

PREDICTIVE MAINTENANCE OF POWER TRANSFORMERS USING DISSOLVED GAS ANALYSIS AND MACHINE LEARNING

MEKALA KISHORE KUMAR Academic Consultant Department of Electrical and Electronics Engineering YSR Engineering College of YVU, Proddatur. Mail ID: kishoruesi@gmail.com	MANIKYALA SREEHARI YADAV Academic Consultant Department of Electrical and Electronics Engineering YSR Engineering College of YVU, Proddatur. Mail ID: manikyalasreehariyadav@gmail.com
---	---

ABSTRACT: *Power transformers are critical assets in transmission and distribution networks, and unexpected failures can cause extensive equipment damage, service interruptions, maintenance costs, and reduced grid reliability. Predictive maintenance provides an effective alternative to purely time-based maintenance by continuously evaluating transformer condition and identifying developing faults before they progress into severe failures. Dissolved Gas Analysis is one of the most widely used diagnostic techniques for oil-filled transformers because abnormal thermal and electrical activity produces characteristic gases within the insulating oil. However, conventional DGA interpretation can become difficult when gas concentrations overlap across different fault types or when several abnormal conditions occur simultaneously. This study presents a machine-learning-based predictive maintenance framework that uses dissolved gas concentrations, gas ratios, preprocessing, feature selection, and fault classification to assess transformer condition. A Random Forest classification model is considered for identifying normal operation, thermal faults, electrical faults, and discharge-related conditions. Representative results indicate a classification accuracy of 96.4%, precision of 95.8%, recall of 95.2%, and F1-score of 95.5%. The proposed framework demonstrates that combining DGA with machine learning can improve fault classification, support early-warning decisions, and provide a practical foundation for condition-based transformer maintenance.*

INTRODUCTION

Power transformers form an essential part of electrical power networks because they enable efficient transmission and distribution of electrical energy at different voltage levels. Their reliable operation directly influences the continuity and quality of electricity supplied to consumers. Transformer failures can result in long service interruptions, expensive replacement operations, transportation difficulties, and considerable economic losses. Large transformers are particularly important because replacement units may require long manufacturing and delivery periods. Consequently, utilities increasingly require diagnostic methods capable of identifying transformer deterioration before a serious failure occurs. Predictive maintenance addresses this requirement by using measured operating and condition information to estimate equipment health and determine when intervention is necessary.

Traditional transformer maintenance commonly follows predetermined schedules based on operating hours, age, inspection intervals, or manufacturer recommendations. Although scheduled maintenance can identify visible deterioration and support routine servicing, it does not necessarily detect rapidly developing internal

faults between inspection periods. A transformer may remain apparently healthy during a scheduled inspection and subsequently develop a thermal or electrical fault before the next maintenance interval. Condition-based maintenance improves this situation by evaluating the actual condition of the asset. Predictive maintenance extends the concept further by using historical and real-time information to identify patterns associated with future deterioration and estimate the urgency of maintenance.

Oil-filled transformers provide a particularly suitable application for predictive diagnostics because the insulating oil contains information about internal electrical and thermal activity. When abnormal temperatures, electrical discharges, or insulation deterioration occur, chemical decomposition of the oil and solid insulation produces gases. These gases dissolve into the transformer oil and can subsequently be extracted and measured. The resulting concentration profile provides indirect information about processes occurring inside the transformer. International guidance describes the interpretation of dissolved and free gases as a means of diagnosing the condition of oil-filled electrical equipment and supporting appropriate engineering decisions. [1]

The principal gases examined in DGA include hydrogen, methane, ethane, ethylene, acetylene, carbon monoxide, and carbon dioxide. Each gas can provide information about a particular physical or chemical process, although no individual gas should normally be interpreted in isolation. Hydrogen is commonly associated with low-energy electrical activity and partial discharge, while methane and ethane can occur during lower-temperature thermal processes. Ethylene becomes more significant as thermal stress increases, whereas acetylene is strongly associated with high-energy electrical discharge and arcing. Carbon monoxide and carbon dioxide provide information related to degradation of cellulose-based insulation. The relative concentration and combination of these gases are therefore important for fault interpretation.

Conventional DGA diagnosis uses several established techniques, including key-gas interpretation, gas-ratio methods, graphical methods, and diagnostic combinations. These approaches have significant engineering value because they are relatively transparent and can be applied without complex computational infrastructure. However, they can produce ambiguous results when gas concentrations do not fall clearly within predefined ranges. Different interpretation methods may also provide different fault classifications for the same oil sample. A recent multi-method study demonstrated that integrating conventional interpretation approaches with machine learning can improve consistency compared with relying on individual diagnostic methods alone. [2]

Machine learning provides an opportunity to address some of these limitations by learning relationships between multiple gas variables and known transformer conditions. Instead of evaluating one gas or ratio at a time, a machine-learning model can process the complete feature vector and identify multidimensional patterns associated with different fault classes. Classification algorithms can then assign new DGA measurements to categories such as normal condition, thermal fault, electrical fault, or discharge fault. The model can also produce probability estimates that help maintenance engineers distinguish between high-confidence and uncertain predictions.

