Article citation info:
Boudjella S, Chaouch D. E, Senoussaoui M. E, Mouffak A, Boudjella F. Z. Comparative analysis of raw and logarithmic feature representations for supervised machine learning-based transformer fault diagnosis using dissolved gas analysis, diagnosis, and condition monitoring of hydraulic systems. Diagnostyka. 2026;27(4):2026404. https://doi.org/10.29354/diag/242709.

DIAGNOSTYKA, 2026, Vol. 27, No. 4

e-ISSN 2449-5220
DOI: 10.29354/diag/242709

COMPARATIVE ANALYSIS OF RAW AND LOGARITHMIC FEATURE REPRESENTATIONS FOR SUPERVISED MACHINE LEARNING-BASED TRANSFORMER FAULT DIAGNOSIS USING DISSOLVED GAS ANALYSIS

Souhila BOUDJELLA [1,*], Djamel Eddine CHAOUCH [1], Mohammed Elamine SENOUSSAOUI [1], Adnane MOUFFAK [1], Fatima Zohra BOUDJELLA [2]

[1] LGPCS Laboratory, Faculty of Sciences and Technology, University Mustapha Stambouli of Mascara, Mamounia Road, Mascara, BP 305-29000, Algeria
[2] Engineering and Sustainable Development Laboratory, Department of Electrical Engineering, University of Ain Temouchent, Ain Temouchent, Algeria
* Corresponding author's E-mail : souhila.boudjella@univ-mascara.dz

Abstract

Early detection of transformer faults is essential for ensuring power system reliability, reducing maintenance costs, and preventing unexpected outages. Traditional dissolved gas analysis (DGA)-based diagnostic approaches often exhibit limited accuracy when handling complex fault conditions. To overcome these limitations, this study investigates machine learning techniques for transformer fault diagnosis using DGA data. Several preprocessing methods were evaluated, and logarithmic transformation was identified as the most effective approach for improving classification performance. Four supervised learning algorithms Fine Tree, Optimizable Tree, Boosted Trees, and Optimizable Ensemble were implemented and evaluated in MATLAB. Performance was assessed using 5-fold and 10-fold cross-validation based on accuracy, precision, recall, and F1-score metrics. The results showed that the Optimizable Ensemble model combined with logarithmic transformation and 10-fold cross-validation achieved the highest performance for fault type detection, reaching an accuracy of 82.1%. For fault severity classification, the same model with logarithmic transformation and 5-fold cross-validation produced the best results, with an accuracy of 82.0%. Comparative analysis demonstrated that the proposed approach consistently outperformed conventional diagnostic methods. These findings confirm the effectiveness of optimized ensemble learning with logarithmic preprocessing for transformer fault diagnosis and severity assessment.

Keywords: transformer fault diagnostics, machine learning, data preprocessing, logarithmic transformation, classification learner.

List of Symbols/Acronyms
A–Abnormal;
N–Normal;
E –Electrical Faults;
T–Thermal Faults;
PD–Partial Discharges;
D1–Low-Energy Discharges;
D2–High-Energy Discharges;
T1–Low-Temperature Faults;
T2–Medium-Temperature Faults;
T3–High-Temperature Faults;
TP–True Positive;
FP–False Positive;
FN–False Negative;
P_0–Observed Agreement Proportion (Classical Accuracy);
P_e–Proportion Agree Expected by Chance;
CV5–5-Fold Cross-Validation;
CV10– 10-Fold Cross-Validation;
DGA– Dissolved Gas Analysis;
IEC–International Electrotechnical Commission;
AI–Artificial Intelligence;
KNN– K-Nearest Neighbor;
SVM –Support Vector Machine;
KPCA –Kernel Principal Components Analysis;
ADASYN– Adaptive Synthetic Sampling;
OLTC: –On-Load Tap Changers;
DT –Duval Triangle ;
RR –Rogers Ratios ;
PPM –Partie Par Million.

1. INTRODUCTION

Electrical transformers play an important role in power generation, transmission and distribution networks [1]. They regulate voltage levels at different stages of the power network, making them inevitable for efficient operation and stability of electrical

Received 2026-07-04; Accepted 2026-11-09; Available online 2026-11-14

© 2026 by the Authors. Licensee Polish Society of Technical Diagnostics (Warsaw. Poland). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

infrastructure. Due to their strategic significance, continuous monitoring and proper maintenance are important to minimize faults, reduce operating costs and guarantee service continuity [2]. Transformers are exposed over time to various thermal, electrical and chemical stresses, resulting in internal faults and compromising their operational reliability [3][4].

Dissolved gas analysis (DGA) is one of the most widely used methods for diagnosing transformer problems [5]. DGA relies on identifying and quantifying generated gases as a result of the decomposition of transformer insulation due to thermal and electrical stresses. [6]. Gases-based diagnostic models such as Rogers ratios, IEC 60599 interpretation guidelines [7][8] and the Duval triangle method [9] are recognized as standard methods for analyzing DGA data in the power industry.

Despite their widespread use, conventional DGA interpretation techniques face challenges with complex, overlapping or noisy data patterns [10] [11]. potentially leading to diagnostic errors or unreliable decision making under critical operating conditions [11]. Recent developments in artificial intelligence, especially in machine learning, provide promising solutions to these challenges [12]. These AI-driven approaches utilize advanced classification algorithms and automated data analysis, thereby substantially improving the diagnostic accuracy of power transformers [12] [13].

The use of artificial intelligence (AI) for transformer health diagnosis has been the subject of several studies. For fault classification, the author of [14] proposed a diagnostic model that integrates traditional (DGA) with K-Nearest neighbor (KNN) algorithm. The model's performance was assessed using standard metrics such as F1-score, accuracy, recall and specificity.

In [15], an intelligent framework was developed, combining ADASYN sampling technique, Kernel Principal components analysis (KPCA), and Support Vector Machine (SVM) classifier optimized with the GD-AHBA algorithm. Applied to DGA data, this approach improved the detection of incipient faults in oil-immersed transformers.

The study in [16] employed a machine learning based decision tree to distinguish true faults from false alarms triggered by gases produced in on-load tap changers (OLTC). This method achieved an accuracy of 85.26% on a new DGA database.

Additionally, ensemble methods are gaining popularity, as they combine multiple weak classifiers to create a more robust model. Predictive maintenance issues have been successfully addressed through the application of techniques such as Random Forest, Bagging, and Boosting [15] [17]. These methods improve generalization, mitigate overfitting, and increase diagnostic precision [17] [18]. In addition, recent research has explored the automatic optimization of hyperparameters in these models using Bayesian methods, thereby increasing their overall performance [19]. To address the limitations of conventional DGA-based transformer fault diagnosis, this study presents a systematic comparative framework to investigate the influence of DGA feature representation and data preprocessing on the performance of supervised machine learning models. The main contribution lies in evaluating both the use of Rogers ratio features as an alternative to raw dissolved gas concentrations expressed in ppm and the effect of logarithmic transformation on classification performance. Unlike studies that primarily focus on classifier development or optimization, this work emphasizes the role of feature representation in improving fault discrimination. Four supervised learning algorithms, namely Fine Tree, Optimizable Tree, Boosted Trees, and Optimizable Ensemble, are evaluated using raw and logarithmically transformed data under both five-fold and ten-fold cross-validation. The models are assessed using accuracy, precision, recall, and F1-score to identify the most effective combinations of feature representation, preprocessing strategy, and classifier. By providing a consistent comparative evaluation, this study offers practical insights into the selection and preprocessing of DGA features for machine-learning-based power transformer fault diagnosis.

2. OVERVIEW OF MATERIAL AND METHODS

This study relies on a methodical approach combining dissolved gas analysis (DGA) with advanced data processing and machine learning techniques for the diagnosis of transformer faults. The initial data, expressed in ppm, were enriched by calculating ratios derived from Rogers' method [20], then subjected to preprocessing including cleaning missing or aberrant values and a logarithmic transformation to reduce the impact of extreme values. After data preparation, the most relevant variables were selected using an adapted selection method, and the classification models were trained and optimized in a specialized development environment. The performance of the models was then evaluated using standard metrics to measure their ability to effectively discriminate between the different classes of defects. This study relies on a methodical approach combining dissolved gas analysis (DGA) with advanced data processing and machine learning techniques for the diagnosis of transformer faults. The initial data, expressed in ppm, were enriched by calculating ratios derived from Rogers' method [20], then subjected to preprocessing including cleaning missing or aberrant values and a logarithmic transformation to reduce the impact of extreme values. After data preparation, the most relevant variables were selected using an adapted selection method, and the classification models were trained and optimized in a specialized development environment. The performance of the models was then evaluated using standard metrics to measure

their ability to effectively discriminate between the different classes of defects, the overall methodology adopted in this study is illustrated in Fig.1.

