International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(9s) 2026 pp. 814-829
https://doi.org/10.70917/ijcisim-2026-3504
Article

Optimization-Driven Intelligent Transformer Fault Diagnosis Using Machine Learning: A Comparative Benchmark of Genetic Algorithm, Particle Swarm Optimization, AutoML, and Conventional DGA Techniques

Parineeta Ritesh Kelkar[1*], Sharanabasava Inamadar[2]

[1*]School of Engineering Ajeenkya D Y Patil university City – Pune Email parineetariteshkelkar@gmail.com Orcid id.- 0009-0005-5014-6627
[2]School of Engineering Ajeenkya D Y Patil university Pune Email - sharan.inamadar@ adypu.edu.in
Orcid id -0009-0006-6741-1711

Abstract: Power transformers are critical assets in electrical power systems, and their reliable operation depends on accurate and timely fault diagnosis. Although Dissolved Gas Analysis (DGA) is widely used for transformer condition monitoring, conventional diagnostic techniques such as the Rogers Ratio and Duval Triangle often exhibit limited accuracy due to predefined decision rules and their inability to effectively handle complex fault patterns. This study proposes an optimization-driven machine learning framework for intelligent transformer fault diagnosis by integrating Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Tree-based Pipeline Optimization Tool (TPOT) AutoML with supervised machine learning classifiers. Unlike existing studies that primarily investigate a single optimization strategy, the proposed framework provides a unified comparative evaluation of three optimization approaches and benchmarks them against conventional DGA techniques under identical experimental conditions. Experimental evaluation using five-fold cross-validation demonstrates that the GA-optimized Logistic Regression model achieved the highest testing accuracy of 96.22%, followed by PSO-optimized Logistic Regression (95.67%) and TPOT AutoML (95.00%). In comparison, the Rogers Ratio and Duval Triangle methods achieved accuracies of only 44.31% and 60.70%, respectively. Overall, the proposed framework improved diagnostic accuracy by 51.91 and 35.52 percentage points over the Rogers Ratio and Duval Triangle methods, respectively. These findings demonstrate that optimization-driven machine learning provides a robust, accurate, and practical solution for intelligent transformer health monitoring and predictive maintenance in modern power systems.
Keywords: Transformer fault diagnosis; Dissolved Gas Analysis (DGA); Genetic Algorithm; Particle Swarm Optimization; Automated Machine Learning (TPOT); Hyperparameter Optimization; Machine Learning; Predictive Maintenance

1. INTRODUCTION

Credit Power transformers are critical assets in modern electrical power systems, ensuring reliable voltage regulation and efficient power transmission across generation, transmission, and distribution networks. Unexpected transformer failures can lead to severe economic losses, power interruptions, equipment damage, and reduced grid reliability. Consequently, electric utilities are increasingly adopting predictive and condition-based maintenance strategies to detect incipient faults at an early stage and improve transformer operational reliability. Among the available condition monitoring techniques, Dissolved Gas Analysis (DGA) is widely regarded as one of the most effective and non-invasive diagnostic methods[1],[2],[3],[4]. By analyzing gases generated due to thermal and electrical degradation of transformer insulation, DGA enables the identification of fault conditions such as partial discharge, electrical arcing, and thermal faults. Conventional interpretation techniques, including the Rogers Ratio and Duval Triangle, remain widely used in industry; however, their dependence on predefined gas-ratio thresholds and deterministic decision rules often limits diagnostic accuracy under complex operating conditions involving overlapping fault signatures and measurement uncertainties[5],[6],[7],[8].

To address these limitations, machine learning (ML) has emerged as an effective data-driven alternative for transformer fault diagnosis by automatically learning nonlinear relationships among dissolved gas concentrations. Supervised learning algorithms, including SVM, DT, RF, GB, and LR, have demonstrated encouraging performance in DGA-based fault classification, while ensemble and deep learning methods have further improved diagnostic capability[9],[10],[11],[12]. Nevertheless, the effectiveness of these models largely depends on appropriate hyperparameter selection. Conventional tuning techniques such as grid search and random search are computationally expensive and frequently fail to identify globally optimal parameter combinations. Consequently, optimization approaches such as GA, PSO, and more recently AutoML have attracted considerable attention for improving model accuracy, generalization, and automation through intelligent hyperparameter optimization[13],[14],[15],[16],[17],[18].

Despite these advances, several research challenges remain. Most existing studies evaluate only a single optimization algorithm or a limited number of classifiers, making it difficult to objectively compare different optimization paradigms. Furthermore, relatively few investigations benchmark optimization-based machine learning models against conventional DGA diagnostic techniques under identical experimental conditions. To address these gaps, this study proposes an optimization-driven machine learning framework integrating GA, PSO, and TPOT AutoML with five supervised classifiers for intelligent transformer fault diagnosis. Unlike previous studies, the proposed framework provides a unified comparative evaluation of three optimization strategies and benchmarks them against the Rogers Ratio and Duval Triangle methods using identical datasets and evaluation metrics. Experimental results demonstrate that the GA-optimized Logistic Regression model achieved the highest testing accuracy of 96.22%, outperforming PSO-optimized Logistic Regression (95.67%), TPOT AutoML (95.00%), Rogers Ratio (44.31%), and Duval Triangle (60.70%).

The major contributions of this work are threefold: (i) the development of an optimization-driven transformer fault diagnosis framework integrating GA, PSO, and TPOT AutoML; (ii) a comprehensive comparative evaluation of five supervised machine learning classifiers under identical optimization settings; and (iii) systematic benchmarking against conventional DGA diagnostic methods, demonstrating the superiority of optimization-enhanced machine learning for predictive transformer maintenance. The remainder of this paper is organized as follows. Section 2 reviews related work, Section 3 presents the proposed methodology, Section 4 discusses the experimental results and comparative analysis, and Section 5 concludes the paper with future research directions.

2. Related Work

2.1 Conventional DGA-Based Transformer Fault Diagnosis

Conventional transformer fault diagnosis primarily relies on DGA, in which characteristic gases generated during insulation degradation are analyzed to identify incipient faults. Diagnostic techniques such as the Rogers Ratio, Duval Triangle, and methods recommended in IEEE C57.104 and IEC 60599 remain widely adopted because of their simplicity, interpretability, and industrial acceptance [19],[20],[21],[22],[23]. These methods classify transformer faults using predefined gas-ratio thresholds and graphical interpretation schemes, enabling practical diagnosis of partial discharge, thermal faults, and electrical discharges. However, their diagnostic capability is limited when transformers exhibit overlapping gas signatures, multiple concurrent fault mechanisms, or measurement uncertainty. Consequently, rule-based DGA interpretation often produces ambiguous or inconsistent fault classifications under complex operating conditions, motivating the development of intelligent data-driven diagnostic approaches[24].

2.2 Machine Learning-Based Transformer Fault Diagnosis

Recent advances in machine learning have significantly improved transformer fault diagnosis by enabling automatic learning of complex relationships among dissolved gas concentrations. Supervised learning algorithms such as SVM, DT, RF, GB, ANN, and LR have demonstrated improved classification accuracy compared with conventional DGA interpretation methods [25],[26],[7]. Ensemble learning techniques further enhance diagnostic robustness by combining multiple classifiers, while deep learning architectures improve feature representation for large datasets. Despite these advancements, the predictive capability of machine learning models remains strongly dependent on appropriate hyperparameter selection and feature optimization. Many studies rely on manually tuned models or evaluate only a limited number of classifiers, making objective performance comparison difficult and reducing reproducibility across different datasets.