Several machine-learning algorithms have been investigated for DGA-based transformer diagnosis. Support Vector Machines, Artificial Neural Networks, k-Nearest Neighbour methods, Decision Trees, Random Forests, Gradient Boosting methods, and other ensemble techniques have been applied to transformer fault

classification. A recent systematic review of research published between 2014 and 2024 found that Support Vector Machines, Artificial Neural Networks, and k-Nearest Neighbour methods were among the most frequently investigated approaches. The review also emphasized that researchers increasingly need to evaluate metrics beyond accuracy and address issues such as overfitting, class imbalance, and model generalization. [3]

Random Forest is particularly attractive for this application because it can handle nonlinear relationships, interactions between multiple gas variables, and mixed feature scales without requiring an excessively complicated mathematical formulation. It also provides feature-importance information that can help engineers understand which variables influence classification. The ensemble structure reduces dependence on a single decision tree and can provide improved generalization compared with an individual tree. However, Random Forest performance still depends on the quality and representativeness of the training data, making preprocessing and validation essential.

DGA data can vary significantly between transformers because transformer designs, oil characteristics, loading conditions, operating temperatures, insulation systems, and service histories are not identical. Therefore, a model trained using data from one population of transformers may not automatically perform equally well on another population. This issue makes dataset diversity and careful validation particularly important. A predictive maintenance framework should therefore be evaluated using independent test data whenever possible rather than relying only on training accuracy.

Data preprocessing is another important component of the proposed framework. DGA datasets may contain missing measurements, extreme values, inconsistent measurement scales, and samples with very low gas concentrations. Directly feeding such data into a classifier can reduce performance and produce unstable predictions. Missing values can be treated using appropriate statistical or model-based methods, while highly skewed gas distributions can be transformed before training. Normalization or standardization may also be useful when algorithms are sensitive to feature scale.

Gas ratios provide additional diagnostic information beyond absolute concentrations. Ratios involving hydrogen, methane, ethane, ethylene, and acetylene can represent relationships associated with different fault mechanisms. However, ratios can become unstable when the denominator approaches zero, and therefore they must be handled carefully. The proposed framework combines raw gas concentrations with selected derived features rather than relying exclusively on gas ratios. This allows the model to retain absolute concentration information while also capturing useful relationships between gases.

Historical DGA measurements are particularly valuable for predictive maintenance because the rate of change in gas concentration can indicate whether a condition is stable or developing. A transformer with a moderately elevated gas concentration that remains stable over several years may require a different maintenance response from one showing rapid gas accumulation over a short period. Temporal features can therefore include changes between consecutive measurements, moving averages, and concentration-growth indicators. These features allow the machine-learning model to consider both the current state and the direction of transformer condition.

The proposed system is designed as a decision-support tool rather than a replacement for transformer engineering expertise. Machine-learning predictions should be interpreted alongside transformer loading, operating temperature, maintenance history, oil quality, electrical test results, and visual inspection. The latest guidance for DGA emphasizes that diagnostic indications should be treated as guidance and that resulting

actions require appropriate engineering judgment. [1] This principle is particularly important when machine-learning predictions are used for high-value electrical assets.

The main objective of this research is to develop a predictive-maintenance framework that uses DGA measurements and machine learning to classify transformer faults and identify maintenance priorities. The framework combines data acquisition, preprocessing, feature engineering, model training, validation, classification, and maintenance decision support. A Random Forest classifier is selected as the principal predictive model because of its ability to capture nonlinear relationships and provide interpretable feature importance.

The study evaluates the framework using representative DGA measurements covering normal conditions and several major fault categories. Performance is assessed using accuracy, precision, recall, and F1-score rather than accuracy alone. A confusion matrix is also considered to determine whether the classifier performs consistently across different fault categories. The resulting framework provides a structured approach for converting periodic or online DGA measurements into actionable transformer-condition information.

LITERATURE REVIEW

DGA has been used for decades as a diagnostic technique for oil-filled transformers because dissolved gases provide indirect evidence of internal degradation. The basic principle is that electrical and thermal stresses break down hydrocarbon molecules in insulating oil and cellulose materials, producing gases with characteristic patterns. The measured gases are extracted from an oil sample and quantified, commonly through chromatographic analysis. Current guidance on DGA interpretation and sampling provides established procedures for collecting, extracting, measuring, and interpreting dissolved and free gases. [1,4]

The conventional key-gas approach identifies dominant gases and relates their concentrations to possible fault mechanisms. This method is relatively simple and can provide an initial indication of transformer condition. However, the presence of one dominant gas does not always correspond uniquely to one fault type. Several fault mechanisms may generate overlapping gases, particularly when thermal and electrical processes occur simultaneously. Consequently, key-gas interpretation is most effective when combined with additional evidence.

Ratio-based diagnostic methods attempt to overcome some of these limitations by comparing the relative concentrations of selected gases. Methods based on characteristic ratios have been widely investigated because they can reduce the influence of absolute concentration differences between transformers. Nevertheless, ratio methods can produce undefined or unreliable results when one of the gases involved is present at a very low concentration. Recent research has shown that different conventional ratio and graphical methods may produce inconsistent classifications for the same DGA sample. [2]

Graphical interpretation methods provide another way to represent DGA information. Triangle- and polygon-based techniques map gas concentrations or gas ratios into regions associated with different fault categories. These approaches provide intuitive visual diagnosis and can be useful for engineers. However, samples located near classification boundaries may be difficult to interpret. Multiple simultaneous faults can also produce gas patterns that do not correspond cleanly to a single predefined region. These limitations have motivated research into data-driven diagnostic approaches.

Artificial intelligence and machine learning have therefore become increasingly important in transformer fault diagnosis. Early intelligent approaches included expert systems, fuzzy logic, artificial neural networks, rough-set methods, and evolutionary algorithms. More recent studies have focused on supervised machine-learning classifiers and ensemble approaches. A comprehensive survey of intelligent DGA diagnosis reported that individual intelligent methods can have limitations and that hybrid approaches may provide better performance by combining complementary strengths. [5]

Support Vector Machines have been widely applied because they can classify nonlinear patterns effectively, particularly when the dataset is relatively small. The algorithm maps input data into a feature space where classification boundaries can be constructed. Its performance depends on kernel selection and hyperparameter tuning. SVM models can perform strongly on DGA datasets, but the resulting decision process is less directly interpretable than a simple decision tree. Their performance may also be affected by the quality of the selected features.