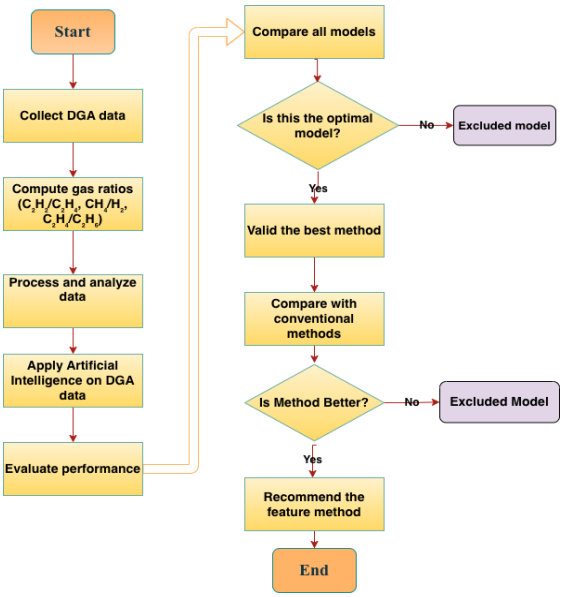


Fig. 1. Experimental Workflow for Transformer Fault Diagnosis

2.1. Data description

In this study, a dataset comprising 710 Dissolved Gas Analysis (DGA) records collected from insulating oil samples of power transformers was used [21]. The gas concentration values reported in this study are actual DGA measurements, expressed in ppm, obtained from oil samples collected from in-service power transformers. Each record corresponds to an individual oil sample analysis.

During the data preprocessing stage, two duplicate records were identified and removed from the dataset. Consequently, the final dataset used in this study consists of 708 valid records.

The dataset includes 216 normal operating conditions (Class N) and 492 fault conditions (Class A). The fault cases are further divided into two major categories: electrical faults (276 cases) and thermal faults (216 cases). The electrical fault category consists of 74 Partial Discharges (PD), 91 Low-Energy Discharges (D1), and 111 High-Energy Discharges (D2), whereas the thermal fault category includes 60 Low-Temperature Faults (T1), 54 Medium-Temperature Faults (T2), and 102 High-Temperature Faults (T3). This hierarchical structure supports a three-level diagnostic framework, including normal/fault discrimination, fault category identification (electrical or thermal), and fault subtype classification.

The input variables correspond to the concentrations (ppm) of five key dissolved gases: hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), and ethane (C_2H_6). The thermal fault subclasses are defined according to their operating temperatures: T1 (below 300 °C), T2 (300–700 °C), and T3 (above 700 °C), while PD, D1, and D2 represent electrical faults of increasing discharge severity (See Fig. 2). Representative concentration ranges for each transformer condition are summarized in Table 1.

Table 1. Representative Gas Concentrations **(ppm)** for Each Transformer Condition

H_2	**CH_4**	**C_2H_2**	**C_2H_4**	**C_2H_6**	**State**	**Number**
523	44	1,5	1,50	17	E	276
2526,3	130,55	1,0	1,53	15		
291	1260	232	6	9	T	216
39111	4291	627	6	12311		
0,1	1,1	2,02	3,1	4,02	N	216
2	2,5	0,5	0	5		

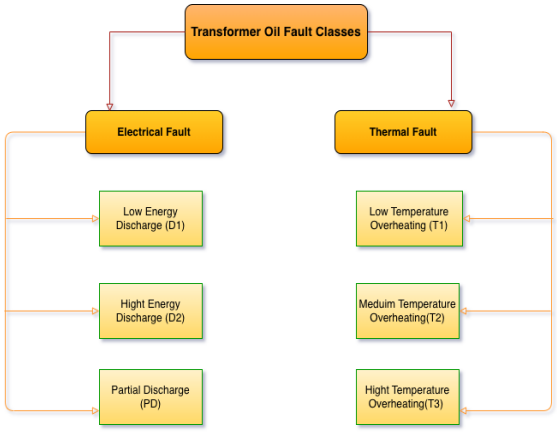


Fig. 2. DGA Fault Classes from Transformer Oil

2.2. Feature Extraction

To incorporate expert knowledge from conventional Dissolved Gas Analysis (DGA) diagnostic techniques, ratio-based diagnostic features were computed. In this study, the raw dissolved gas concentration measurements were complemented by the calculation of four diagnostic ratios according to the modified Rogers method: CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, and C_2H_6/CH_4 (See Table 2) [20]. The five raw dissolved-gas concentrations constitute the original DGA measurements, while the four

calculated Rogers ratios were used exclusively as input features for the machine learning models.
These ratios reduce the influence of absolute gas concentration variations while emphasizing the characteristic signatures of different fault types. By normalizing the relationships between the gases, they enhance the discrimination between thermal and electrical faults and improve the robustness of the diagnostic process against operating condition variations and differences in data scale.

Table 2. Ratio-Base Extracted DGA Features

Method	Gas Ratio	Label
Modified Rogers' Ratio Method	CH_4/H_2	R1
	C_2H_2/C_2H_4	R2
	C_2H_4/C_2H_6	R3
	C_2H_6/CH_4	R4

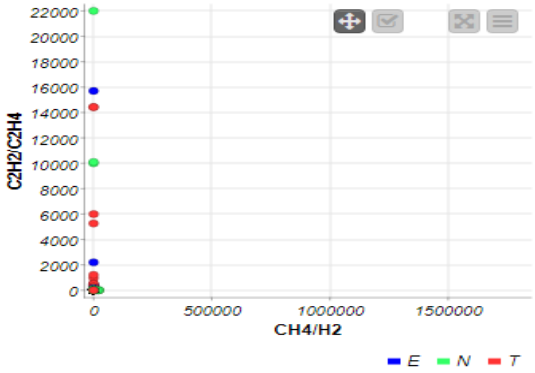
2.3. **Data Cleaning and Pre-processing**

Pre-processing is an essential step to ensure the reliability and performance of diagnostic models. The dissolved gas measurements (DGA) from the 710 cases studied may contain missing, inconsistent, or substantially overlapping values between classes, due to the diversity of sources and sampling conditions.
To ensure the quality and consistency of the dataset, several preprocessing steps were performed prior to model training.

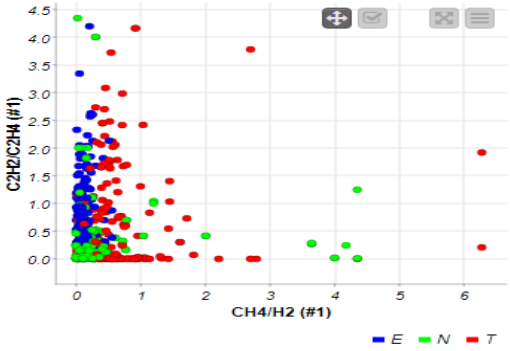
- **Calculation of derived DGA features** from the corresponding gas ratios.
- **Data verification and cleaning** to identify inconsistent or invalid observations.
- **Treatment of missing or invalid values** using the specified imputation procedure.
- **Removal of duplicate records and repeated observations** to avoid potential bias during model training.
- **Application of logarithmic transformation** When the LOG configuration was evaluated, the logarithmic transformation Xlog=log10(x+1) was applied to the calculated gas ratios derived from the dissolved-gas concentrations expressed in ppm.
- **mRMR feature selection** to identify the most relevant and least redundant features for classification.