2.3 Optimization-Based Machine Learning

To improve diagnostic accuracy and model generalization, recent studies have increasingly integrated metaheuristic optimization techniques with machine learning classifiers. GA and PSO have been widely employed to optimize classifier hyperparameters, improving predictive performance while reducing manual parameter tuning [27],

[28],[14],[16]. More recently, AutoML frameworks have gained attention by simultaneously performing feature selection, algorithm selection, pipeline generation, and hyperparameter optimization with minimal human intervention [17],[18]. Although these optimization approaches have demonstrated promising results individually, most published studies focus on a single optimization strategy or a limited number of classifiers. Comprehensive comparative evaluations involving multiple optimization paradigms under identical experimental conditions remain relatively scarce.

2.4 Research Gap

Based on the reviewed literature, several important research gaps remain. First, existing studies predominantly investigate individual optimization algorithms, making objective comparison among different optimization paradigms difficult. Second, relatively few works systematically evaluate multiple supervised machine learning classifiers using identical datasets and evaluation protocols. Third, direct benchmarking of optimization-enhanced machine learning models against conventional industrial diagnostic methods such as the Rogers Ratio and Duval Triangle is rarely reported. Finally, limited research has investigated the integration of GA, PSO, and TPOT AutoML within a unified transformer fault diagnosis framework. Addressing these gaps provides the primary motivation for the optimization-driven framework proposed in this study.

3. Proposed Optimization-Driven Framework

3.1 Overview of the Proposed Optimization-Driven Framework

The proposed optimization-driven framework is designed to improve the diagnostic accuracy of transformer fault identification by integrating intelligent hyperparameter optimization techniques with multiple supervised machine learning classifiers. The overall workflow consists of six sequential stages: DGA data acquisition, data preprocessing, baseline machine learning model development, optimization using Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Tree-based Pipeline Optimization Tool (TPOT) AutoML, followed by comparative performance evaluation against conventional DGA diagnostic methods. The complete framework is illustrated in Figure 1.

Initially, dissolved gas analysis (DGA) data are collected and organized into fault-specific categories representing different transformer operating conditions. The acquired dataset is then pre-processed to improve data quality through missing-value inspection, feature normalization, and dataset partitioning into training and testing subsets. These preprocessing steps ensure consistent feature representation and reduce potential bias during model training.

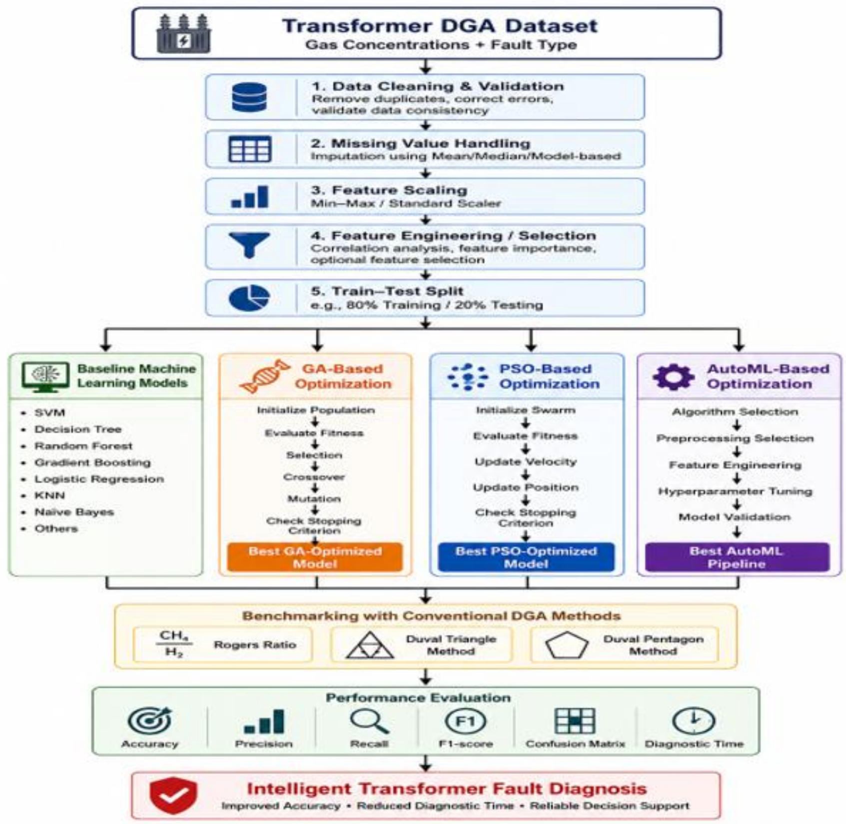

Figure 1:Proposed System Architecture

Following preprocessing, five supervised machine learning classifiers, SVM, DT, RF, GB, and LR, are developed as baseline diagnostic models. These classifiers were selected because they represent diverse learning paradigms, including margin-based, tree-based, ensemble, and probabilistic learning approaches. Training all classifiers using the same dataset and evaluation protocol enables an unbiased comparison of their predictive capabilities.

To enhance classification performance, each baseline classifier is independently optimized using three different optimization strategies. GA employs an evolutionary search process to identify suitable hyperparameter combinations, whereas PSO performs swarm-based iterative optimization by updating candidate solutions according to both individual and global search experiences. In contrast, TPOT AutoML automatically constructs complete machine learning pipelines by simultaneously performing feature selection, model selection, and hyperparameter optimization. Applying these optimization strategies independently to the same classifiers facilitates a comprehensive evaluation of their effectiveness under identical experimental conditions.

The optimized classifiers are subsequently evaluated using five-fold cross-validation and multiple performance metrics, including accuracy, precision, recall, F1-score, and confusion matrix. Finally, the best-performing optimization strategy is benchmarked against conventional transformer diagnostic techniques, namely the Rogers Ratio and Duval Triangle methods, to quantify the practical improvement achieved through optimization-driven machine learning. This systematic evaluation framework enables objective comparison of conventional and intelligent diagnostic approaches while identifying the most effective optimization strategy for transformer fault diagnosis.

3.2 DGA Dataset and Data Preprocessing

The proposed framework was evaluated using a publicly available DGA dataset obtained from Kaggle, comprising transformer oil test records corresponding to multiple transformer fault conditions. Each record contains the concentrations of key dissolved gases generated during transformer operation, including Hydrogen (H_2), Methane (CH_4), Ethane (C_2H_6), Ethylene (C_2H_4), Acetylene (C_2H_2), Carbon Monoxide (CO), and Carbon Dioxide (CO_2). These gases serve as diagnostic indicators for identifying electrical and thermal insulation faults within power transformers.

The dataset was selected because it represents a wide range of operating conditions and has been widely adopted for machine learning-based transformer fault diagnosis.