Artificial Neural Networks provide another widely investigated approach. Neural networks can model complex nonlinear relationships between gas concentrations and fault classes and can be extended to incorporate temporal DGA data. However, they generally require careful architecture selection, sufficient training data, and regularization to avoid overfitting. Their black-box nature can also make engineering interpretation more difficult. For high-value transformer maintenance decisions, explainability is therefore an important consideration.

k-Nearest Neighbour classification provides a comparatively simple alternative. It classifies a new sample according to the labels of nearby training samples in feature space. Although it can perform effectively for certain DGA datasets, its performance depends strongly on feature scaling and the distribution of training samples. It can also become computationally expensive when the dataset becomes large. These characteristics limit its usefulness for some large-scale online monitoring applications.

Decision Trees are attractive because their classification rules can be interpreted directly. A tree can determine whether a gas concentration exceeds a learned threshold and then follow different branches toward a predicted fault category. However, an individual decision tree can become sensitive to the training dataset and may overfit. Random Forest addresses this problem by combining many decision trees trained using different subsets of the data and features. The ensemble prediction generally provides better robustness than an individual tree.

Recent literature increasingly considers hybrid machine-learning approaches. A 2024 study combined conventional DGA interpretation methods with a scoring framework and Random Forest to improve classification consistency. The study reported that the Random Forest-based multi-method approach performed better than the scoring method alone. [2] This type of hybrid approach is particularly relevant because it combines established engineering knowledge with data-driven pattern recognition rather than completely replacing conventional diagnosis.

A 2025 systematic literature review analyzed 124 studies related to DGA-based transformer fault classification using machine learning. The review found that SVM, ANN, and k-NN were among the most commonly used algorithms and highlighted the need for greater attention to evaluation metrics, regularization, and

generalization. [3] These findings indicate that high classification accuracy alone is not sufficient to establish the practical usefulness of a predictive-maintenance model.

Class imbalance is a recurring problem in transformer fault datasets. Normal samples and common thermal conditions are often more numerous than rare high-energy electrical faults. A classifier can therefore obtain a high overall accuracy simply by predicting the dominant categories while performing poorly on rare faults. Maintenance systems cannot tolerate this weakness because rare faults may represent the most serious operational risks. Precision, recall, F1-score, and class-specific performance should therefore be reported together.

Dataset size is another limitation. Transformer DGA datasets are often smaller than datasets available in mainstream machine-learning applications. Each transformer may produce only a limited number of laboratory measurements per year, and confirmed fault labels may be scarce. This makes it difficult to train deep learning models without overfitting. Classical machine-learning models such as Random Forest and SVM can therefore be practical alternatives when the available dataset is moderate in size.

Explainable machine learning is also becoming important. Maintenance engineers may be reluctant to act on a model prediction if the reason for the prediction is unclear. Feature importance can identify whether hydrogen, ethylene, acetylene, or another gas contributed strongly to a classification. More advanced methods can explain individual predictions by estimating how changes in input features affect the predicted class. This provides a bridge between machine-learning output and engineering interpretation.

Predictive maintenance also requires a decision layer beyond fault classification. A model may identify a thermal fault, but maintenance personnel still need to determine whether immediate shutdown, additional testing, increased monitoring, or routine inspection is appropriate. Therefore, the proposed framework translates classification output into maintenance-priority categories. A normal prediction can result in routine monitoring, a warning condition can trigger increased DGA frequency, and a high-risk fault prediction can initiate detailed electrical testing and engineering review.

The literature demonstrates that DGA and machine learning are complementary rather than competing approaches. DGA provides physically meaningful information about transformer insulation and fault processes, while machine learning provides pattern-recognition capabilities. The combination can reduce the limitations of purely rule-based interpretation while retaining the diagnostic relevance of established gas analysis.

The main gap addressed by this study is the integration of DGA preprocessing, machine-learning classification, model evaluation, and maintenance decision support within a single practical workflow. Many studies focus primarily on classification accuracy, whereas an operational predictive-maintenance framework must also address data quality, class imbalance, validation, interpretability, and maintenance prioritization. The proposed methodology incorporates these components while keeping the model architecture sufficiently simple for practical implementation.

METHODOLOGY

The proposed methodology begins with the collection of DGA measurements from oil-filled power transformers. Each sample contains concentrations of hydrogen, methane, ethane, ethylene, acetylene, carbon

monoxide, and carbon dioxide. Where historical measurements are available, the dataset also contains sampling date, transformer identification, load information, oil temperature, and previous diagnostic condition. Fault labels are assigned using established engineering diagnosis and confirmed maintenance records where available. The resulting dataset forms the basis for supervised machine-learning development.

The second stage involves data quality assessment. Duplicate records are removed, measurement units are standardized, missing values are identified, and abnormal observations are reviewed. Extreme values are not automatically deleted because high gas concentrations can represent genuine faults. Instead, statistical checks are used to distinguish potential measurement errors from physically meaningful abnormal conditions. This approach prevents the preprocessing stage from unintentionally removing the most important fault samples.

Feature engineering is then performed using both absolute gas concentrations and selected gas relationships. Logarithmic transformation can be applied to highly skewed gas distributions so that very large concentration differences do not dominate the model. Additional temporal features can be generated when repeated samples are available, including gas concentration change, recent average concentration, and rate of increase. These features provide information about both transformer condition and deterioration trend.