First, missing or invalid gas measurements were handled by imputing the corresponding values with the median computed from the training dataset. Next, duplicate records and repeated gas patterns were removed to prevent potential bias during the learning process. Finally, when the LOG configuration was evaluated, a logarithmic transformation was applied to the calculated gas ratios derived from the dissolved gas concentrations to reduce the influence of extreme observations and mitigate the effects of the highly skewed distributions inherent in DGA data.
This transformation compressed the dynamic range of the measurements, improved the interpretability of the data distribution, enhanced the separation between fault classes, and contributed to better classification performance.

$$x_{log} = \log_{10}(x+1) \quad (1)$$



(A) Scatter Plot of CH_4/H_2 vs. C_2H_2/C_2H_4 Ratios – Raw Data

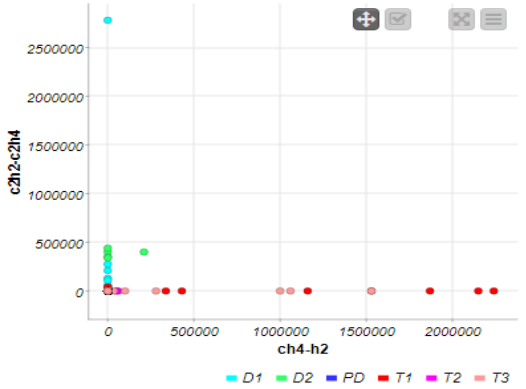


(B) Scatter Plot of CH_4/H_2 vs. C_2H_2/C_2H_4 Ratios – Log-Transformed Data

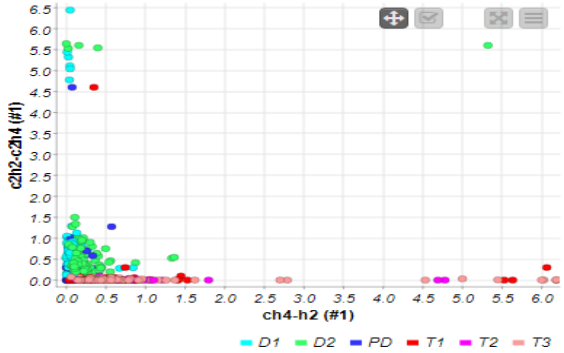
Fig. 3. Dataset distribution for transformer fault detection

The scatter plot constructed from the CH_4/H_2 ratios (X-axis) and C_2H_2/C_2H_4 ratios (Y-axis), illustrates the distribution of defects in the feature space. In the raw version (See Fig. 3A), without transformation, the classes E (electrical faults), T (thermal faults), and N (normal) represented in blue, red, and green, respectively show a marked overlap, particularly in the ranges of low values. The lack of clear clustering and the presence of outliers reduce the interpretability of the plot, making the visual interpretation of the classes difficult. The logarithmic transformation improves the data distribution by mitigating the influence of extreme values and revealing more distinct clustering structures (See Fig. 3B). Notably, electrical faults predominantly align along the y-axis, indicating a strong correlation with the C_2H_2/C_2H_4 ratio, while thermal faults are positioned in the upper area of the graph, marked by the simultaneous increase of both ratios. The normal state is located in the lower left corner, suggesting low

concentrations for these indicators. These graphical improvements justify the use of logarithmic preprocessing as a preliminary step to any classification. A similar analysis was conducted on the severity levels (T1, T2, T3, D1, D2, PD), where the raw cloud shows an even more pronounced overlap (See Fig. 4A). There too, the logarithmic transformation allows for better visual separation of the subclasses, thereby reinforcing the interest in rigorous preprocessing for reliable interpretation (See Fig. 4B).



(A) Scatter Plot of CH_4/H_2 vs. C_2H_2/C_2H_4 Ratios – Raw Data



(B) Scatter Plot of CH_4/H_2 vs. C_2H_2/C_2H_4 Ratios – Log-Transformed Data

Fig. 4. Dataset distribution for transformer fault severity

2.4. Model Development and Training

Classification algorithms are an essential element for building a reliable model. An inappropriate choice can lead to erroneous decisions and reduce the overall performance of the diagnostic system. This study is not limited to a single method but offers a comparative analysis of several approaches from the decision tree family and ensemble methods (See Fig. 5). These techniques were selected because they represent different supervised learning strategies and are widely recognized for their effectiveness in processing complex and heterogeneous data. In this work, only supervised learning, where the algorithm learns from labeled examples, was considered. The performance of the models was then evaluated by measuring their ability to correctly predict new observations.

Grid Search was used to optimize the hyperparameters of all the classification models considered in this study. For each classifier, a predefined set of candidate values was specified for the relevant hyperparameters. The different combinations were systematically evaluated according to the selected validation criterion to identify the configuration that achieved the best classification performance. The final hyperparameter configurations obtained for each model are reported in Table 3.

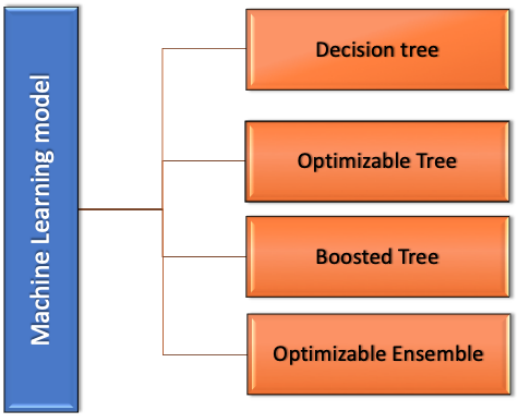


Fig. 5. Classification Algorithms Used for Transformer Fault Diagnosis

Table 3. Hyperparameters setting for the Classification Models

Model	Hyperparameters
Fine Tree	Split criterion=Gini´s diversity index MaxNumSplits=100 MinleafSize=4
Optimizable Tree	Split criterion=Gini´s diversity index MaxNumSplits=39
Boosted Tree	Ensemble method= Adaboost Number of learners=30 MaxNumSplits=20 Learning rate=0.1
Optimizable Ensemble	Ensemble method= Bag Number of Learners=98 Learning rate=0.1

2.5. Model validation

In this study, the performance of classification algorithms Fine Tree, Optimizable Tree, Boosted Trees, and Optimizable Ensemble was evaluated using two cross-validation methods: 10-fold and 5-fold. Both 5-fold and 10-fold cross-validation were performed using MATLAB Classification Learner. For classification tasks, the data were partitioned into folds while maintaining approximately the same proportion of response classes in each fold as in the training dataset.

The principle of cross-validation (k-fold cross-Validation) involves dividing the entire dataset into k subsets of equal size. At each iteration, (k − 1) subsets are used for training, while the remaining subset is used for testing. This process is repeated k times, and the average of the obtained results provides a reliable estimate of the overall performance. The use of 10-fold and 5-fold configurations allows for a comparison of the impact of the number of divisions on the accuracy and robustness of the models, while taking into account the trade-off between computation time and evaluation quality.

The rigorous evaluation of classification models is an essential step to ensure their reliability and relevance in a diagnostic context. In this study, several performance metrics were employed to evaluate the quality of predictions produced by the Fine Tree, Optimizable Tree, Boosted Trees, and Optimizable Ensemble models, considering the multi-class nature of the problem [26].

For the multi-class classification tasks, Precision, Recall, and F1-score were calculated using macro-averaging, where the metrics were computed independently for each class and then averaged with equal weight across all classes.

A) *Accuracy*

Accuracy represents the proportion of correctly classified examples relative to the total number of samples. It is a global measure of performance, often used as the main indicator. However, in unbalanced datasets, it can be misleading because it does not take into account the distribution between classes.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

B) *Precision*

Precision evaluates the model's ability not to incorrectly classify a sample as belonging to a class when it should not. It is defined as the ratio between true positives and the total number of positive predictions:

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

High precision indicates few false positives, which is essential when the consequences of a false diagnosis are critical.

C) *Recall*

Recall measures the model's ability to detect all truly positive samples. It is particularly important in the context of diagnosis, where failing to detect a defect (false negative) can have serious consequences:

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

D) *F1-Score (F- Measure)*

The F1-Score combines precision and recall into a single harmonized metric, which allows for a balanced evaluation when one of the two is significantly weaker than the other:

$$F1 - score = 2.\frac{Precision.Recall}{Precision+Recall} \quad (5)$$

This metric is particularly useful for assessing the performance of a classifier in contexts where the classes are imbalanced.

E) *Cohen kappa*

Cohen's kappa is a statistical measure of the agreement between two raters or classification models when they assign categories to a set of items. Unlike simple accuracy, Cohen's kappa takes into account the agreement that may occur by chance, making it a more reliable measure in situations where the classes are unbalanced.

$$k = \frac{P_0-P_e}{1-P_e} \quad (6)$$

- P_0: Observed Agreement Proportion (Classical Accuracy)
- P_e : Proportion Agree Expected by Chance

2.6. Feature selection

For the selection of relevant variables, the MRMR (Minimum Redundancy Maximum Relevance) algorithm was applied to rank the variables according to their importance in predicting defects. The mRMR feature selection procedure was performed separately for each classification task. A distinct mRMR selection was therefore applied for fault detection and fault severity classification, ensuring that the selected features were specifically adapted to the objective of each task.

