research in transmission lines lightning overvoltage protection, Insulation coordination of transmission lines, and high-voltage insulation and testing. Their current project is the application of AI techniques to improve transmission lines lightning protection performance. He is an Assistant Chief Engineer at NASENI, Nigeria. He can be contacted at email: ahmednasiryahaya@gmail.com.

Hazlee Azil Illias was born in Kuala Lumpur, Malaysia. He received a bachelor's degree in electrical engineering from the Universiti Malaya, Malaysia in 2006. He immediately joined Freescale Semiconductors Malaysia as a product engineer before pursuing his Ph.D. studies in 2008. He obtained a Ph.D. Degree in Electrical Engineering from the University of Southampton, United Kingdom, in 2011. He was a senior lecturer from 2011 to 2016 and an associate professor from 2016 to 2024 at the Department of Electrical Engineering, Universiti Malaya. Currently, he is a professor at the Department of Electrical Engineering, Universiti Malaya. Since joining Universiti Malaya, he has received funding and grants of nearly RM 3 million, published more than 80 journal papers in ISI-indexed publications and 80 refereed conference papers, and successfully supervised the completion of 13 Ph.D. and 6 master's degrees by research and 37 master's degree projects (by coursework) candidates. He is currently an associate editor of IET Science Measurement and Technology (since 2019) and the Head of Universiti Malaya High Voltage Laboratory and Research Group (since 2011). His current management post at Universiti Malaya is the Director of Corporate Ranking Centre (CRC) under the Associate Vice Chancellor (Corporate Strategy). He was the Head of Department of Electrical Engineering from September 2013 to July 2014, the Secretary of IEEE Dielectrics and Electrical Insulation Society (DEIS) Malaysia Chapter from 2015 to 2019, and the Counselor of IEEE Universiti Malaya Student Branch from 2018 to 2024. He is now a Senior Member of IEEE (Institute of Electrical and Electronics Engineers), Fellow of IET (Institution of Engineering and Technology), Corporate Member of IEM (Institution of Engineers Malaysia), a registered Chartered Engineer (C.Eng.), United Kingdom, and a Professional Engineer (P.Eng.), Malaysia. He is an active reviewer for numerous high-quality journals, which include publishers from IEEE, IET, Elsevier and IOP. His main research interests include modelling and measurement of partial discharge phenomena in dielectric insulation materials, condition monitoring, insulation system diagnosis, lightning overvoltage, transmission line modelling, optimization methods and artificial intelligence techniques. He can be contacted at email: h.illias@um.edu.my.

Smail Latifa is an associate professor in the Electrical Engineering Department at the Nour Bachir University of El-Bayadh in Algeria. She received her Ph.D. degree in electrical engineering from Djilali Liabès University, Sidi Bel Abbès, Algeria, in 2022. She is currently a lecturer in the Electrical Engineering Department at the Nour Bachir University of El-Bayadh Centre in Algeria. Her research interests include power grid systems, voltage regulation, renewable energy integration, and flexible AC transmission systems (FACTS). She has been involved in several studies related to improving power quality and enhancing renewable energy contribution to the grid. She can be contacted at email: l.smail@cu-elbauadh.dz.

Reriballah Hafidha is a lecturer in the Electrical Engineering and Automation Department at Ahmed Zabana University of Relizane, Algeria. She received her B.Eng. and M.Eng. degrees in electrical engineering from Djillali Liabes University of Sidi Bel Abbès, Algeria, in 2015 and 2017, respectively, and her Ph.D. degree in electrical engineering from the same university in 2021. Her research interests include renewable energy, voltage regulation, electrical power systems, and FACTS devices. She can be contacted at email: hafidha.reriballah@univ-relizane.dz.