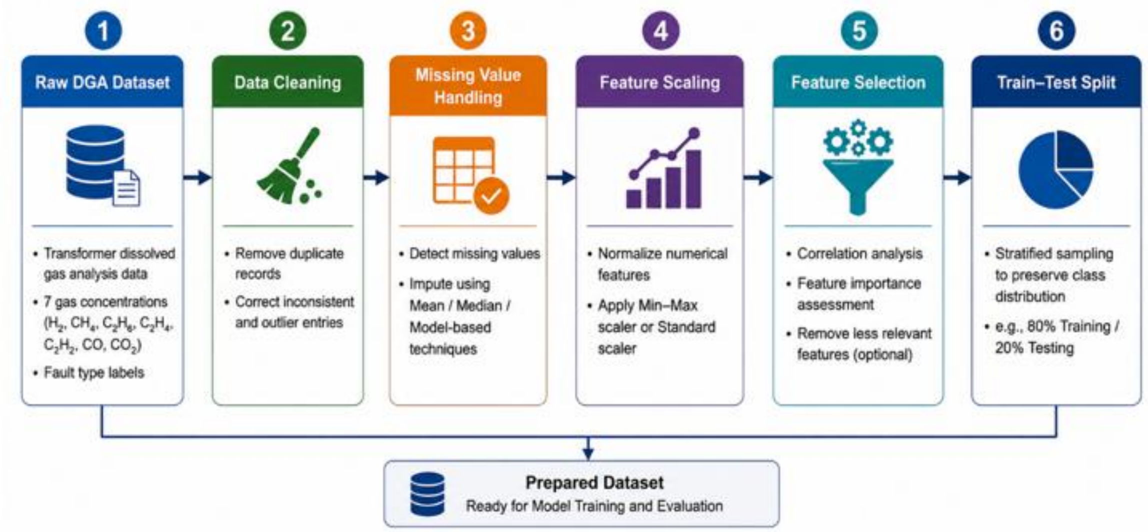

Figure 2: Overall Workflow of Dataset Preparation and Pre-processing

Prior to model development, the dataset underwent a systematic preprocessing procedure to improve data quality and ensure reliable model training. Initially, the dataset was inspected for missing values, duplicate records, and inconsistent entries. Missing or invalid observations, if present, were handled using appropriate data cleaning procedures to maintain dataset consistency. Subsequently, numerical features were normalized using Min-Max Scaling, transforming all input variables into a common range to prevent features with larger numerical values from dominating the learning process.

Following normalization, the dataset was randomly partitioned into training and testing subsets using an 80:20 ratio. The training dataset was used for classifier training and optimization, whereas the testing dataset was reserved exclusively for performance evaluation. To further improve model robustness and reduce estimation bias during hyperparameter optimization, five-fold cross-validation was employed throughout the optimization process. Figure 2 illustrates the overall data preprocessing workflow, while the dataset characteristics are summarized in Table 1.

Table 1:Summary of the DGA Dataset

Parameter	Description
Dataset Source	Kaggle DGA Dataset
Application	Transformer Fault Diagnosis
Number of Classes	4
Input Features	H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2
Output	Transformer Fault Category
Feature Scaling	Min-Max Normalization
Train-Test Split	80% : 20%
Cross Validation	5-Fold

3.3 Machine Learning Models

Five supervised machine learning classifiers were selected to evaluate the effectiveness of the proposed optimization framework for transformer fault diagnosis. The selected models include SVM, DT, RF, GB, and LR. These classifiers were chosen because they represent diverse learning paradigms, including margin-based, tree-based, ensemble-based, and probabilistic learning techniques, thereby providing a comprehensive evaluation of optimization performance across different model characteristics.

All classifiers were implemented using the Scikit-learn library under identical experimental conditions to ensure a fair comparison. Before optimization, each model was trained using the same pre-processed training dataset with default hyperparameter settings to establish baseline performance. Subsequently, the hyperparameters of each classifier were optimized independently using GA, PSO, and TPOT AutoML. The optimized models were then evaluated using the same testing dataset and five-fold cross-validation protocol.

The selected classifiers exhibit complementary learning characteristics. SVM constructs an optimal separating hyperplane for handling nonlinear decision boundaries through kernel functions. DT performs hierarchical rule-based classification by recursively partitioning the feature space. RF improves predictive stability by aggregating multiple DT trained on bootstrap samples, while GB sequentially combines weak learners to minimize classification error through iterative residual learning. LR, although comparatively simple, provides an efficient probabilistic classification framework and has demonstrated excellent generalization capability for structured tabular datasets such as DGA measurements.

Rather than modifying the internal learning mechanisms of these classifiers, the proposed framework focuses on hyperparameter optimization, enabling each algorithm to operate under its most suitable configuration. This strategy allows objective comparison of optimization techniques while preserving the intrinsic characteristics of each learning algorithm. The principal hyperparameters optimized for each classifier are summarized in Table 2.

Table 2:Machine Learning Classifiers and Optimized Hyperparameters

Classifier	Learning Paradigm	Key Hyperparameters Optimized
SVM	Margin-based	C, Kernel, Gamma
DT	Tree-based	Max Depth, Min Samples Split, Criterion
RF	Ensemble	Number of Trees, Max Depth, Max Features
GB	Ensemble Boosting	Learning Rate, Number of Estimators, Max Depth
LR	Probabilistic	Regularization (C), Solver, Penalty

3.4 Optimization Framework

To improve the predictive performance of the selected ML classifiers, three optimization strategies were investigated: GA, PSO, and TPOT AutoML. Each optimization technique was independently applied to all five classifiers using the same training dataset and evaluation protocol, thereby enabling a fair and unbiased comparison of their optimization capabilities. The objective of each optimization strategy was to identify the optimal hyperparameter configuration that maximized the five-fold cross-validation accuracy while maintaining good generalization on the testing dataset. Figure 3 illustrates the optimization framework adopted in this study. Unlike conventional manual hyperparameter tuning, the proposed optimization framework automatically searches the hyperparameter space using different optimization principles. GA performs evolutionary optimization through selection, crossover, and mutation operations; PSO updates candidate solutions using swarm intelligence; whereas TPOT AutoML simultaneously optimizes feature selection, model selection, and hyperparameters by automatically generating machine learning pipelines. The optimized classifiers were subsequently evaluated using identical testing data and multiple performance metrics to identify the most effective optimization strategy for transformer fault diagnosis.

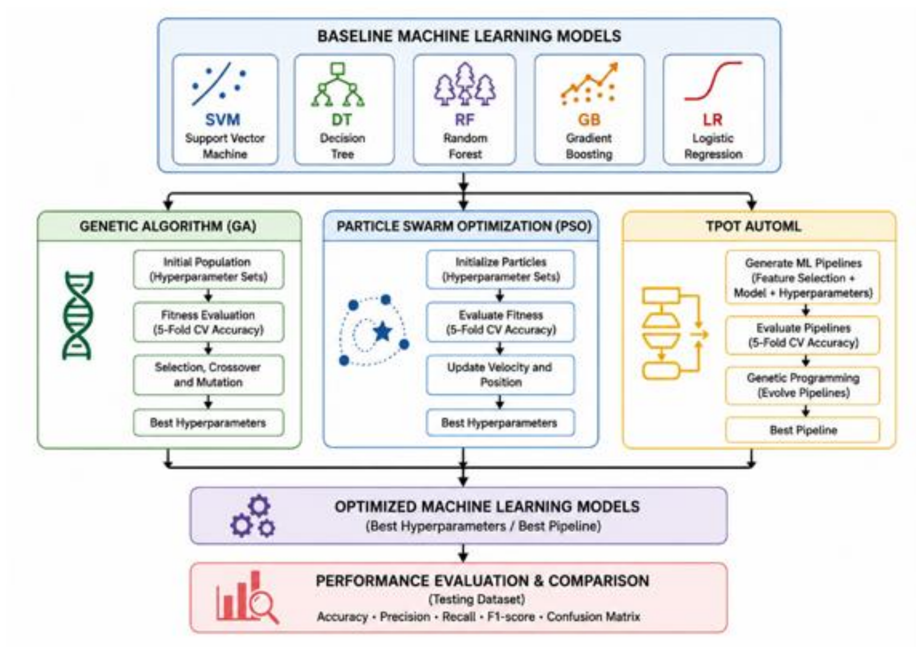

Figure 3:Unified optimization framework employed for hyperparameter tuning of machine learning classifiers.

