SRTTU

Journal of Electrical and Computer Engineering Innovations

(JECEI)

Journal homepage: http://www.jecei.sru.ac.ir

JECEI

Research paper

Ensemble Learning Algorithm for Power Transformer Health Assessment Using Dissolved Gas Analysis

K. Gorgani Firouzjah *, J. Ghasemi

Department of Electrical Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran.

Article Info

Article History:

Received 12 November 2024
Reviewed 25 December 2025
Revised 26 January 2025
Accepted 29 January 2025

Keywords:

Power transformer
Health assessment
Ensemble learning
Dissolved gas analysis
Predictive maintenance

*Corresponding Author's Email Address: *k.gorgani@umz.ac.ir*

Abstract

Background and Objectives: Power transformer (PT) health assessment is crucial for ensuring the reliability of power systems. Dissolved Gas Analysis (DGA) is a widely used technique for this purpose, but traditional DGA interpretation methods have limitations. This study aims to develop a more accurate and reliable PT health assessment method using an ensemble learning approach with DGA.

Methods: The proposed method utilizes 11 key parameters obtained from real PT samples. In this way, synthetic data are generated using statistical simulation to enhance the model's robustness. Twelve different classifiers are initially trained and evaluated on the combined dataset. Two novel indices (a risk index and an unnecessary cost index) are introduced to assess the classifiers' performance alongside traditional metrics such as accuracy, precision, and the confusion matrix. An ensemble learning method is then constructed by selecting classifiers with the lowest risk and cost indices.

Results: The ensemble learning approach demonstrated superior performance compared to individual classifiers. The learning algorithm achieved high accuracy (99%, 92%, and 86% for three health classes), a low unnecessary cost index (6%), and a low misclassification risk (16%). This result indicates the effectiveness of the ensemble approach in accurately detecting PT health conditions.

Conclusion: The proposed ensemble learning method provides a reliable and accurate assessment of PT health using DGA data. This approach effectively optimizes maintenance strategies and enhances the overall reliability of power systems by minimizing misclassification risks and unnecessary costs.

Introduction

Power transformers (PTs) are considered one of the essential elements in power supply systems. The main task of this equipment is to manage voltage levels to ensure compatibility between generation sources and various electrical loads [1]. In this regard, the reliable performance of PTs essential to ensure a continuous power supply. This equipment (as a critical network asset) to reduce the risk of unexpected failures requires regular maintenance and condition assessment. In fact, carrying out these preventive measures improves the reliability of the network. They also reduce downtime and enhance the overall stability of the power system [2]. In order to protect personnel and equipment, it is essential to identify PT issues in their early phases, thereby preventing costly repairs and mitigating potential safety hazards. Modern diagnostic tools and intelligent decision-making procedures drive predictive maintenance techniques, which are essential for ensuring a continuous energy supply and improving the functionality of PTs. Asset managers that proactively identify potential problems and meticulously plan maintenance measures can successfully reduce downtime, extend PT operating lifespans, and significantly reduce overall operational expenses. In addition to enhancing grid stability, this innovative approach fosters a more robust and sustainable energy system. The literature on PT fault diagnosis provides most of the techniques for identifying early faults and preventing catastrophic failures. The researchers employ a variety of methods in this procedure. Among these are vibration data analysis, thermographic image processing [3], dissolved gas analysis (DGA), and acoustic emission analysis [4]. However, each of these diagnostic techniques has its own unique advantages and drawbacks [5]. One popular and successful method for diagnosing PT faults is DGA [6].