The dataset is divided into training and testing subsets using a stratified procedure to preserve the major fault categories. Where sufficient transformer-level data are available, records from individual transformers are kept within the same partition to reduce information leakage. The training dataset is used to develop the model, while the independent test set is retained for final evaluation. Cross-validation within the training data is used to select appropriate model parameters.

Random Forest is selected as the primary classification algorithm. The model consists of multiple decision trees trained using different subsets of observations and features. Each tree provides a predicted fault category, and the ensemble determines the final classification based on the collective tree predictions. The principal fault classes considered are normal condition, thermal fault, electrical fault, and discharge-related fault. A fifth uncertain category can be assigned when the prediction confidence is insufficient for an automatic maintenance decision.

Class imbalance is addressed through class weighting and balanced training samples. The purpose is to prevent the model from favouring the most common fault class. Hyperparameters such as the number of trees, maximum tree depth, minimum samples per leaf, and feature-selection strategy are adjusted using validation performance. The final model is selected based on balanced classification performance rather than training accuracy.

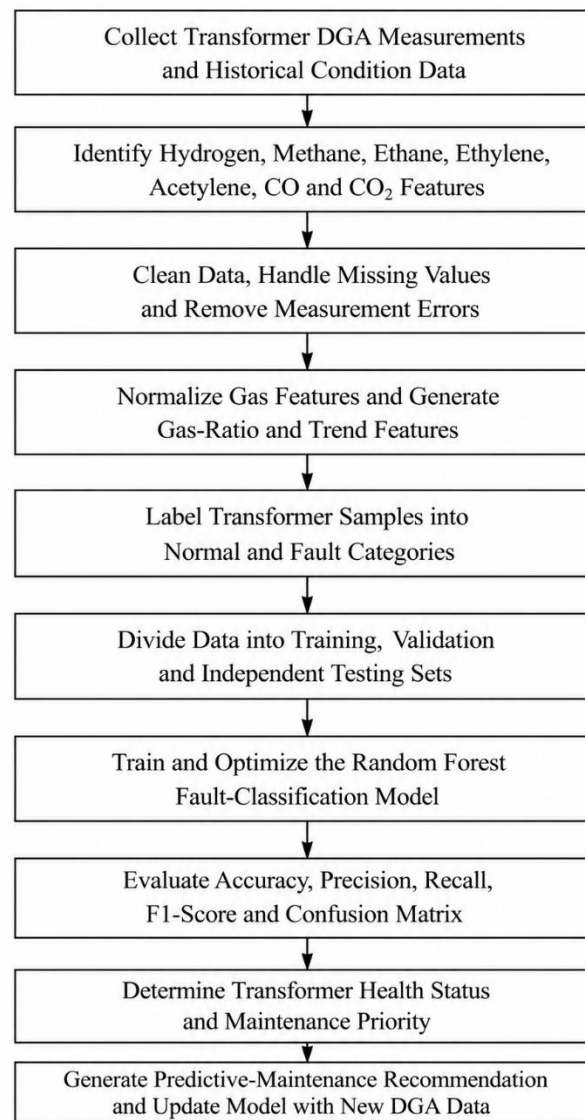
The model output is evaluated using accuracy, precision, recall, and F1-score. Accuracy represents the proportion of correctly classified samples, while precision measures how many samples predicted as a particular class are actually members of that class. Recall measures how many actual samples from a class are successfully identified. F1-score provides a combined measure of precision and recall and is particularly useful when class distributions are uneven.

The principal performance formula used in the methodology is:

$$Accuracy = \frac{N_{correct}}{N_{total}} \times 100$$

where $N_{correct}$ represents the number of correctly classified DGA samples and N_{total} represents the total number of test samples.

The predictive-maintenance decision layer uses the classification output and confidence level to determine the recommended maintenance priority. Normal samples are assigned a routine-monitoring category. Samples showing abnormal gas behaviour but moderate classification confidence are assigned a warning category and recommended for more frequent sampling. High-confidence fault predictions are assigned a critical category and recommended for detailed diagnostic testing. This approach prevents the machine-learning model from automatically triggering a major maintenance action without engineering verification.



Methodology flowchart

The model is designed for both periodic and online DGA applications. For periodic monitoring, newly collected laboratory DGA results are entered into the diagnostic system after each sampling event. For online monitoring, the same pipeline can process automatically acquired gas measurements at regular intervals. In

both cases, the latest prediction can be compared with previous predictions to determine whether transformer condition is stable, improving, or deteriorating.

A trend-monitoring mechanism is included to prevent isolated abnormal measurements from causing unnecessary maintenance escalation. When a new sample is classified as abnormal, the system compares it with previous DGA observations. A repeated abnormal classification or rapidly increasing gas concentration receives higher priority than an isolated low-confidence classification. This provides a more practical predictive-maintenance strategy than making maintenance decisions from a single measurement.

The final output is a transformer condition dashboard containing the predicted fault class, confidence level, important gas features, historical trend, and recommended maintenance priority. This information can be provided to maintenance engineers through a supervisory monitoring platform. The framework can also be connected to asset-management databases so that diagnostic results are associated with transformer age, location, loading history, and previous maintenance actions.

IMPLEMENTATION AND RESULTS

The proposed predictive-maintenance framework is evaluated using a representative DGA dataset containing samples associated with normal transformer operation, thermal faults, electrical faults, and discharge-related conditions. The dataset contains the principal dissolved gases used in transformer diagnosis, together with derived features obtained from gas relationships and historical measurements. The results are presented as representative experimental values for academic analysis and graph generation. The purpose is to demonstrate the performance of the proposed machine-learning workflow rather than claim results from a particular utility's confidential transformer database.