Algorithm 1: Proposed Optimization-Driven Transformer Fault Diagnosis Framework

---

Input: DGA dataset

Preprocess dataset

Train baseline classifiers

For each optimization method (GA, PSO, TPOT)

  Optimize classifier hyperparameters

  Train optimized classifier

  Evaluate using 5-fold CV

End

Compare optimized models

Benchmark against Rogers Ratio and Duval Triangle

Output: Best optimized diagnostic model

3.4.1 Genetic Algorithm Optimization

A Genetic Algorithm (GA) was employed to optimize the hyperparameters of each machine learning classifier through an evolutionary search process. Each chromosome represented a candidate hyperparameter configuration, while the fitness function was defined as the five-fold cross-validation accuracy obtained on the training dataset. During each generation, candidate solutions were evolved using selection, crossover, and mutation operators to progressively improve classifier performance. The optimization process continued until the maximum number of generations or convergence criterion was reached. The optimized hyperparameters were subsequently used to train the final classifier, which was evaluated on the independent testing dataset. Applying GA independently to all five classifiers enabled a consistent assessment of its capability to improve transformer fault diagnosis performance.

3.4.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) was used as an alternative hyperparameter optimization strategy to explore the search space through cooperative swarm intelligence. Each particle represented a candidate hyperparameter configuration, while its position was iteratively updated according to both its individual best solution and the global best solution identified by the swarm. Similar to the GA implementation, the optimization objective was to maximize the five-fold cross-validation accuracy of each classifier. After convergence, the optimal hyperparameters obtained by PSO were used to retrain the corresponding classifier, and the resulting model was evaluated using the same testing dataset. This procedure ensured an objective comparison between evolutionary optimization and swarm-based optimization under identical experimental conditions.

3.4.3 TPOT AutoML Optimization

Unlike GA and PSO, which optimize predefined classifiers, TPOT AutoML performs automated optimization of the complete machine learning pipeline. In this study, TPOT was configured to automatically perform feature selection, classifier selection, and hyperparameter optimization using a genetic programming framework. The optimization process employed a five-fold cross-validation strategy, with a maximum search duration of 30 minutes and a maximum evaluation time of 5 minutes for each candidate pipeline. The best-performing pipeline generated by TPOT was subsequently evaluated using the independent testing dataset. This automated optimization framework enables objective comparison between manually optimized classifiers (GA and PSO) and a fully automated machine learning approach while significantly reducing human intervention during model development.

Table 3:Optimization Configuration Used in This Study

Parameter	**GA**	**PSO**	**TPOT AutoML**
Optimization Objective	CV Accuracy	CV Accuracy	CV Accuracy
Cross Validation	5-Fold	5-Fold	5-Fold
Search Strategy	Evolutionary	Swarm Intelligence	Genetic Programming
Hyperparameter Optimization	✓	✓	✓
Feature Selection	✗	✗	✓
Pipeline Optimization	✗	✗	✓

3.5 Performance Evaluation Metrics

The performance of the proposed optimization framework was evaluated using widely accepted classification metrics, including Accuracy, Precision, Recall, and F1-score. Accuracy measures the overall proportion of correctly classified samples, whereas Precision evaluates the proportion of correctly predicted positive samples among all predicted positives. Recall quantifies the ability of the classifier to correctly identify actual positive instances, while the F1-score provides the harmonic mean of Precision and Recall, enabling balanced assessment for multi-class classification. In addition to these metrics, confusion matrices were analyzed to evaluate class-wise prediction performance and identify potential misclassification patterns. These evaluation measures collectively provide a comprehensive assessment of classifier accuracy, robustness, and generalization capability.

4. Experimental Results and Discussion

4.1 Experimental Configuration

The proposed optimization-driven transformer fault diagnosis framework was implemented using Python with the Scikit-learn, TPOT, NumPy, Pandas, and Matplotlib libraries. All machine learning models were trained and evaluated using the same pre-processed DGA dataset to ensure a fair comparison. The dataset was divided into 80% training and 20% testing subsets, while five-fold cross-validation was employed during hyperparameter optimization to reduce estimation bias and improve model generalization.

Five supervised machine learning classifiers SVM, DT, RF, GB, and LR were initially developed as baseline models. Their hyperparameters were subsequently optimized using GA, PSO, and TPOT AutoML under identical experimental conditions. Model performance was evaluated using Accuracy, Precision, Recall, F1-score, and Confusion Matrix analysis. Finally, the best-performing optimized models were benchmarked against the conventional Rogers Ratio and Duval Triangle methods to evaluate the practical effectiveness of the proposed framework.

Table 4:Experimental Configuration

Parameter	Value
Programming Language	Python 3.11
ML Library	Scikit-learn
AutoML Library	TPOT
Data Analysis	NumPy, Pandas
Visualization	Matplotlib
Dataset Split	80% Training: 20% Testing
Cross Validation	5-Fold
Optimization Methods	GA, PSO, TPOT AutoML
Evaluation Metrics	Accuracy, Precision, Recall, F1-score

4.2 Baseline Machine Learning Performance

Prior to applying optimization techniques, baseline experiments were conducted to establish the diagnostic performance of the selected machine learning classifiers. Five individual classifiers and four hybrid models were evaluated using identical training and testing datasets. The baseline testing accuracies obtained by these models are presented in Table 5, while Figure 4 provides a graphical comparison of their diagnostic performance.

Table 5:Baseline Performance of Machine Learning Models

Model	Test Accuracy (%)
SVM	93.00
RF	95.00
GB	95.00
DT	92.00
LR	95.00
SVM + Fuzzy	95.00
SVM + RF	95.00
GB + DT	92.00
RF + GB	95.00

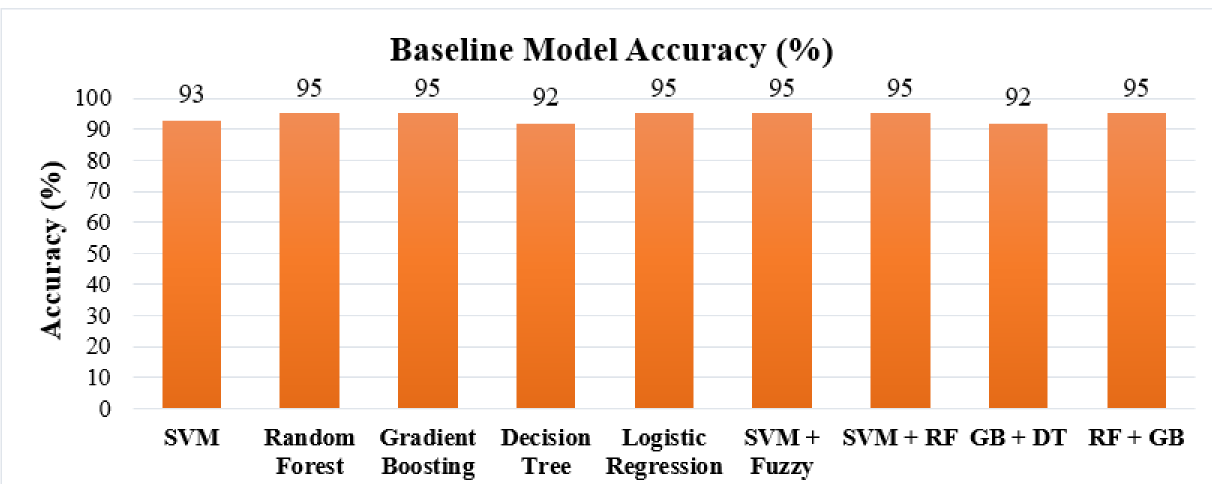

Figure 4: Baseline model accuracy comparison

The baseline results indicate that all evaluated classifiers achieved satisfactory diagnostic performance, with testing accuracies ranging from 92% to 95%. Among the individual classifiers, Random Forest, Gradient Boosting, and Logistic Regression achieved the highest baseline accuracy of 95%, whereas Decision Tree produced the lowest accuracy (92%). Similarly, the hybrid classifiers SVM + Fuzzy, SVM + RF, and RF + GB also achieved 95% accuracy, indicating that combining classifiers alone does not necessarily yield significant performance improvements.