This method allows for the retention of features that have the greatest relevance to the target variable, while minimizing redundancy among them, which enhances the quality of the model. Four gas ratios were evaluated: CH_4/H_2, C_2H_4/C_2H_6, C_2H_2/C_2H_4, and C_2H_6/CH_4. The results show that the CH_4/H_2 ratio is the most informative variable, with an MRMR score of 0.2346, followed by C_2H_4 /C_2H_6 (0.1216) and C_2H_2/C_2H_4 (0.0657). On the other hand, C_2H_6/CH_4 has negligible importance (0.0033) and was therefore excluded. The first three variables were automatically selected according to using the "Select highest ranked features" option, with the number of variables set to 3. This selection step optimizes the model's performance by retaining only the most discriminating features. To further interpret the contribution of the selected features to the classification process, SHAP analysis was performed, and the resulting feature importance rankings are presented in Fig. 6.

SHAP analysis was performed using the Optimizable Ensemble classifier for both fault detection and fault severity classification, using the previously described dataset with the calculated Rogers ratios as input features. For the LOG configuration, the logarithmic transformation was applied to the calculated gas ratios before the SHAP analysis.

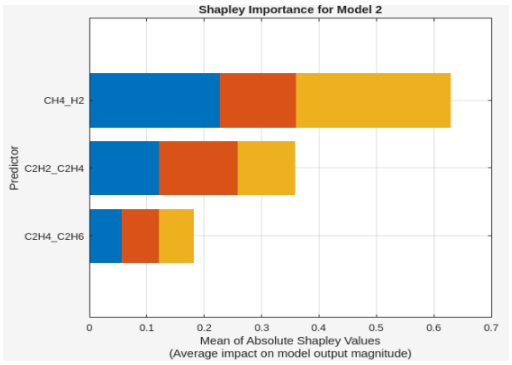
The interpretation of SHAP values allows for the quantification of the relative importance of different dissolved gas ratios in the classification of the state and severity of faults in power transformers. For the state classification model, the results show that the CH_4/H_2 ratio is the most influential predictor, particularly for identifying thermal faults (T). This ratio, an indicator of oil overheating phenomena, plays an essential role in the early detection of temperature-related abnormal conditions. The C_2H_2/C_2H_4 ratio, on the other hand, is strongly correlated with class E (electrical faults), particularly internal arcs, due to the marked generation of acetylene during such phenomena.

The C_2H_4/C_2H_6 ratio is of more moderate importance, primarily serving to refine the recognition of thermal faults according to their degree of evolution.

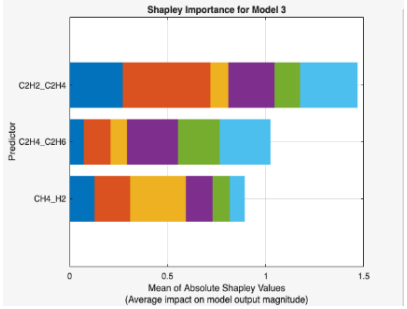
Regarding the fault severity classification model, the SHAP results indicate that C_2H_2/C_2H_4 is the most important ratio, effectively differentiating between electrical faults (D1, D2), partial discharges (PD), and severe thermal faults (T2, T3). This ratio thus appears as a good marker of the increasing intensity of the faults. The C_2H_4/C_2H_6 ratio proves particularly useful for discriminating the severity levels of thermal defects, distinguishing between the phases of progressive heating (T1) up to severe thermal degradation (T3). Finally, CH_4/H_2, although less specific, provides valuable complementary information for all classes, making it a general support predictor.

3. EXPERIMENTS AND DISCUSSION OF RESULTS

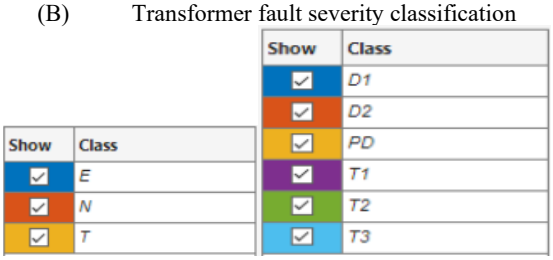
A precise assessment of the condition of electrical equipment is a strategic tool, allowing either to optimize their lifespan when they are in service or to serve as a starting point for targeted maintenance of defective units. In this context, this study relied on advanced machine learning techniques to improve the reliability and accuracy of the diagnosis, by combining data preprocessing methods tailored to high-performance classification algorithms.



(A) Transformer fault state classification



(B) Transformer fault severity classification



(C)

Fig. 6. SHAP Values for Feature Importance in Dissolved Gas Analysis-Based Transformer Diagnosis

Table 4. Performance Comparison of Transformer Fault Detection Models Under Transformation Methods and Cross-Validation Schemes

Algorithm		Transformation	Precision	Recall	F1-score	Accuracy
Fine Tree	CROSS 5	RAW	72,8	72,5	72,6	72,53
		LOG	75,3	74,8	74,9	74,85
	CROSS 10	RAW	76,3	76,2	76,3	76,23
		LOG	76,2	76,1	76,1	76,08
Optimizable Tree	CROSS 5	RAW	76	75,8	75,8	75,77
		LOG	77,7	76,9	76,9	75,85
	CROSS 10	RAW	78,2	77,8	77,9	77,78
		LOG	78,1	77,8	77,8	77,78
ENSEMBLE Boosted Trees	CROSS 5	RAW	76,5	76,2	76,3	76,23
		LOG	77,5	77,2	77,2	77,16
	CROSS 10	RAW	76,6	76,5	76,5	76,54
		LOG	76,6	76,5	76,5	76,54
Optimizable Ensemble	CROSS 5	RAW	80,8	80,7	80	80,71
		LOG	81,4	81,3	81,3	81,33
	CROSS 10	RAW	81,5	81,3	81,3	82
		LOG	82,2	82,1	82,1	82,1

The models studied include fine tree, optimizable tree, boosted trees, and ensemble methods, selected for their ability to efficiently handle complex and heterogeneous data. The experimentation was conducted on a dataset of 710 samples, incorporating both raw data and preprocessed data, in order to evaluate and compare the performance of each approach in various analysis and defect detection scenarios.

3.1. Fault detection

This resulted in improved predictive accuracy. The results show that applying a logarithmic transformation to the input data consistently improves the model's performance across all evaluated metrics, whether using 5-fold cross-validation (CV5) or 10-fold cross-validation (CV10). Table 4 presents the precision, recall, F1-score, and overall accuracy obtained for different algorithms tested under MATLAB's Classification Learner on a set of 710 samples. Two data configurations (raw and logarithmic) and two cross-validation schemes (5-fold and 10-fold) were applied.

The analysis shows that the Optimizable Ensemble achieves the best performance in all scenarios. With logarithmic transformation and 10-fold cross-validation, it achieves 82.2% precision, 82.1% recall, and 82.1% F1-score. This model consistently exceeds 80% accuracy, confirming its superiority over the other models.

The Optimizable Tree ranked second, with a maximum accuracy of 77.78%, demonstrating the positive impact of hyperparameter optimization. The logarithmic transformation provides a slight gain, particularly under 5-fold cross-validation.

The Boosted Trees Ensemble shows stable performance, close to that of the Optimizable Tree, while the Fine Tree serves as a basic reference, with a maximum accuracy of 76.23%.

Overall, the logarithmic transformation generally enhances or maintains classification performance, with its effectiveness varying according to the model and cross-validation configuration. For example, Fine Tree shows an improvement in CV5, whereas its performance remains nearly unchanged in CV10.

Similarly, the Optimizable Tree and Boosted Trees exhibit slight improvements under CV5, while their performance remains stable under CV10. The Optimizable Ensemble benefits from the logarithmic transformation under both CV5 and CV10.

Overall, the performance remains higher with CV10 than with CV5, regardless of the logarithmic transformation method. This can be explained by the fact that a larger proportion of training data in each fold allows the model to capture more representative patterns.

To summarize, the choice of cross-validation strategy and the use of logarithmic transformation both influence classification performance, with the Log + CV10 configuration providing the best overall results across the evaluated metrics (Fig. 7).

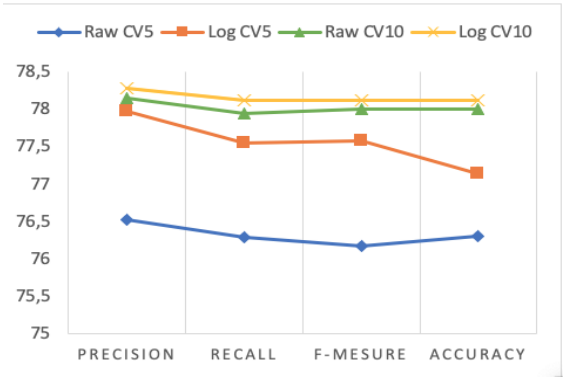


Fig. 7. Impact of Data Transformation and Cross-Validation Strategy on Overall Model Performance for Transformer Fault Detection

3.1.1. Comparison between matrix confusion

The comparative evaluation of the two model configurations reveals a superior performance of the Optimizable Ensemble combined with a logarithmic transformation and 10-fold cross-validation, compared to the Fine Tree model applied to raw data with 5-fold cross-validation (Fig. 8) .The Fine Tree, although simple and interpretable, exhibits high sensitivity to amplitude variations and extreme values of ppm concentrations, which results in notable confusions between thermal (T) and normal (N) classes, as well as between electrical (E) and normal classes. More precisely, the T class is incorrectly predicted as N in 20.6% of cases, and the E class is confused with N in 12.3% of cases, reflecting an imperfect separation of decision boundaries when the data is not transformed.