This strategy is based on the fact that different types of PT faults produce distinct gases in the oil. Accordingly, assessing the quantities of these gases can identify the kind and extent of the fault [7]. Many metrics obtained from dissolved gases in oil have led to the identification of key gases. These include carbon monoxide (CO), carbon dioxide (CO_2), hydrogen, methane (CH_4), ethylene (C_2H_4), ethane (C_2H_6), and acetylene (C_2H_2) [8]. However, we have also exploited other practical features like the oil's moisture content and its insulation breakdown voltage level. As indicated before, DGA is a robust oil assessment instrument. Key gas method (KGM) is a popular DGA data analysis tool [9]. KGM, a DGA subsidiary, specializes on fault-related gases. This strategy improves fault type and severity identification. Measurements of dissolved gases in PT insulating oil indicate fault occurrence and type. This method leverages the fact that every inaccuracy leads to different ratios of oil to dissolved gas. It is possible to ascertain the nature and severity of the fault by comparing the gas concentrations to the threshold values [10]. The characteristics of these gases can identify various faults. For instance, internal faults can result in the production of hydrogen, CO, CO_2, and CH_4 d. However, cellulose faults have the ability to produce other gases, such as CH_4, C_2H_6, and C_2H_4 [10]. On the other hand, higher hydrocarbon temperatures can lead to higher CH_4 and C_2H_6 concentrations. Moreover, studies in [4] indicate that an electrical arc or partial discharge may raise the concentration of hydrogen. The IEEE C57.104-2008 standard provides useful guidelines for testing, interpreting, and decision-making of various PT faults. The aim of this standard is to standardize and improve the accuracy and reliability of the DGA method. This standard, by establishing uniform procedures and specific requirements, helps increase the efficiency and accuracy of the DGA method for detecting faults and assessing the health status of PTs.

In the literature, the DGA process includes various steps, the first of which is data preparation. This involves collecting oil samples and conducting necessary tests. The procedure is done to measure the concentration of dissolved gases. After data preparation, the concentration of gases (such as hydrogen, methane, ethylene, etc.) is identified. Then, the PT fault type is determined using various analytical methods (such as key gas ratios, Rogers ratios, IEC ratios, and the Duval triangle [11]). In the next step, the obtained results from the analytical methods are compared with real data to evaluate the accuracy and efficiency of the method. Eventually (based on the obtained results), decisions will be made regarding the health status of the PT and the necessary maintenance actions. Gas concentrations in oil must be within limits. The references [12]-[14] provide the concentration limits for gases in the DGA method. These boundaries are used for fault diagnosis and health assessment. For example, if the hydrogen concentration in the oil exceeds the permissible limit, it may indicate windings or core fault . However, interpreting DGA parameters takes careful consideration of several aspects. In fact, environment, oil type, age, and loads affect gas concentrations and fault detection. Furthermore, the relation between the concentrations of different gases and the types of fractures can be complex and nonlinear. For this reason, the use of advanced statistical methods and mathematical models is essential for the accurate interpretation of results.

Recent years have seen extensive research into advanced DGA interpretation methods. Researchers used neural networks (NNs) [15]-[17], genetic algorithms [18], and fuzzy logic [19], [20]. These approaches are promising PT fault detectors because they can learn complex data patterns and correlations [21]. Traditional DGA interpretation relies on empirical rules and expert knowledge. While these methods have been helpful in many circumstances, they may not be enough to reliably estimate transformer health under all operating conditions. Many studies have shown that classifier algorithms including support vector machines (SVM) [22], [23], k-nearest neighbors (KNN) [24], [25], random forests, decision tree, and Naïve Bayes can categorize DGA data and discover different faults. For instance, utilizing DGA data, Benmahamed et al. [26] suggested a unique method for improving the precision of transformer problem identification. They use a bat algorithm for parameter optimization, a Gaussian classifier, and a SVM classifier in their method. The proposed method were able to accurately classify six different types of faults than with traditional DGA methods. This was done by optimizing the SVM parameters and using the concentration of five combustible gases as input. In [27], they used SVM and the optimization procedure to enhance the model parameters and increase the accuracy of fault detection. Haque et al.