Although the baseline models demonstrated competitive diagnostic capability, the relatively narrow accuracy range suggests that simply selecting a different classifier is insufficient to achieve substantial performance gains. This observation highlights the importance of hyperparameter optimization, which aims to improve model generalization by identifying more suitable parameter configurations rather than increasing model complexity. Consequently, the following subsections investigate the effectiveness of Genetic Algorithm, Particle Swarm Optimization, and TPOT AutoML for enhancing transformer fault diagnosis performance.

4.3 Effect of Genetic Algorithm-Based Hyperparameter Optimization

To improve the diagnostic capability of the baseline classifiers, GA was employed for hyperparameter optimization. The optimization objective was to maximize the five-fold cross-validation accuracy of each classifier while maintaining robust generalization on unseen testing data. GA independently optimized the hyperparameters of all five machine learning classifiers and four hybrid models under identical experimental conditions. The optimized results are summarized in Table 6, while the corresponding improvement over baseline performance is illustrated in Figure 5.

Table 6: Performance of GA-Optimized Models

Model	CV Accuracy	Train Accuracy	Test Accuracy	Macro Precision	Macro Recall	Macro F1
GA SVM	0.9533	0.9786	0.9600	0.89	0.89	0.89
GA RF	0.9495	0.9962	0.9467	0.90	0.82	0.85
GA GB	0.9452	1.0000	0.9478	0.88	0.85	0.86
GA DT	0.9324	0.9662	0.9300	0.83	0.80	0.81
GA LR	**0.9533**	**0.9738**	**0.9622**	**0.90**	**0.91**	**0.91**
GA SVM + Fuzzy	0.9281	0.9705	0.9300	0.82	0.76	0.78
GA SVM + RF	0.9452	1.0000	0.9522	0.90	0.84	0.87
GA GB + DT	0.9390	0.9857	0.9378	0.85	0.82	0.84
GA RF + GB	0.9476	1.0000	0.9489	0.88	0.84	0.86

Table 7: Improvement Achieved Using GA Optimization

Model	Baseline (%)	GA (%)	Improvement (%)
SVM	93.00	96.00	+3.00
RF	95.00	94.67	−0.33
GB	95.00	94.78	−0.22
DT	92.00	93.00	+1.00
LR	95.00	96.22	+1.22

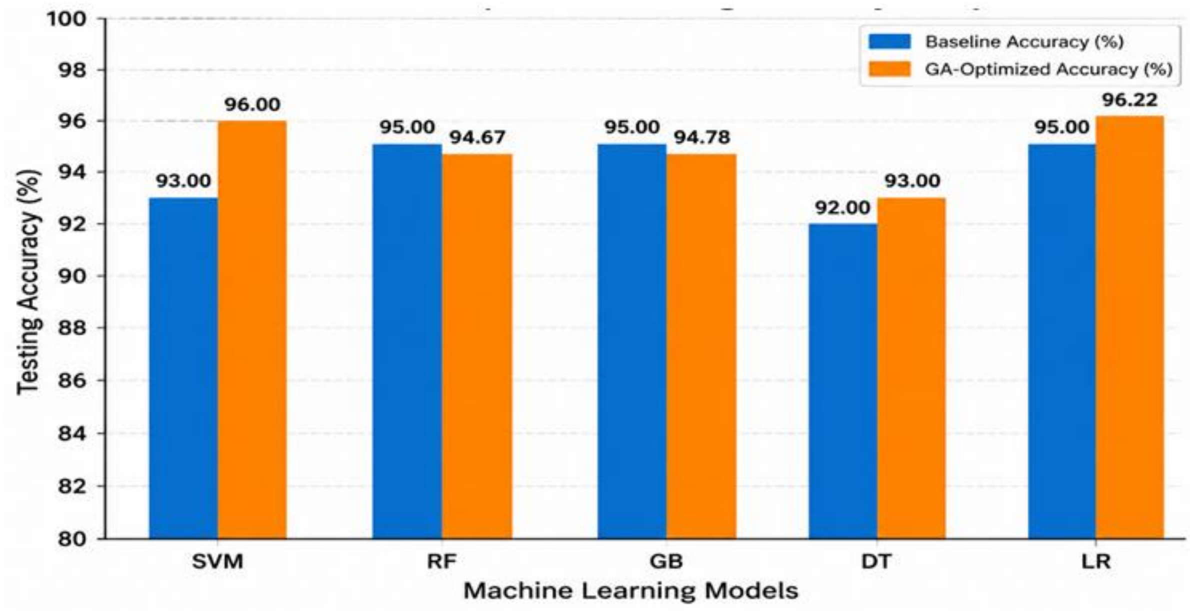

Figure 5:Baseline Vs GA-Optimized Accuracy Comparison

The results demonstrate that GA-based hyperparameter optimization produced measurable improvements for several classifiers. Support Vector Machine exhibited the largest gain, with testing accuracy increasing from 93.00% to 96.00%, representing a 3.00 percentage point improvement. Logistic Regression also benefited significantly from optimization, improving from 95.00% to 96.22%, while Decision Tree achieved a modest improvement of 1.00 percentage point.

Conversely, RF and GB exhibited marginal reductions in testing accuracy after optimization. This behaviour suggests that these ensemble models were already operating close to their optimal configurations under the baseline settings, leaving limited scope for further improvement through hyperparameter optimization. Such observations indicate that the effectiveness of optimization is model-dependent and influenced by the complexity and stability of the underlying learning algorithm.

Among all evaluated models, the GA-optimized Logistic Regression classifier achieved the highest testing accuracy (96.22%) together with the best macro precision (0.90), macro recall (0.91), and macro F1-score (0.91). The comparatively smaller gap between training and testing accuracy further indicates superior generalization capability and reduced overfitting. These findings demonstrate that evolutionary hyperparameter optimization can substantially enhance diagnostic performance without increasing model complexity, making GA-LR the most effective optimization strategy evaluated in this study.

4.4 Effect of Particle Swarm Optimization-Based Hyperparameter Optimization

The same experimental protocol used for GA-based optimization was applied to Particle Swarm Optimization (PSO) to optimize the hyperparameters of the selected machine learning classifiers. PSO was evaluated to determine whether swarm-intelligence-based search could improve diagnostic accuracy compared with the baseline models. The optimized PSO results are presented in Table 8, while the improvement over baseline performance is summarized in Table 9 & Figure 6 illustrates the comparison between baseline and PSO-optimized testing accuracies.