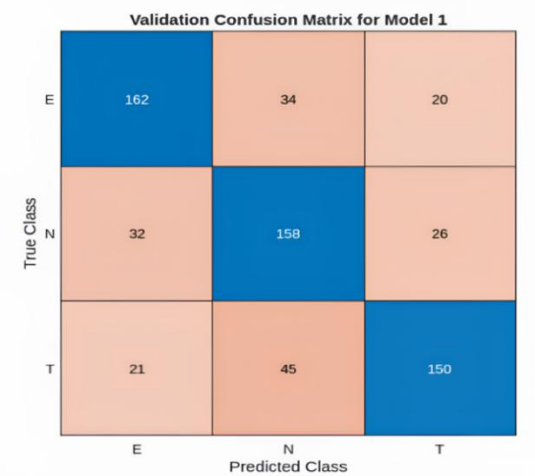
Conversely, the Optimizable Ensemble benefits from the logarithmic transformation that compresses high values and accentuates differences in low concentrations, thereby reducing the dispersion effect and improving inter-class separability. This configuration reduces critical T↔N confusions by more than half and improves the classification accuracy of classes N and T by +10.2% and +10.1% respectively, while increasing the overall stability of the model. The use of 10-fold cross-validation also provides a better estimate of actual performance by exposing the algorithm to a greater diversity of training and test samples, thereby limiting overfitting.

These results show that the combination of adapted data transformation with an optimized overall approach constitutes a robust and effective strategy for multi-class diagnosis through dissolved gas analysis in oil, offering a tangible improvement in the reliability and accuracy of power transformer maintenance decisions.

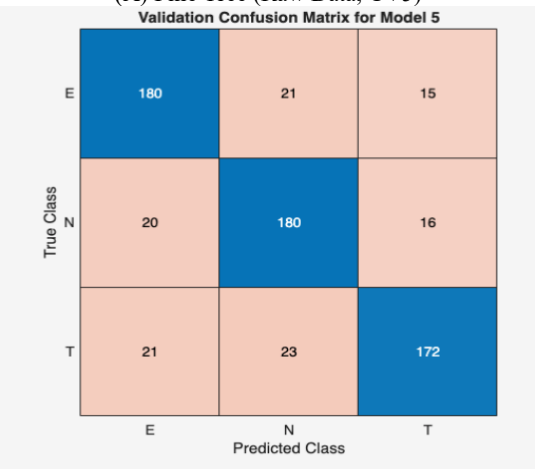
3.2. Degree of severity of a specific defect

Detecting a fault and identifying its cause is not enough; it is imperative to assess the severity of the

fault to effectively prioritize maintenance actions on transformers. This assessment helps determine which units need to be immediately taken out of service for repair, which ones can be repaired without shutting down operations, and which ones can be deferred without major risk.



(A) Fine Tree (Raw Data, CV5)



(B) Optimizable Ensemble (Log Transformation, CV10)

Fig. 8. Comparative Confusion Matrix for Transformer Fault Detection

The analysis of the performance of algorithms applied to the classification of the transformer faults severity (PD/D1/D2 for electrical faults and T1/T2/T3 for thermal faults) highlights clear differences between simple and advanced models (Table 5).

Fine Tree-type decision trees show the most modest scores, with accuracies limited to 75–77%, reflecting a restricted ability to capture the complexity of relationships between variables. Conversely, the optimized (Optimizable Tree) and ensemble models (Boosted Trees and Optimizable Ensemble) regularly exceed 81%, reaching a maximum of 82% for Optimizable Ensemble with logarithmic transformation and 5-fold cross-validation. The logarithmic transformation brings a notable gain mainly in the 5-fold cross-validation scenarios, where it reduces the impact of extreme values and improves generalization, whereas in 10-fold cross-validation the effect is more limited, even negative in some cases, due to the increased stability offered by a larger training volume.

Overall, Optimizable Ensemble and Boosted Trees stand out for their robustness and consistency, making them the most relevant candidates for an industrial application of predictive maintenance, particularly in contexts where the data exhibits high variability and class imbalance.

The results obtained for the severity data confirm the significant influence of the logarithmic transformation on the model's performance, although its effect varies depending on the cross-validation strategy used. With CV5, the application of the logarithmic transformation improves all metrics, indicating a better ability of the model to extract relevant relationships from the data.

On the other hand, with CV10, the raw data slightly better performance than those trained on logarithmically transformed data. This suggests that when the model has a larger training set per fold, the logarithmic transformation does not provide additional benefits and may even dampen some useful signals.

Overall, these results highlight that the impact of data transformation is not uniform and depends on the learning configuration. The optimal choice therefore results from an appropriate combination of preprocessing and cross-validation strategy, taking into account the nature and distribution of the processed data (See Fig. 9).

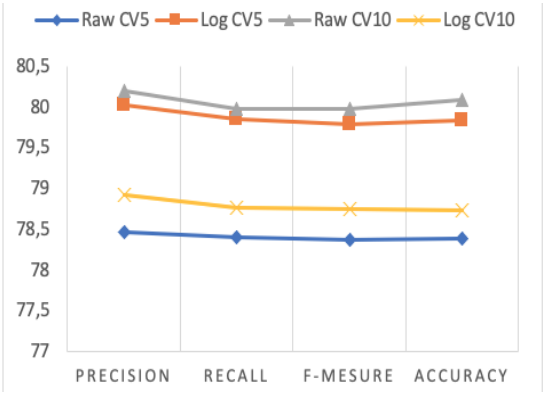


Fig. 9. Impact of Data Transformation and Cross-Validation Strategy on Overall Model Performance for Transformer fault severity classification

3.2.1. Comparison between matrix confusion

The comparative analysis of confusion matrices highlights a notable improvement in performance when transitioning from the Fine Tree model applied to raw data with 10-fold cross-validation to the Optimizable Ensemble trained on logarithmically transformed data with 5-fold cross-validation. The Fine Tree shows a tendency to confuse certain closely related classes in their gas signatures, particularly between electrical faults D1 and D2, as

Table 5. Performance Comparison of Transformer fault severity Models Under Different Transformation Methods and Cross-Validation Schemes

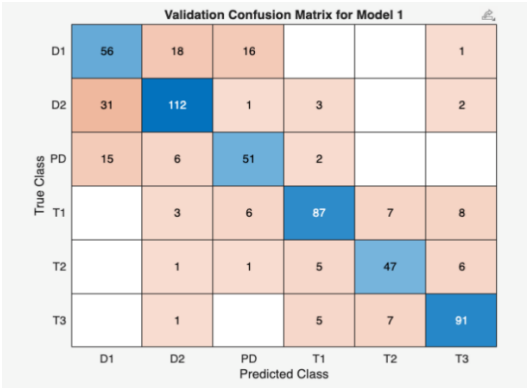
Algorithm		Transformation	Precision	Recall	F1-score	Accuracy
Fine Tree	CROSS 5	RAW	75,9	75,7	75,8	75,72
		LOG	77,1	77,1	77	77,08
	CROSS 10	RAW	75,3	74,9	75	75,36
		LOG	76	75,7	75,8	75,55
Optimizable Tree	CROSS 5	RAW	79	78,8	78,8	78,78
		LOG	79,6	79,3	79,3	79,29
	CROSS 10	RAW	81,6	81,5	81,5	81,49
		LOG	78,2	78,3	78,1	78,27
Boosted Trees	CROSS 5	RAW	79,9	79,8	79,8	79,8
		LOG	81,2	81	81	80,98
	CROSS 10	RAW	81,9	81,7	81,6	81,66
		LOG	79,9	79,8	79,8	79,8
Optimizable Ensemble	CROSS 5	RAW	79,1	79,3	79,1	79,29
		LOG	82,2	82	81,9	82
	CROSS 10	RAW	82	81,8	81,8	81,83
		LOG	81,6	81,3	81,3	81,32

well as between PD and T1, which reflects limitations in separating decision boundaries in the face of the natural variability of DGA measurements. For example, the PD class is correctly predicted in 51 cases but still confused with D1 in 15 cases and with T1 in 2 cases, while D1 is often misclassified as D2 (18 cases) or PD (16 cases). On the other hand, the Optimizable Ensemble, benefiting from the logarithmic transformation that mitigates the impact of extreme values and improves inter-class separation, reduces these confusions. Thus, PD sees its correct predictions increase to 54 cases, with less dispersion toward D1 and D2, while D1 and D2 gain 5 and 18 additional correct predictions respectively compared to the Fine Tree. The thermal classes (T1, T2, T3) also benefit from this configuration, with better accuracy and fewer cross-errors among them, notably an increase in correct predictions from T1 from 87 to 93 and from T3 from 91 to 94. This overall improvement demonstrates the relevance of jointly using a logarithmic transformation and an optimized ensemble method, which enhances the model's robustness against data variations and improves the reliability of multi-class transformer fault diagnosis via DGA analysis.