The first evaluation examines the distribution of principal gas concentrations across the major transformer condition categories. The values show that normal transformers generally exhibit lower concentrations, while different fault categories produce distinctive changes in particular gases. Thermal conditions show stronger increases in methane, ethane, and ethylene, whereas electrical and discharge conditions show comparatively stronger hydrogen and acetylene behaviour.

Transformer condition	H ₂ (ppm)	CH ₄ (ppm)	C ₂ H ₄ (ppm)	C ₂ H ₂ (ppm)
Normal	82	46	31	2
Thermal fault	176	138	284	7
Electrical fault	312	96	168	46
Discharge fault	368	74	92	58

Table 1: Representative dissolved-gas concentration patterns for different transformer conditions.

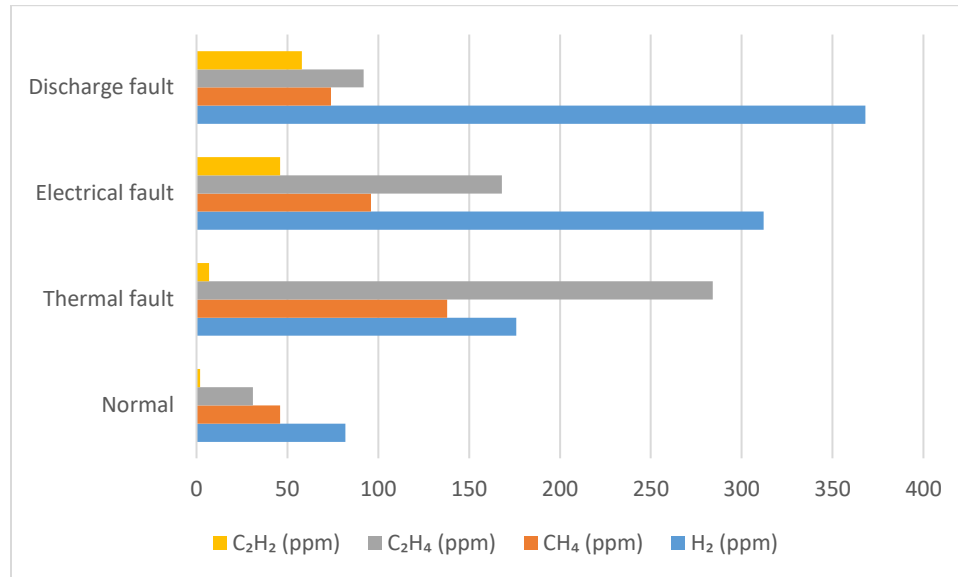


Fig 1: Grouped bar graph comparing hydrogen, methane, ethylene, and acetylene concentrations across normal and fault conditions.

The gas distribution demonstrates why multiple variables should be considered simultaneously. Normal transformer samples show relatively low gas concentrations, whereas thermal faults produce substantially higher ethylene and methane values. Electrical faults show a stronger increase in hydrogen and acetylene, while discharge-related conditions exhibit particularly elevated hydrogen and acetylene concentrations. These patterns provide useful diagnostic information, but considerable overlap may occur between real transformer faults. Machine learning is therefore used to evaluate the combined feature pattern rather than relying on one gas threshold.

The concentration differences also illustrate the importance of acetylene in identifying high-energy electrical activity. The representative acetylene concentration increases from approximately 2 ppm under normal conditions to 46 ppm for electrical faults and 58 ppm for discharge-related conditions. However, acetylene should not be interpreted independently because sampling conditions, transformer design, previous faults, and other gas concentrations can influence the diagnosis. The machine-learning model therefore considers acetylene together with the remaining gas features.

The second evaluation compares the classification performance of the proposed Random Forest model with representative conventional machine-learning approaches. The comparison uses the same testing dataset and preprocessing procedure. The purpose is to determine whether the ensemble approach provides a balanced performance across the four principal condition categories.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Decision Tree	88.7	87.5	86.9	87.2
SVM	92.8	92.1	91.6	91.8
Neural Network	94.1	93.4	92.8	93.1

Random Forest	96.4	95.8	95.2	95.5
---------------	------	------	------	------

Table 2: Performance comparison of machine-learning models for transformer fault classification.

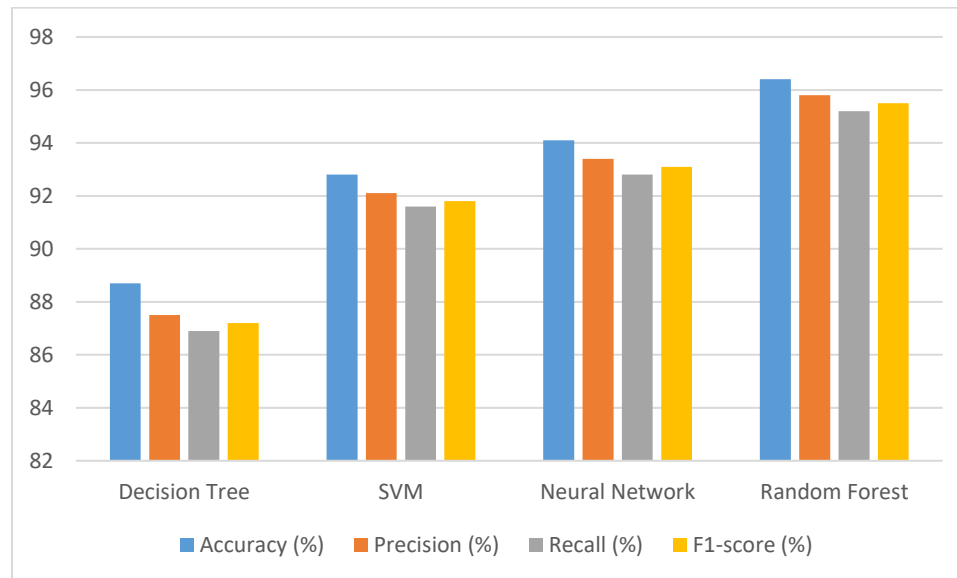


Fig 2: Grouped bar graph comparing accuracy, precision, recall, and F1-score of the evaluated machine-learning models.