Table 8: Performance of PSO-Optimized Machine Learning Models

Model	**CV Accuracy (%)**	**Train Accuracy (%)**	**Test Accuracy (%)**	**Precision (Macro)**	**Recall (Macro)**	**F1-score (Macro)**
SVM	95.33	97.19	**95.44**	0.86	0.87	0.86
RF	94.95	99.62	94.67	0.90	0.82	0.85
GB	94.52	100.00	95.11	0.89	0.85	0.87
DT	93.24	96.67	92.89	0.83	0.79	0.81
LR	**95.62**	97.48	**95.67**	0.88	**0.89**	**0.89**

Table 9: Improvement Achieved Using PSO Optimization

Model	Baseline (%)	PSO (%)	Improvement (%)
SVM	93.00	95.44	+2.44
RF	95.00	94.67	−0.33
GB	95.00	95.11	+0.11
DT	92.00	92.89	+0.89
LR	95.00	95.67	+0.67

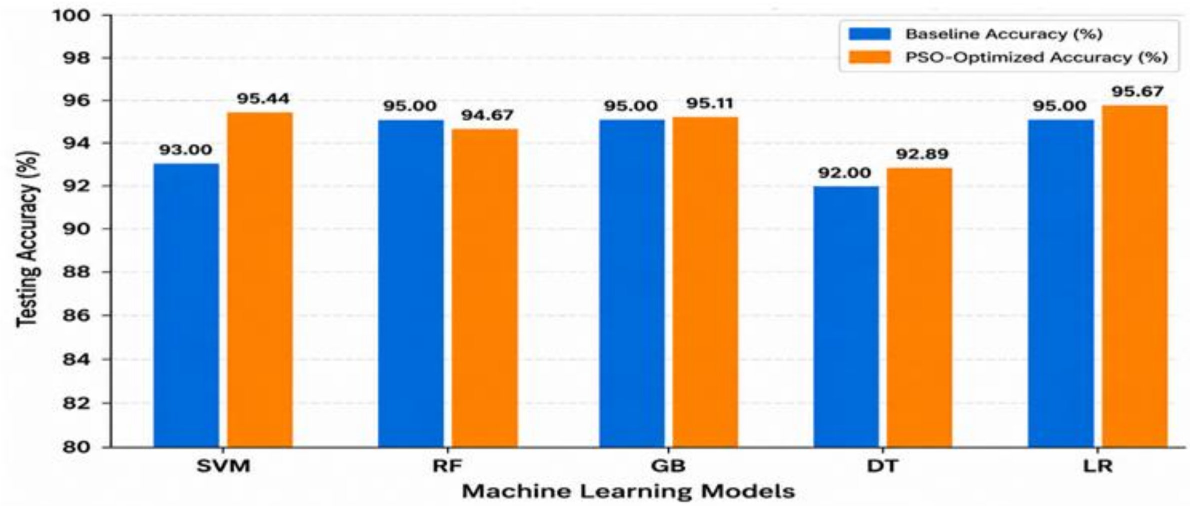

Figure 6: Baseline Vs PSO-Optimized Testing Accuracy Comparison

PSO optimization produced noticeable improvements for selected classifiers. The SVM classifier showed the highest improvement, increasing from 93.00% to 95.44%, corresponding to a 2.44 percentage point gain. Logistic Regression also improved from 95.00% to 95.67%, while Gradient Boosting showed a marginal improvement of 0.11 percentage points. Decision Tree improved modestly from 92.00% to 92.89%.

Among all PSO-optimized models, PSO-optimized Logistic Regression achieved the best overall performance with a testing accuracy of 95.67%, macro precision of 0.88, macro recall of 0.89, and macro F1-score of 0.89. Although PSO improved several models, its performance was slightly lower than the GA-optimized Logistic Regression model. These results indicate that swarm-intelligence-based hyperparameter optimization is effective for improving model generalization, but its improvement is model-dependent.

Overall, PSO provided competitive diagnostic performance and confirmed that intelligent hyperparameter optimization can enhance transformer fault classification without increasing model complexity. However, the comparatively higher performance obtained using GA suggests that evolutionary optimization provided a slightly better search capability for the selected DGA dataset.

4.5 Effect of TPOT AutoML-Based Pipeline Optimization

Unlike Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), which optimize predefined classifier hyperparameters, Tree-based Pipeline Optimization Tool (TPOT) AutoML automatically constructs complete machine learning pipelines by integrating feature selection, model selection, and hyperparameter optimization. The objective of this experiment was to evaluate whether a fully automated optimization framework could achieve competitive diagnostic performance while minimizing manual intervention. The TPOT AutoML pipeline was configured using a linear search space, five-fold cross-validation, a maximum optimization time of 30 minutes, and a maximum pipeline evaluation time of 5 minutes. The best-performing pipeline identified by TPOT consisted of LightGBM with feature selection, and its performance is summarized in Table 10.

Table 10: Performance of TPOT AutoML

Parameter	Value
Search Space	Linear
Cross Validation	5-Fold
Maximum Optimization Time	30 min

Maximum Evaluation Time	5 min
Best Pipeline	LightGBM + Feature Selection
Train Accuracy	100%
Test Accuracy	95.00%
Macro Precision	0.88
Macro Recall	0.84
Macro F1-score	0.86

Table 11: Improvement Achieved Using TPOT AutoML

Model	**Baseline (%)**	**TPOT AutoML (%)**	**Improvement (%)**
Best Baseline Model (LR)	95.00	95.00	0.00

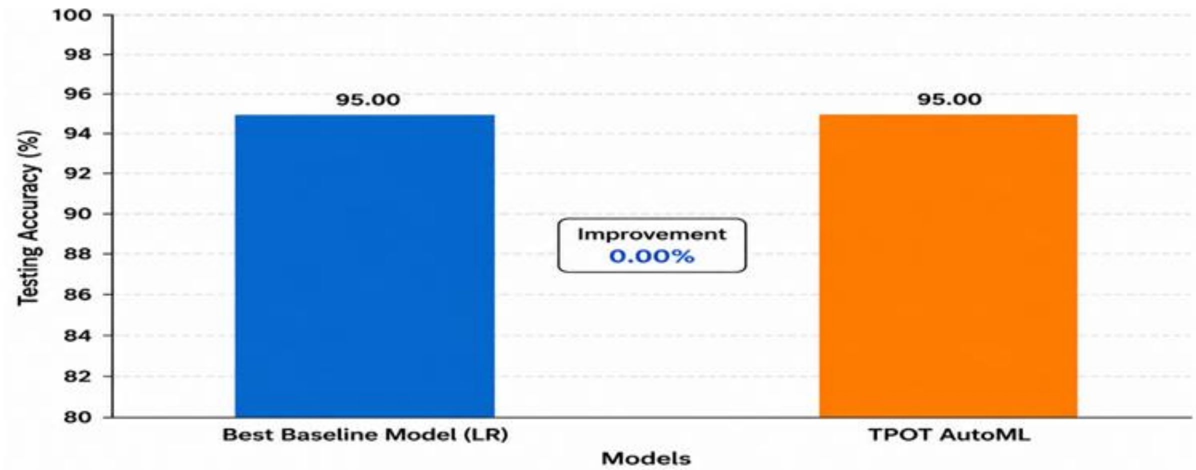

Figure 7: Baseline Vs TPOT AutoML Testing Accuracy Comparison

4.6 Comparative Evaluation with Conventional DGA Diagnostic Techniques

Following the individual evaluation of GA, PSO, and TPOT AutoML, the best-performing optimized models were benchmarked against the conventional transformer fault diagnosis techniques, namely the Rogers Ratio and Duval Triangle methods. This comparative analysis was conducted using the same DGA dataset to ensure consistency and fairness in performance evaluation. The objective was to determine whether optimization-driven machine learning models provide a significant improvement over traditional rule-based diagnostic approach. The comparative results are summarized in Table 12, while Figure 8 presents the corresponding graphical comparison.