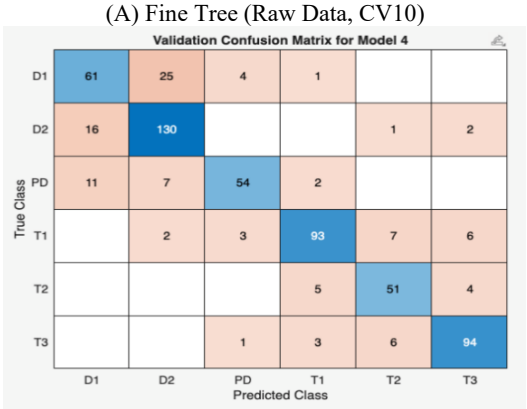
3.3 Comparison with Conventional methods

3.3.1 Fault detection

To evaluate the effectiveness of the proposed model, a comparative analysis was conducted against the three most widely used conventional DGA interpretation techniques: the Duval Triangle, the Rogers Ratio Method, and the IEC 60599 Ratio Method. All four methods were evaluated on exactly the same dataset and samples, using the same fault-class definitions and evaluation conditions, to ensure a fair and directly comparable assessment of their classification performance. For a fair comparison, the fault subclasses were grouped into two main categories, namely thermal faults (T1 + T2 + T3 → T) and electrical faults (PD + D1 + D2 → E), while the normal operating condition (N) was retained as a separate class. The comparative results summarized in Table 6 indicate that the conventional ratio-based



(A) Fine Tree (Raw Data, CV10)



(B) Optimizable Ensemble (Log Transformation, CV (5)

Fig. 10. Comparative Confusion Matrix for Transformer fault severity classification

Table 7. Comparison of Optimizable Ensemble Cross 5-LOG with Conventional Methods of transformer fault severity Models

Classifier	Overall Recall (%)	Per Class Recall (%)					
		PD	**D1**	**D2**	**T1**	**T2**	**T3**
DT [27] [28]	51.27	37.84	74.73	15.44	49.55	50.00	94.23
RR [27] [28]	54.16	89.19	48.35	15.44	47.75	85.00	78.85
IEC [27] [28]	59.42	70.27	49.45	22.15	82.88	85.00	74.04
Optimizable Ensemble (Log CV5)	82.04	72.97	67.03	87.25	83.78	85	90.38

methods achieve overall recall ranging from 71.8% to 76.5%, whereas the proposed Optimizable Ensemble model (Cross 10-LOG) attains an overall recall of 82.1%. Similar improvements are observed for the individual classes, where the proposed model achieves recall values of 83.33%, 79.63%, and 83.33% for the electrical (E), thermal (T), and normal (N) classes, respectively.

The observed performance improvement highlights the limitations of traditional ratio-based diagnostic methods when dealing with complex fault patterns and overlapping gas distributions. Conventional approaches rely on fixed decision thresholds and predefined diagnostic regions, making them less effective when gas concentrations are close to boundary conditions. In contrast, the proposed ensemble model exploits the nonlinear relationships and interactions among dissolved gas features, resulting in a more robust and reliable classification. Nevertheless, conventional DGA methods remain valuable in industrial practice due to their simplicity, interpretability, and compliance with international standards, while machine learning models require representative datasets and careful validation before deployment in real-world transformer condition monitoring.

Table 6. Comparison of Optimizable Ensemble Cross10-LOG with Conventional Methods of Transformer Fault Detection Models

Method	Overall Recall (%)	Per Class Recall (%)		
		E	**T**	**N**
DT [19] [21]	71.8	69.5	72.4	77.3
RR [22] [23]	74.2	72.8	75.1	79.2
IEC [24] [25]	76.5	75	77.3	81
Optimizable Ensemble (Cross10-LOG)	82.1	83.33	79.63	83.33

3.3.2 Degree of severity of a specific defect

The performance of the proposed Optimizable Ensemble (Log CV5) model was compared with three conventional DGA fault diagnosis methods, namely the Duval Triangle Method (DTM), Rogers' Ratio Method (RRM), and the IEC 60599 Method (IECM). As presented in Table 7, the proposed model achieved the highest overall recall of 82.04%, clearly outperforming the conventional methods, whose overall Recall ranged from 51.27% to 59.42%.

A detailed comparison of the per-class results further confirms the effectiveness of the proposed approach. The Optimizable Ensemble (Log CV5) model achieved recall values of 72.97%, 67.03%, 87.25%, 83.78%, 85.00%, and 90.38% for the PD, D1, D2, T1, T2, and T3 fault classes, respectively. In particular, the proposed model demonstrated excellent performance for D2, T2, and T3, while maintaining strong performance across most fault categories. In contrast, the conventional methods exhibited highly variable performance depending on the fault type. For example, although the Duval Triangle Method achieved a high recall for T3 (94.23%) and the Rogers' Ratio Method performed well for PD (89.19%), both methods yielded very poor recall for D2 (15.44%). Similarly, the IEC 60599 Method showed improved performance for T1 (82.88%) and T2 (85.00%), but its performance remained substantially lower than that of the proposed model in terms of overall recall values.

These results demonstrate that the proposed Optimizable Ensemble (Log CV5) model provides a more balanced and reliable diagnostic performance across all transformer fault categories. Unlike conventional ratio-based methods, which rely on fixed thresholds and predefined decision rules, the proposed ensemble learning approach effectively captures the complex nonlinear relationships among dissolved gas features, thereby improving fault discrimination and enhancing the robustness of transformer fault diagnosis. **External Validation on Unseen Data**

3.3.3 Fault detection

The use of different cross-validation configurations for the two classification tasks was determined from the comparative experiments reported in Tables 4 and 5. Both CV5 and CV10 were evaluated for each task using the same preprocessing strategies and classification algorithms. For fault detection, the Optimizable

Ensemble with logarithmic transformation achieved its highest validation accuracy with CV10 (82.1%), whereas for fault severity classification, the highest accuracy was obtained with the Optimizable Ensemble using logarithmic transformation and CV5 (82.0%). Therefore, the Log-CV10 and Log-CV5 configurations were respectively selected for external validation of fault detection and fault severity classification. This selection was based on the experimentally observed performance of each configuration for its corresponding task and does not imply that one cross-validation strategy is universally superior to the other.

To further evaluate the generalization capability of the proposed diagnostic framework, an external validation was performed using an independent dataset composed of 90 transformer oil samples that were completely excluded from the training, validation, and preprocessing stages. [21][29] The best-performing model, namely the Optimizable Ensemble (Log CV10), was directly applied to these unseen samples without any parameter adjustment.

For the external validation, an independent dataset, completely separate from the dataset used for model training, was employed to evaluate both the proposed machine-learning models and the conventional DGA diagnostic methods. The external dataset consisted of the same type of DGA data used in the machine-learning analysis. For the machine-learning approach, the external data underwent the same preprocessing steps described in the training phase before being provided to the trained models, without any retraining or parameter adjustment. In contrast, the conventional DGA methods, namely the Duval Triangle, Rogers Ratio, and IEC 60599, were directly applied to the original DGA data without preprocessing, according to their respective diagnostic procedures. Thus, both approaches were evaluated on the same independent external dataset, while each method was applied according to its appropriate and standard processing procedure.

Table 8 summarizes the classification performance of the proposed model on the independent external validation dataset, while the corresponding confusion matrix is presented in Figure 11. The model achieved a recall of 96.67% for the Electrical (E) class, with 29 out of 30 samples correctly classified. For the Normal (N) class, 27 out of 30 samples were correctly classified, corresponding to a recall of 90.00%, whereas the Thermal (T) class achieved a 100% recall, with all 30 samples correctly classified. Overall, 86 out of 90 samples were correctly classified, resulting in an overall recall of 95.56%.