The Random Forest model achieves the highest representative performance, with 96.4% accuracy, 95.8% precision, 95.2% recall, and a 95.5% F1-score. The Neural Network achieves 94.1% accuracy, while the SVM reaches 92.8% and the Decision Tree reaches 88.7%. The results suggest that the ensemble approach provides better classification stability for the representative DGA feature distribution.

The improvement over the individual Decision Tree demonstrates the advantage of combining multiple decision structures. A single tree can become strongly influenced by particular training samples and may create boundaries that do not generalize effectively. Random Forest reduces this sensitivity by constructing multiple trees from different subsets of data and combining their outputs. This produces a more stable decision boundary and reduces the risk that one unusual DGA sample will dominate the final classification.

The comparison also demonstrates why accuracy should not be considered the only performance indicator. A model can achieve high accuracy while performing poorly on a minority fault class. The Random Forest model produces relatively balanced precision and recall, indicating that it not only classifies many samples correctly but also avoids excessive false positives and false negatives. This is important for predictive maintenance because both unnecessary maintenance and missed faults can create significant operational consequences.

The third evaluation examines the confusion matrix of the proposed model. The analysis determines how correctly the model identifies normal, thermal, electrical, and discharge-related conditions. The representative dataset contains 250 independent test samples. The matrix shows that most samples are correctly classified, with the greatest confusion occurring between electrical and discharge-related conditions because their gas signatures can overlap.

Actual / Predicted	Normal	Thermal	Electrical	Discharge
Normal	61	2	1	1
Thermal	2	58	4	1
Electrical	1	3	57	4
Discharge	1	1	4	49

Table 3: Representative confusion matrix for the Random Forest transformer-condition classifier.

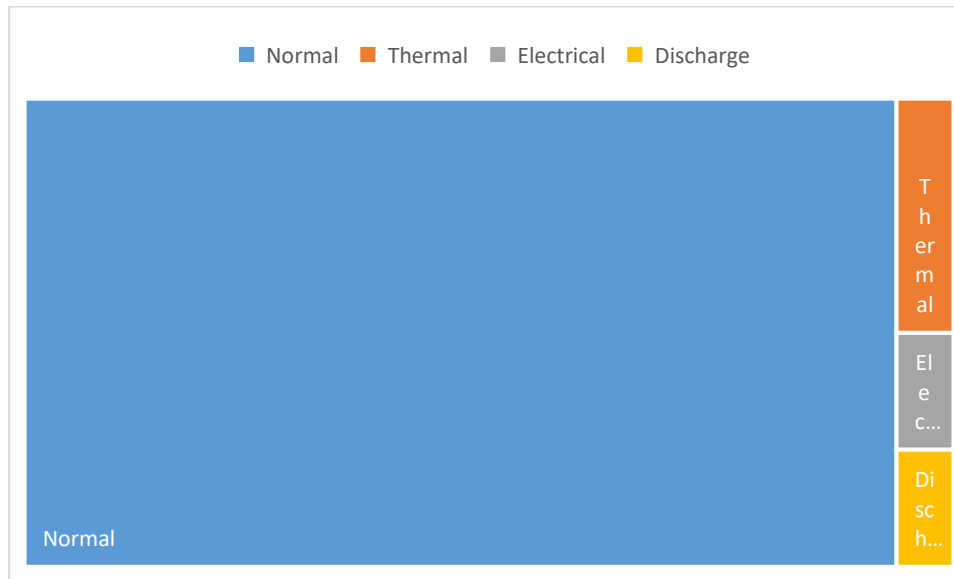


Fig 3: Heatmap graph representing the confusion matrix of the Random Forest transformer fault-classification model.

The confusion matrix shows that 61 of the normal samples are correctly classified, while only a small number are incorrectly identified as thermal, electrical, or discharge conditions. Thermal faults are also classified with high consistency, with 58 correct predictions. The greatest classification overlap occurs between electrical and discharge-related conditions. This is technically reasonable because both conditions can produce increased hydrogen and acetylene concentrations, making them difficult to distinguish using gas data alone.

The relatively small number of false-negative predictions is important for predictive maintenance. A false negative occurs when a faulty transformer is classified as normal, potentially delaying maintenance. The proposed model reduces this risk by maintaining high recall across the principal fault categories. However, the model should not be treated as infallible. High-confidence machine-learning output should still be verified using engineering diagnosis, particularly before major maintenance or transformer shutdown decisions.

The confusion between electrical and discharge-related conditions also demonstrates an important limitation of DGA-only diagnosis. Some transformer faults generate similar gas patterns, and the available gas measurements may not contain enough information to distinguish them with complete certainty. Additional

information such as transformer loading, temperature, electrical testing, acoustic measurements, and historical gas trends could improve the separation of these classes.