Table 6:Comparative Performance of Conventional and Optimization-Based Diagnostic Methods

Method	**Testing Accuracy (%)**
Rogers Ratio	**44.31**
Duval Triangle	**60.70**
TPOT AutoML	**95.00**
PSO-Optimized Logistic Regression	**95.67**
GA-Optimized Logistic Regression	**96.22**

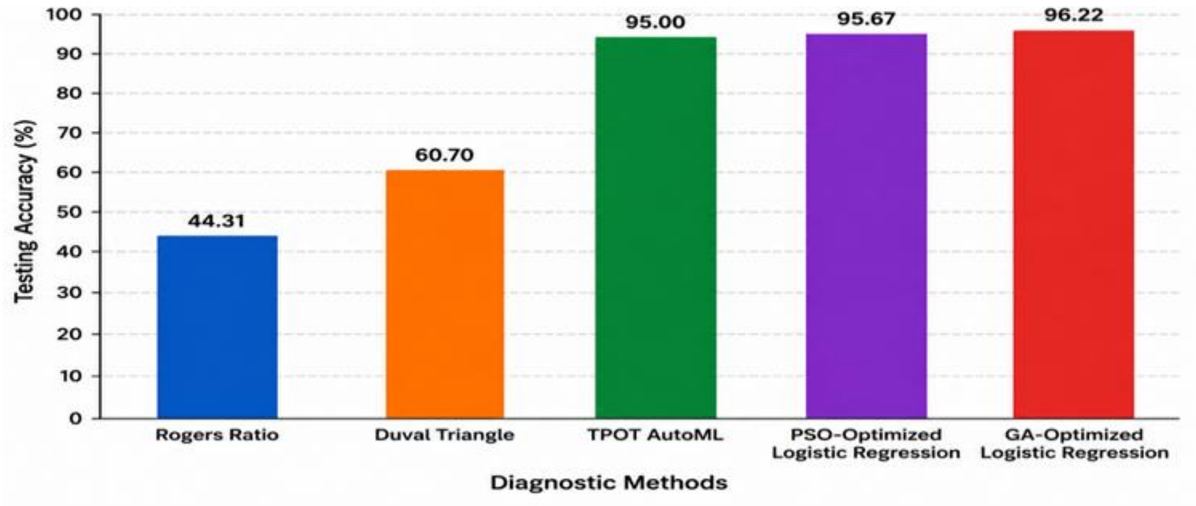

Figure 8:Accuracy Comparison of Conventional and Proposed Optimization-Based Diagnostic Methods

The comparative results clearly demonstrate the superiority of optimization-driven machine learning over conventional DGA interpretation methods. The GA-optimized Logistic Regression model achieved the highest testing accuracy of 96.22%, followed by PSO-optimized Logistic Regression (95.67%) and TPOT AutoML (95.00%). In contrast, the conventional Duval Triangle and Rogers Ratio methods achieved accuracies of only 60.70% and 44.31%, respectively.

The substantial performance gap can be attributed to the inherent limitations of conventional diagnostic techniques. Both the Rogers Ratio and Duval Triangle methods rely on predefined gas-ratio thresholds and deterministic decision boundaries, which are often inadequate for handling overlapping gas signatures, noisy measurements, and multiple simultaneous transformer faults. Consequently, their diagnostic capability decreases significantly when applied to complex real-world operating conditions.

In contrast, optimization-driven machine learning models automatically learn complex nonlinear relationships among dissolved gas concentrations and optimize classifier parameters to maximize predictive performance. Among the evaluated optimization strategies, GA produced the most effective hyperparameter configuration, enabling the Logistic Regression classifier to achieve the highest diagnostic accuracy while maintaining excellent generalization capability. The relatively small difference between the training and testing accuracies further indicates that the optimized model successfully avoided excessive overfitting.

Table 7:Performance Improvement over Conventional Diagnostic Methods

Optimized Method	Improvement over Rogers Ratio (%)	Improvement over Duval Triangle (%)
TPOT AutoML	+50.69	+34.30
PSO-Optimized LR	+51.36	+34.97
GA-Optimized LR	+51.91	+35.52

The improvement analysis further highlights the practical significance of the proposed framework. Compared with the conventional Rogers Ratio, the GA-optimized Logistic Regression model achieved an absolute improvement of 51.91 percentage points, while outperforming the Duval Triangle method by 35.52 percentage points. Similar improvements were observed for the PSO-optimized Logistic Regression and TPOT AutoML models, demonstrating that optimization-based machine learning consistently provides substantially higher diagnostic accuracy than conventional DGA interpretation techniques.

These findings validate the effectiveness of the proposed optimization-driven framework for intelligent transformer fault diagnosis. Beyond improving classification accuracy, the framework minimizes dependence on manually defined diagnostic rules and expert interpretation, making it more suitable for automated transformer condition monitoring and predictive maintenance applications in modern smart grid environments.

4.6 Discussion

The experimental results demonstrate that optimization-based machine learning significantly enhances transformer fault diagnosis compared with conventional DGA interpretation methods. Among the evaluated optimization strategies, the GA-optimized Logistic Regression model achieved the best overall performance, attaining a testing accuracy of 96.22%, followed by PSO-optimized Logistic Regression (95.67%) and TPOT AutoML (95.00%). These findings indicate that evolutionary optimization is more effective in identifying optimal hyperparameter configurations for the selected DGA dataset.

The baseline experiments further revealed that hyperparameter optimization is not equally beneficial for all classifiers. While SVM and Logistic Regression exhibited noticeable performance improvements after optimization, Random Forest and Gradient Boosting showed only marginal changes, suggesting that these ensemble models were already operating close to their optimal configurations. This highlights that the effectiveness of optimization depends on the intrinsic characteristics of the underlying classifier.

A comparison with the Rogers Ratio and Duval Triangle methods confirmed the superiority of the proposed optimization-driven framework. Unlike conventional rule-based techniques that rely on fixed gas-ratio thresholds, the optimized machine learning models effectively captured complex nonlinear relationships among dissolved gas concentrations, resulting in substantially higher diagnostic accuracy. These results demonstrate the potential of optimization-driven machine learning to support intelligent transformer condition monitoring and predictive maintenance in modern power systems.

5. Conclusion & Future Scope

This study presented an optimization-driven machine learning framework for intelligent transformer fault diagnosis using Dissolved Gas Analysis (DGA) data. The framework integrated Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and TPOT AutoML with multiple supervised machine learning classifiers to enhance diagnostic performance through automated hyperparameter optimization. Experimental results demonstrated that the GA-optimized Logistic Regression model achieved the highest testing accuracy of 96.22%, outperforming both the PSO-optimized Logistic Regression (95.67%) and TPOT AutoML (95.00%) models. Furthermore, all optimization-based approaches significantly surpassed the conventional Rogers Ratio (44.31%) and Duval Triangle (60.70%) diagnostic techniques. These findings confirm that optimization-driven machine learning provides a reliable and effective solution for transformer fault diagnosis, offering improved diagnostic accuracy, reduced dependence on manual interpretation, and enhanced suitability for predictive maintenance in modern power systems.

Future research will focus on validating the proposed framework using large-scale real-time transformer monitoring data collected from utility substations to further assess its robustness under practical operating conditions. The framework can also be extended by incorporating deep learning, Explainable Artificial Intelligence (XAI), and online learning techniques to improve model interpretability and adaptive fault diagnosis. Additionally, integrating the proposed approach with IoT-enabled monitoring systems, digital twins, and smart grid infrastructures could facilitate continuous transformer health assessment and support next-generation predictive maintenance applications.