As shown in the confusion matrix in Figure 11, one Electrical (E) sample was misclassified as Normal (N), while three Normal (N) samples were misclassified: one as Electrical (E) and two as Thermal (T). No Thermal (T) samples were misclassified, indicating complete recognition of this class in the external validation dataset. These results indicate that the proposed model maintains a high level of predictive performance when applied to previously unseen data. Nevertheless, given the size and composition of the external validation dataset, these results should be interpreted as additional experimental evidence of the model's predictive capability rather than as definitive evidence of generalization or absence of overfitting. Overall, the findings support the potential applicability of the proposed diagnostic framework for transformer condition monitoring based on dissolved gas analysis.

Table 8. Classification performance on the external validation dataset for fault detection.

Fault Detection	Test Samples	Correctly Classified	Recall (%)
Electrical (E)	30	29	96.67
Normal (N)	30	27	90
Thermal (T)	30	30	100.0
Overall	**90**	**86**	**95.56**

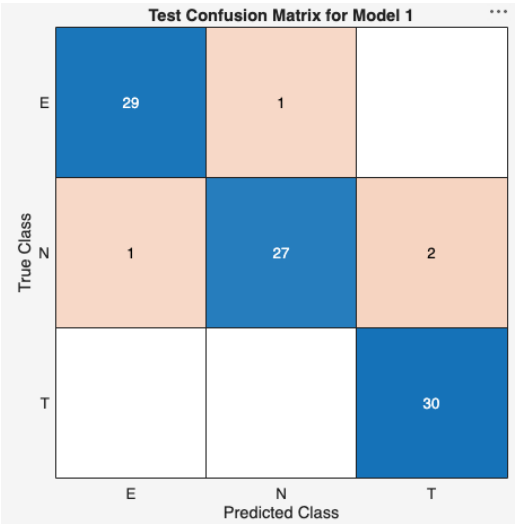


Fig. 11. Test Confusion Matrix for Transformer fault detection classification

3.3.4 External Validation for Defect Severity Classification

To further evaluate the generalization capability of the proposed diagnostic framework for detailed fault identification, an external validation was performed using an independent dataset of 30 unseen transformer oil samples, equally distributed among the six fault severity classes (PD, D1, D2, T1, T2, and T3). These samples were completely excluded from the training, validation, and model optimization stages to ensure an unbiased assessment of the model's predictive capability.

The best-performing model, namely the Optimizable Ensemble (Log-CV5), was directly applied to the independent external dataset without any retraining or parameter adjustment. The model correctly classified **56 out of 60 samples**, achieving an overall recall of **93.33%**. The corresponding confusion matrix is presented in **Figure 12**,

providing a detailed visualization of the classification performance across the six fault severity classes.

As summarized in Table 9, the proposed model achieved a **100% recall** for the D1, D2, T1, T2, and T3 fault classes, with all samples in these classes correctly classified. In contrast, the PD class achieved a recall of **60.0%**, with **6 out of 10 samples correctly classified**. As shown in the confusion matrix, the four misclassified PD samples were classified as D1 (one sample) and T1 (three samples). No misclassification was observed for the D1, D2, T1, T2, or T3 classes.

Overall, the external validation resulted in **56 correctly classified samples out of 60**, corresponding to an accuracy of **93.3%**. The high recognition rates obtained for five of the six fault severity classes indicate that the proposed Optimizable Ensemble (Log-CV5) model maintains a high level of predictive performance when applied to previously unseen data. However, the lower recall observed for the PD class highlights a specific classification limitation, with some PD samples being confused with D1 and T1. This behavior may be related to similarities in the dissolved gas patterns associated with these fault categories. Given the size and composition of the external dataset, these results should be interpreted as additional experimental evidence of the model's predictive capability rather than as definitive evidence of generalization or absence of overfitting. Nevertheless, the results demonstrate the potential applicability of the proposed framework for transformer condition monitoring and fault severity classification based on dissolved gas analysis.

Table 9. Classification performance on the external validation dataset for fault severity classification

Fault Severity	Test Samples	Correctly Classified	Recall (%)
PD	10	6	60.0
D1	10	10	100.0
D2	10	10	100.0
T1	10	10	100.0
T2	10	10	100.0
T3	10	10	100.0
Overall	**60**	**56**	**93.33**

4. CONCLUSION

This study demonstrated the effectiveness of supervised learning, particularly models from the decision tree family and ensemble methods, for diagnosing faults in power transformers using dissolved gas analysis (DGA) data. The adopted approach integrated several key steps, ranging from data preparation and the transformation of gas ratios according to Rogers' method to hyperparameter optimization and performance evaluation using precision, recall, F1-score, and overall accuracy.

The results show that the effect of logarithmic transformation depends on the classification model and validation configuration. While logarithmic transformation improved the performance of several models, it did not consistently outperform the raw representation in all cases. For fault detection, the Optimizable Ensemble combined with logarithmic transformation and CV10 achieved the highest validation accuracy of 82.1%, whereas for fault severity classification, the same model with logarithmic transformation and CV5 achieved the highest validation accuracy of 82.0%. These results indicate that logarithmic transformation can be beneficial for specific model configurations, but its effect should be evaluated according to the classification task and validation strategy rather than considered as a universally superior representation. The external validation results further showed accuracies of 95.56% for fault detection and 93.33% for fault severity classification on the independent datasets.

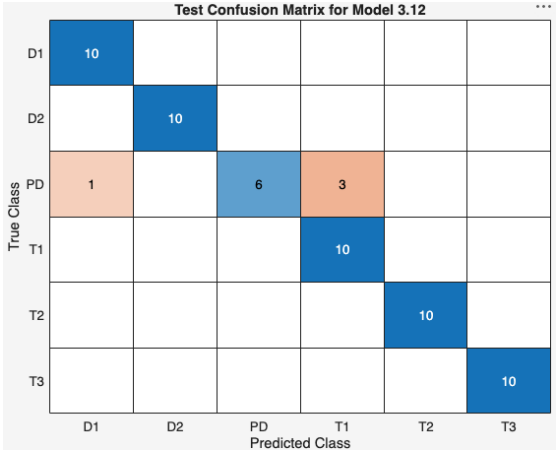


Fig. 12. Test Confusion Matrix for Transformer fault severity classification

The explainability analysis using SHAP values highlighted the predominant role of certain gas ratios in the detection and classification of faults. In particular, the CH_4/H_2 ratio proved important for identifying thermal faults, while C_2H_2/C_2H_4 served as an important indicator of electrical faults and their severity. This information provides additional insight into the model's behavior and contributes to the interpretability of the proposed diagnostic framework.

In conclusion, this study highlights the importance of combining appropriate data preparation, relevant algorithm selection, comparative evaluation of data representations, and interpretable model analysis to develop an effective diagnostic framework for power transformers. The findings indicate that logarithmic transformation can improve performance in specific configurations, but its benefits are not uniform across all models and validation settings. Future work could focus on integrating deep learning methods and real-time data

from connected sensors to further investigate early fault detection and predictive maintenance of power transformers.

Source of funding: *This research did not receive any outside financial support.*

Acknowledgment: *Optional - this data is not required.*

Author contributions: *research concept and design, B.S.; Collection and/or assembly of data, B.S.; Data analysis and interpretation, B.S.; Writing the article, B.S.; Critical revision of the article, B.S., C.D.E., S.M.E., F.Z.B.; Final approval of the article, B.S., C.D.E., S.M.E., A.M., F.Z.B.*

Declaration of competing interest: *The authors affirm that the work described in this publication was not impacted by any known competing financial interests or personal ties.*

AI Use Statement: *The authors used an artificial intelligence (AI) tool to assist with language editing, including grammar correction, spelling correction, and text rephrasing. The AI tool was not used for data analysis, interpretation of results, scientific content generation, or drawing conclusions.*