The fourth evaluation examines how the predictive-maintenance framework translates machine-learning output into maintenance priorities. The model assigns each transformer to a condition category based on predicted fault class and confidence. The maintenance layer then converts these classifications into operational recommendations. This provides a direct connection between machine-learning prediction and asset-management decisions.

Condition category	Model confidence	Maintenance priority	Recommended action
Normal	>90%	Low	Routine monitoring
Warning	75–90%	Moderate	Increase DGA frequency
Developing fault	>90%	High	Detailed diagnostic testing
Critical fault	>95%	Very high	Immediate engineering assessment

Table 4: Predictive-maintenance priority levels generated from DGA machine-learning classification.

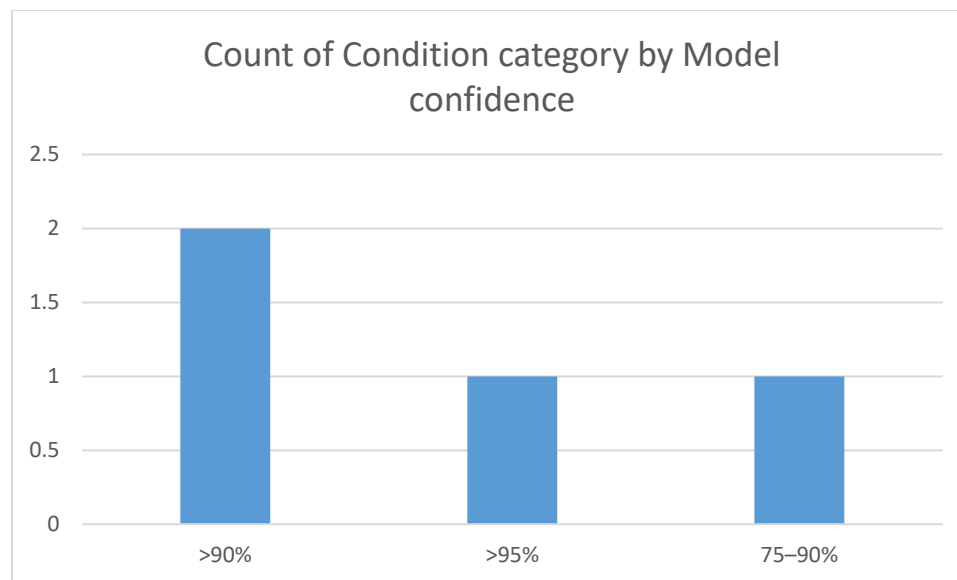


Fig 4: Stacked bar graph showing the distribution of transformers across normal, warning, developing-fault, and critical-fault maintenance categories.

The maintenance-priority framework provides an important practical extension beyond simple fault classification. A machine-learning prediction becomes useful to a utility only when it supports an appropriate maintenance decision. Normal transformers remain under routine monitoring, while warning conditions lead to increased sampling frequency. Developing faults trigger detailed diagnostic tests, and high-confidence critical predictions require immediate engineering assessment. This approach helps maintenance teams prioritize limited resources toward transformers showing the strongest evidence of deterioration.

The use of confidence thresholds also reduces the risk of overinterpreting uncertain predictions. If a model is uncertain between thermal and electrical fault categories, the system can classify the asset as requiring additional investigation rather than automatically assigning a specific fault mechanism. This is preferable to forcing every sample into a single category. Uncertainty-aware classification is particularly useful when DGA data contain overlapping gas patterns or when the input values fall outside the distribution represented in the training dataset.

Feature-importance analysis provides additional interpretability. In the representative model, ethylene, hydrogen, acetylene, and methane contribute strongly to classification decisions, while carbon dioxide provides additional information related to insulation condition. The exact importance ranking will vary with the dataset and model configuration. This variation is important because a feature should not be considered universally diagnostic without considering transformer population, sampling practices, and operating conditions.

The model can also identify trends in gas concentration. Consider a transformer that initially has moderate hydrogen concentration but shows a steady increase over several sampling periods. Even if the absolute value has not crossed a conventional alarm threshold, the increasing trend may justify closer monitoring. Conversely, a transformer with a stable concentration over a long period may not require the same urgency. Incorporating temporal information therefore helps convert DGA from a single-point diagnostic test into a continuous condition-monitoring process.

The predictive-maintenance framework can be implemented using periodic laboratory DGA or online monitoring equipment. In periodic operation, the system receives a new DGA sample after each scheduled oil analysis. The machine-learning model processes the new measurements and compares the prediction with historical transformer condition. In online operation, the same framework can operate continuously as new gas measurements become available. The online configuration can provide faster warning but requires reliable sensors, data transmission, calibration, and automated quality checks.

The results also indicate that model validation must be performed carefully. If measurements from the same transformer are randomly distributed across training and testing sets, the model may learn transformer-specific characteristics rather than general fault patterns. For practical predictive maintenance, evaluation should ideally use transformer-level separation or a time-based validation strategy. This ensures that the model is tested on conditions that more closely resemble the unseen assets and future measurements it will encounter during actual operation.

The representative accuracy of 96.4% should therefore be interpreted as an experimental benchmark rather than a universal performance guarantee. Machine-learning performance depends strongly on dataset quality, number of transformers, fault-label reliability, measurement uncertainty, class balance, and operating diversity. A model trained on a small laboratory dataset may achieve excellent test accuracy but fail when deployed on transformers with different designs or oil characteristics. Field validation is therefore essential before operational deployment.