References

1. H. Cao, C. Zhou, Y. Meng, J. Shen, and X. Xie, “Advancement in transformer fault diagnosis technology,” Front. Energy Res., vol. 12, p. 1437614, Jul. 2024, doi: 10.3389/FENRG.2024.1437614/FULL.
2. Y. Zhang, Y. Tang, Y. Liu, and Z. Liang, “Fault diagnosis of transformer using artificial intelligence: A review,” Front. Energy Res., vol. 10, p. 1006474, Sep. 2022, doi: 10.3389/FENRG.2022.1006474/TEXT.
3. A. Kumar, A. R. Saxena, and V. Shrivastava, “DGA Based Transformer Fault Analysis Using Machine Learning,” 2024 International Conference on Electrical, Electronics and Computing Technologies, ICEECT 2024, 2024, doi: 10.1109/ICEECT61758.2024.10738926.
4. M. Alenezi, F. Anayi, M. Packianather, and M. Shouran, “Enhancing Transformer Protection: A Machine Learning Framework for Early Fault Detection,” Sustainability 2024, Vol. 16, vol. 16, no. 23, Dec. 2024, doi: 10.3390/SU162310759.
5. G. Santamaria-Bonfil, G. Arroyo-Figueroa, M. A. Zuniga-Garcia, C. G. Azcarraga Ramos, and A. Bassam, “Power Transformer Fault Detection: A Comparison of Standard Machine Learning and autoML Approaches,” Energies 2024, Vol. 17, vol. 17, no. 1, Dec. 2023, doi: 10.3390/EN17010077.
6. T. Hubana and M. Hodzic, “Artificial Intelligence Based Fault Detection and Classification in Power Systems: An Automated Machine Learning Approach,” 2024 23rd International Symposium INFOTEH-JAHORINA, INFOTEH 2024 - Proceedings, 2024, doi: 10.1109/INFOTEH60418.2024.10496038.
7. Suwarno, H. Sutikno, R. A. Prasojo, and A. Abu-Siada, “Machine learning based multi-method interpretation to enhance dissolved gas analysis for power transformer fault diagnosis,” Heliyon, vol. 10, no. 4, p. e25975, Feb. 2024, doi: 10.1016/J.HELIYON.2024.E25975.
8. G. Xu, M. Zhang, W. Chen, and Z. Wang, “Transformer Fault Diagnosis Utilizing Feature Extraction and Ensemble Learning Model,” Information 2024, Vol. 15, vol. 15, no. 9, Sep. 2024, doi: 10.3390/INFO15090561.

9. V. M. N. Dladla and B. A. Thango, “Fault Classification in Power Transformers via Dissolved Gas Analysis and Machine Learning Algorithms: A Systematic Literature Review,” Applied Sciences 2025, Vol. 15, vol. 15, no. 5, Feb. 2025, doi: 10.3390/APP15052395.
10. R. A. Prasojo et al., “Precise transformer fault diagnosis via random forest model enhanced by synthetic minority over-sampling technique,” Electric Power Systems Research, vol. 220, p. 109361, Jul. 2023, doi: 10.1016/J.EPSR.2023.109361.
11. S. Rao, G. Zou, S. Yang, and S. Barmada, “A feature selection and ensemble learning based methodology for transformer fault diagnosis,” Appl. Soft Comput., vol. 150, p. 111072, Jan. 2024, doi: 10.1016/J.ASOC.2023.111072.
12. A. Nanfak, E. Samuel, I. Fofana, F. Meghnefi, M. G. Ngaleu, and C. Hubert Kom, “Traditional fault diagnosis methods for mineral oil-immersed power transformer based on dissolved gas analysis: Past, present and future,” IET Nanodielectrics, vol. 7, no. 3, pp. 97–130, Sep. 2024, doi: 10.1049/NDE2.12082;WEBSITE:WEBSITE:IETRESEARCH;ISSUE:ISSUE:DOI.
13. S. Zhang, L. Fu, and J. Wang, “Research on Machine Learning-Based Fault Diagnosis and Prediction Techniques for Power Transformers,” Proceedings of 2024 2nd International Conference on Signal Processing and Intelligent Computing, SPIC 2024, pp. 699–703, 2024, doi: 10.1109/SPIC62469.2024.10691397.
14. 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 2025 15:1, vol. 15, no. 1, pp. 8263-, Mar. 2025, doi: 10.1038/s41598-024-78293-7.
15. O. Attallah, R. A. Ibrahim, and N. E. Zakzouk, “A lightweight deep learning framework for transformer fault diagnosis in smart grids using multiple scale CNN features,” Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 14505-, Apr. 2025, doi: 10.1038/s41598-025-96290-2.
16. A. A. Adekunle et al., “Multiclass Fault Diagnosis in Power Transformers Using Dissolved Gas Analysis and Grid Search-Optimized Machine Learning,” Energies 2025, Vol. 18, vol. 18, no. 13, Jul. 2025, doi: 10.3390/EN18133535.
17. M. A. A. Putra, Suwarno, and R. A. Prasojo, “Improving Transformer Health Index Prediction Performance Using Machine Learning Algorithms with a Synthetic Minority Oversampling Technique,” Energies 2025, Vol. 18, vol. 18, no. 9, May 2025, doi: 10.3390/EN18092364.
18. J. Wang, X. Zhang, F. Zhang, J. Wan, L. Kou, and W. Ke, “Review on Evolution of Intelligent Algorithms for Transformer Condition Assessment,” Front. Energy Res., vol. 10, p. 904109, May 2022, doi: 10.3389/FENRG.2022.904109/ENDNOTE.
19. “IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers,” Art. no. C57.104-2008, Jun. 2019, doi: 10.1109/IEEESTD.2019.8890040.
20. L. Londo et al., “Hybrid Dissolved Gas-in-Oil Analysis Methods,” Journal of Power and Energy Engineering, vol. 3, no. 6, pp. 10–19, Jun. 2015, doi: 10.4236/JPEE.2015.36002.
21. R. R. Rogers, “Ieee And IEc Codes To Interpret Incipient Faults In Transformers/ Using Gas In Oil Analysis,” IEEE Transactions on Electrical Insulation, vol. EI-13, no. 5, pp. 349–354, 1978, doi: 10.1109/TEI.1978.298141.
22. M. Duval and A. DePablo, “Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases,” IEEE Electrical Insulation Magazine, vol. 17, no. 2, pp. 31–41, Mar. 2001, doi: 10.1109/57.917529.
23. M. Duval, “A review of faults detectable by gas-in-oil analysis in transformers,” IEEE Electrical Insulation Magazine, vol. 18, no. 3, pp. 8–17, May 2002, doi: 10.1109/MEI.2002.1014963.
24. H. Cao, C. Zhou, Y. Meng, J. Shen, and X. Xie, “Advancement in transformer fault diagnosis technology,” Front. Energy Res., vol. 12, p. 1437614, Jul. 2024, doi: 10.3389/FENRG.2024.1437614/FULL.
25. Y. Zhang, Y. Tang, Y. Liu, and Z. Liang, “Fault diagnosis of transformer using artificial intelligence: A review,” Front. Energy Res., vol. 10, p. 1006474, Sep. 2022, doi: 10.3389/FENRG.2022.1006474/BIBTEX.
26. M. Alenezi, F. Anayi, M. Packianather, and M. Shouran, “Enhancing Transformer Protection: A Machine Learning Framework for Early Fault Detection,” Sustainability 2024, Vol. 16, Page 10759, vol. 16, no. 23, p. 10759, Dec. 2024, doi: 10.3390/SU162310759.
27. A. Dhiman and R. Kumar, “Fault Diagnosis of a Transformer using Fuzzy Model and PSO optimized SVM,” 2023 IEEE 8th International Conference for Convergence in Technology, I2CT 2023, 2023, doi: 10.1109/I2CT57861.2023.10126260.
28. S. Zhang, L. Fu, and J. Wang, “Research on Machine Learning-Based Fault Diagnosis and Prediction Techniques for Power Transformers,” Proceedings of 2024 2nd International Conference on Signal Processing and Intelligent Computing, SPIC 2024, pp. 699–703, 2024, doi: 10.1109/SPIC62469.2024.10691397.