REFERENCES

1. Arvind DK, S. Khushdeep, K. Deepak. Condition monitoring of power transformer: A review. A review IEEE/PES Transmission and Distribution Conference and Exposition. 2008;1:1-6. https://doi.org/10.1109/TDC.2008.4517046.
2. Mackenzie EA, Crossey J, dePablo A, Ferguson W. On-line monitoring and diagnostics for power transformers. IEEE International Symposium on Electrical Insulation. 2010:1-5. https://doi.org/10.1109/ELINSL.2010.5549734.
3. Meysam Beheshti Asl, Fofana I, Fethi Meghnefi. Review of various sensor technologies in monitoring the condition of power transformers. energies. 2024; 17(14):3533–3. https://doi.org/10.3390/en17143533.
4. Dhara S, Koley C, Chakravorti S. An ultrawideband partial discharge sensor for high-voltage power transformers. IEEE Sensors Journal. 2024;24(10):16261–9. https://doi.org/10.1109/JSEN.2024.3383064.
5. Ali MS, Abu Bakar AH, Omar A, Abdul Jaafar AS, Mohamed SH. Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review. Electric Power Systems Research. 2023 Mar; 216:109064. https://doi.org/10.1016/j.epsr.2022.109064.
6. N'cho J, Fofana I, Hadjadj Y, Beroual A. Review of physicochemical-based diagnostic techniques for assessing insulation condition in aged transformers. Energies. 2016 May 13;9(5):367. https://doi.org/10.3390/en9050367.
7. Rogers R. IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis. IEEE Transactions on Electrical Insulation. 1978 Oct; EI-13(5):349–54. https://doi.org/10.1109/TEI.1978.298141.
8. International Electrotechnical Commission. IEC, 60599. Mineral Oil-Impregnated Electrical Equipment in Service–Guide to the Interpretation of Dissolved and Free Gases Analysis", International Electrotechnical Commission: Geneva, Switzerland, 2015. https://webstore.iec.ch/en/publication/23323?utm.
9. Duval M. A review of faults detectable by gas-in-oil analysis in transformers. IEEE Electrical Insulation Magazine [Internet]. 2002 May 1 [cited 2021 Aug 26];18(3):8–17. Available from: https://doi.org/10.1109/MEI.2002.1014963.
10. Wanjare SB, Swami PS. DGA Interpretation for increasing the percent of accuracy by different methods: A Review. 2018 Internat2018 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)ional conference on computation of power, energy, Information and Communication (ICCPEIC). 2018 Mar;458–61. https://doi.org/10.1109/ICCPEIC.2018.8525179.
11. Arnaud Nanfak, Samuel E, Fofana I, Fethi Meghnefi, Martial Gildas Ngaleu, Kom CH. Traditional fault diagnosis methods for mineral oil-immersed power transformer based on dissolved gas analysis: Past, present and future. IET Nanodielectrics. 2024 Apr 22;7(3):97–130. https://doi.org/10.1049/nde2.12082.
12. Khan Mam. Ai and machine learning in transformer fault diagnosis: a systematic review. SSRN Electronic Journal. 2025. https://dx.doi.org/10.2139/ssrn.5190585.
13. Wang L, Littler T, Liu X. Hybrid AI model for power transformer assessment using imbalanced DGA datasets. IET renewable power generation. 2023 Apr 13;17(8):1912–22. https://doi.org/10.1049/rpg2.12733.
14. Bargate V, Yadav A, Gupta S. Combining KNN Algorithm with Traditional DGA Techniques for Transformer Defect Identification and Analysis. 2024 4th Asian Conference on Innovation in Technology (ASIANCON). 2024;23:1–6. https://doi.org/10.1109/ASIANCON62057.2024.10837587.
15. Liu G, Liu Y, Li H, Liu K, Gao J, Zhong L. Enhancing DGA transformer fault diagnosis using data reconstruction and GD-AHBA-SVM model. IEEE Transactions on Dielectrics and Electrical Insulation. 2024;31(6):2966–73. https://doi.org/10.1109/TDEI.2024.3395235.
16. Bustamante S, Lastra JLM, Manana M, Arroyo A. Use of a decision tree to define transformer oil contamination from on-load tap-changer gases to ensure power quality in the grid. Renewable Energy and Power Quality Journal. 2024 Aug 5;119–24. https://doi.org/10.52152/4014.
17. Breiman L. Bagging predictors. Machine learning. 1996 Aug;24(2):123–40. Available from: https://doi.org/10.1007/BF00058655.
18. Friedman JH. Greedy function approximation: a gradient boosting machine. The Annals of Statistics. 2001;29(5):1189–232. https://www.jstor.org/stable/2699986.
19. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16. 2016 Aug

13;1(1):785–94. https://doi.org/10.1145/2939672.2939785.
20. Taha IBM, Ghoneim SS, Duaywah AS. Refining DGA methods of IEC Code and Rogers four ratios for transformer fault diagnosis. IEEE Power and Energy Society General Meeting (PESGM). 2016;1:1–5. https://doi.org/10.1109/PESGM.2016.7741157,
21. Senoussaoui MEA, Brahami M, Fofana I. Combining and comparing various machine-learning algorithms to improve dissolved gas analysis interpretation. IET Generation, Transmission & Distribution. 2018;12(15):3673–9. https://doi.org/10.1049/iet-gtd.2018.0059.
22. Dhiman A, Azmeera HK, Reddy CC. Intelligent Diagnosis of Transformer Faults Using DGA. 2024 IEEE 7th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON). 2024 Nov 22;166–70. https://doi.org/10.1109/CATCON60527.2024.10831256.
23. Sarma S. Performance assessment of customized LSTM based deep learning model for predictive maintenance of transformer [Internet]. Forexjournal.co.in. 2023. Availablefrom: https://ijeer.forexjournal.co.in/archive/volume-11/ijeer-110220.html.
24. Chen T. Introduction to boosted trees. University of Washington Computer Science, vol. 22, no 115, pp. 14-40,2014. https://web.njit.edu/~usman/courses/cs675_fall16/BoostedTree.pdf.
25. Moradi E, Mahmoud K, Darwish MMF. Advanced ensemble learning classifiers for power transformer fault diagnosis with imbalanced DGA Data. 25th International Middle East Power System Conference (MEPCON). IEEE. 2024:1-5. https://doi.org/10.1109/MEPCON63025.2024.10850169.
26. Naidu G, Zuva T, Sibanda EM. A Review of evaluation metrics in machine learning algorithms. Lecture notes in networks and systems. 2023; 724:15–25. https://doi.org/10.1007/978-3-031-35314-7_2.
27. Tarade RS, Kundu P. A performance comparison of fault diagnosis methods for power transformer using machine learning techniques based on DGA. 2026 International Conference on Electric Power and Renewable Energy (EPREC). 2026;2:1–5. https://doi.org/10.1109/EPREC66546.2026.11412063.
28. Mansour DEA. Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases. IEEE Transactions on Dielectrics and Electrical Insulation. 2015;22(5):2507–12. https://doi.org/10.1109/TDEI.2015.004999.
29. Boudjella FZ, Boudjella S, Ahmed NY, Illias HA, Latifa S, Hafidha R. Transformer fault diagnosis using dissolved gas analysis: A hybrid ensemble model with data preprocessing. International Journal of Power Electronics and Drive Systems (IJPEDS), 2026;*17*(3):1780. https://doi.org/10.11591/ijpeds.v17.i3.pp1780-1793.



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.
E-mail: souhila.boudjella@univ-mascara.dz.



Djamel Eddine Chaouch is Associate Professor since 2004 at electrical engineering departement, faculty of sciences and technology, Mustapha Stambouli University of Mascara, Algeria. He received the Magister and Ph.D. degree in Electrical Engineering, from USTO University at Oran, Algeria in 2004 and 2016. The university habilitation in 2021. His research areas are Robotic, Computer automation, control of industrial processes and supervis artificial intelligence in power system applications, Identification of process, energy renewable, soft computing. He is member of a research team at LIS2T Laboratory.
E-mail: d.chaouch@univ-mascara.dz



Mohammed El Amine Senoussaoui was born in Mascara Algeria. February 7th, 1986 He is graduated from University Mustapha Stambouli of Mascara in2009 He received the Master and Ph.D. degree in electrical engineering, from Djilali Liabes University at Sidi Blabbess, Algeria in2012 and 2019. He is currently an associate professor in the department of electrical engineering, Mascara University, Algeria. His research interests are the application of artificial intelligence techniques in diagnosis of electrical equipment.
E-mail : m.senoussaoui@univ-mascara.dz



Adnane Mouffak is Associate Professor since 2004 at electrical engineering departement, faculty of sciences and technology, Mustapha Stambouli University of Mascara, Algeria. He Holds a PhD degree in Electrical Engineering with specialization in signal processing and electronics. His research areas are signal processing, artificial intelligence in power system applications as well as instrumentation and modelization of industrial processes. He is head of a research team at LGPCS Laboratory and project manager in diagnostics power systems project.
E-mail : mouffakadn@univ-mascara.dz .



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.
e-mail: fatima.boudjella@univ-temouchent.edu.dz