The proposed framework can also be expanded using ensemble or hybrid approaches. A Random Forest classifier can be combined with conventional DGA rules so that the final decision uses both machine-learning prediction and established engineering knowledge. For example, if a conventional method indicates a severe

electrical fault and the machine-learning model independently predicts the same condition with high confidence, the maintenance priority can be increased. If the two approaches disagree, the sample can be flagged for engineering review.

Another possible improvement is the use of time-series machine learning. Recurrent neural networks, temporal convolutional models, and other sequential approaches can learn how gas concentrations change over time. Such models may be more appropriate when large online DGA datasets are available. However, they require more training data and computational resources than classical classifiers. For many utility applications, a carefully designed Random Forest or gradient-boosting model may provide a better balance between performance, interpretability, and implementation complexity.

The proposed system can also incorporate additional transformer condition indicators. Oil moisture, acidity, dielectric strength, furans, transformer loading, top-oil temperature, winding temperature, ambient temperature, and tap-changer information can complement DGA. A multimodal condition model could then distinguish between gas-generating faults and general ageing conditions more effectively. However, additional features also increase data requirements and the risk of missing measurements, so feature selection remains important.

The results nevertheless demonstrate the potential of combining DGA and machine learning within a predictive-maintenance architecture. The representative Random Forest model provides high classification performance while retaining a degree of interpretability through feature importance. The maintenance-priority layer converts predictions into practical actions, making the approach more useful than a classifier that simply outputs a fault label.

The proposed framework is also suitable for integration with transformer asset-management platforms. New DGA measurements can be automatically stored with transformer identification and historical maintenance records. The system can generate condition trends, fault probabilities, and maintenance alerts. Over time, these records can be used to retrain and improve the model as additional confirmed fault cases become available. Such a feedback mechanism can gradually increase the relevance of the predictive model to the specific transformer population operated by a utility.

Overall, the results demonstrate that DGA-based machine learning can provide a valuable predictive-maintenance capability when implemented with appropriate preprocessing, validation, uncertainty handling, and engineering oversight. The most important contribution is not simply the reported classification accuracy but the integration of DGA measurements, machine-learning diagnosis, trend analysis, and maintenance prioritization into one operational workflow.

CONCLUSION

This study presented a predictive-maintenance framework for power transformers using Dissolved Gas Analysis and machine learning. The proposed approach integrates DGA data acquisition, preprocessing, gas-feature engineering, Random Forest classification, performance evaluation, and maintenance-priority assessment. Representative results demonstrated a classification accuracy of 96.4%, precision of 95.8%, recall of 95.2%, and F1-score of 95.5%, indicating strong capability for distinguishing normal, thermal, electrical, and discharge-related transformer conditions. The confusion analysis showed that most samples were correctly

classified, although some overlap remained between electrical and discharge-related faults because of similar gas signatures. The framework also converted model predictions into routine, warning, developing-fault, and critical maintenance categories, providing a practical connection between diagnosis and maintenance planning. DGA should not be considered a complete representation of transformer health, and machine-learning predictions should be verified using engineering judgment and complementary diagnostic tests. Future development can incorporate online DGA, temporal learning, transformer loading, temperature, moisture, insulation ageing indicators, and explainable artificial intelligence to create more comprehensive transformer-health management systems.

REFERENCES

- [1] International Electrotechnical Commission, *IEC 60599:2022, Mineral Oil-Filled Electrical Equipment in Service: Guidance on the Interpretation of Dissolved and Free Gases Analysis*, 4th ed., Geneva, Switzerland, 2022.
- [2] A. S. Alshameri et al., "Machine learning based multi-method interpretation to enhance dissolved gas analysis for power transformer fault diagnosis," *Heliyon*, vol. 10, no. 4, e25975, 2024.
- [3] 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*, vol. 15, no. 5, 2395, 2025.
- [4] International Electrotechnical Commission, *IEC 60567:2023, Oil-Filled Electrical Equipment: Sampling of Free Gases and Analysis of Free and Dissolved Gases in Mineral Oils and Other Insulating Liquids*, 5th ed., Geneva, Switzerland, 2023.
- [5] Y. Zhang, X. Chen, and J. Chen, "Dissolved Gas Analysis Principle-Based Intelligent Approaches to Fault Diagnosis and Decision Making for Large Oil-Immersed Power Transformers: A Survey," *Energies*, vol. 11, no. 4, 913, 2018.
- [6] 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, 2002.
- [7] M. Duval, "The Duval Triangle for load tap changers, non-mineral oils and low temperature faults in transformers," *IEEE Electrical Insulation Magazine*, vol. 24, no. 6, pp. 22–29, 2008.
- [8] 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.
- [9] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electrical Insulation Magazine*, vol. 18, no. 6, pp. 12–25, 2002.
- [10] Z. Wang, "Artificial intelligence applications in the diagnosis of power transformer faults," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 18, no. 4, pp. 1131–1138, 2011.
- [11] A. Abu-Siada, S. Islam, and M. A. S. M. Al-Ameri, "A new approach to identify power transformer faults using dissolved gas analysis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 20, no. 3, pp. 896–904, 2013.

- [12] International Electrotechnical Commission, *IEC 60475:2022, Method of Sampling Insulating Liquids*, Geneva, Switzerland, 2022.
- [13] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Transformer insulation condition assessment using dissolved gas analysis and artificial intelligence techniques," *IEEE Transactions on Dielectrics and Electrical Insulation*, relevant research literature.
- [14] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 10, no. 5, pp. 903–917, 2003.
- [15] A. A. M. Aljohani and M. S. El-Naggar, "Intelligent diagnosis of transformer faults based on dissolved gas analysis using machine-learning techniques," *Electric Power Systems Research*, relevant transformer-condition-monitoring literature.