[28] proposed a novel method for fault diagnostics in PTs, employing DGA and a Random Forest classifier. They classified several fault types with excellent accuracy using their approach, which combines a modified Duval pentagon method with Euclidean distance characteristics and density-based clustering. Using DGA data, a study [29] assessed the diagnostic performance of Naive Bayes and the KNN algorithm for transformer oil insulation states. An approach for finding faults in oil-immersed PTs uses DGA, a mixed KNN algorithm, and a decision tree [24]. The literature review demonstrates that feature selection enhances detection process. However, some issues with the literature may make it less useful and accurate in practice. The limitations of prior studies:

- Some research has mostly looked at small datasets or certain kinds of mistakes. This problem makes the model less reliable and useful in the real world.
- Some studies have not used ensemble learning methods to improve the accuracy and stability of the models (by combining the predictions of models).
- Many studies have ignored the importance of mistakes and the costs associated with erroneously recognizing issues. This problem may lead to the selection of models that are extremely accurate but carry serious risks and high costs. For example, misdiagnosing a malfunctioning PT as functional might result in irreparable damage, whereas misdiagnosing a working PT as a challenge in high maintenance costs.

This research has presented a novel ensemble learning method to address the limitations of previous studies in PT health status assessment using DGA data. This algorithm uses a big dataset and combines 8 successful classifiers in an ensemble learning method. It aims to develop a model to determine the PT health across various scenarios. This approach considers diagnostic accuracy and error costs to make cost-effective decisions. This generates synthetic data and standardizes it to improve performance. Additionally, in order to measure the PT's health status, limit high-risk mistakes, and save unnecessary costs, we have developed two risk and unnecessary cost indicators. The technologies used in this study should improve PT condition monitoring systems and reduce unexpected breakdowns. Consequently, we have organized the paper as follows: Section 2 explains the suggested approach. Section 3 compares the performance of the proposed technique with other methods, showcasing the implementation details and evaluation results on real datasets. Results and suggestions for future work are summarized in Section 4.

Proposed Method

This section outlines a proposed methodology that employs DGA data and machine learning algorithms to assess the health status of transformers. The primary objective of this method is to address the shortcomings of conventional techniques and to enhance the precision and dependability in assessing the health condition of the transformer. The suggested technique categorizes DGA data into three clusters: Healthy, needs retesting in the future and needs immediate retesting. This classification allows power system operators to evaluate the PT health status accurately. Also, they make appropriate decisions for PT maintenance and repair. Fig. 1 shows the flowchart of the proposed algorithm. As shown, the algorithm includes multiple steps as follows. At statistical distribution extraction step, the appropriate distribution for DGA parameters (gas concentrations and their ratios) in each of the three classes is determined. The algorithm generate synthetic DGA parameters to increase the of training data size and enhance the performance of machine learning models. These data are generated using the obtained statistical distributions. The algorithm normalize both real and synthetic DGA data using an appropriate method to ensure uniformity in their scale and measurement units. The data is randomly shuffled and assigned to training and testing sets for machine learning analysis. In order to establish an effective model, twelve classifiers are trained and assessed. Consequently, an ensemble approach is then employed to further enhance the accuracy and reliability of health predictions. The following sections will provide complete descriptions of the steps.

A. Step 1: Retrieval of Initial Data

This paper outlines a method for evaluating the health of transformers using historical DGA data. These data are collected through the analysis of oil samples extracted from operational PTs and stored in a dedicated database. These include the concentrations and ratios of each of the gases, which collectively provide unique insights into the transformer's state. The eleven important parameters are then reviewed and explained in terms of their technical importance. The average breakdown voltage measures the dielectric strength of the oil, which shows how well it can handle electrical breakdowns when voltage stress is applied. The drop in breakdown voltage shows that the oil insulation is breaking down. This can be caused by contamination, oxidation, or the buildup of breakdownproducts. This makes it more likely that there will be partial discharges and, eventually, insulation failure. Moisture in transformer oil, even in small amounts (ppm), can make it less effective at insulating and can greatly reduce the oil's dielectric strength.

This can speed up the breakdown of the paper insulation, cause acids and sludge to form, and further weaken the transformer's integrity. Carbon monoxide solubilized in oil serves as the primary indication of thermal stress on cellulose insulation (paper). The elevation in CO levels signifies that the insulation is experiencing overheating, either attributable to overload, inadequate cooling, or isolated hot spots within the PT.

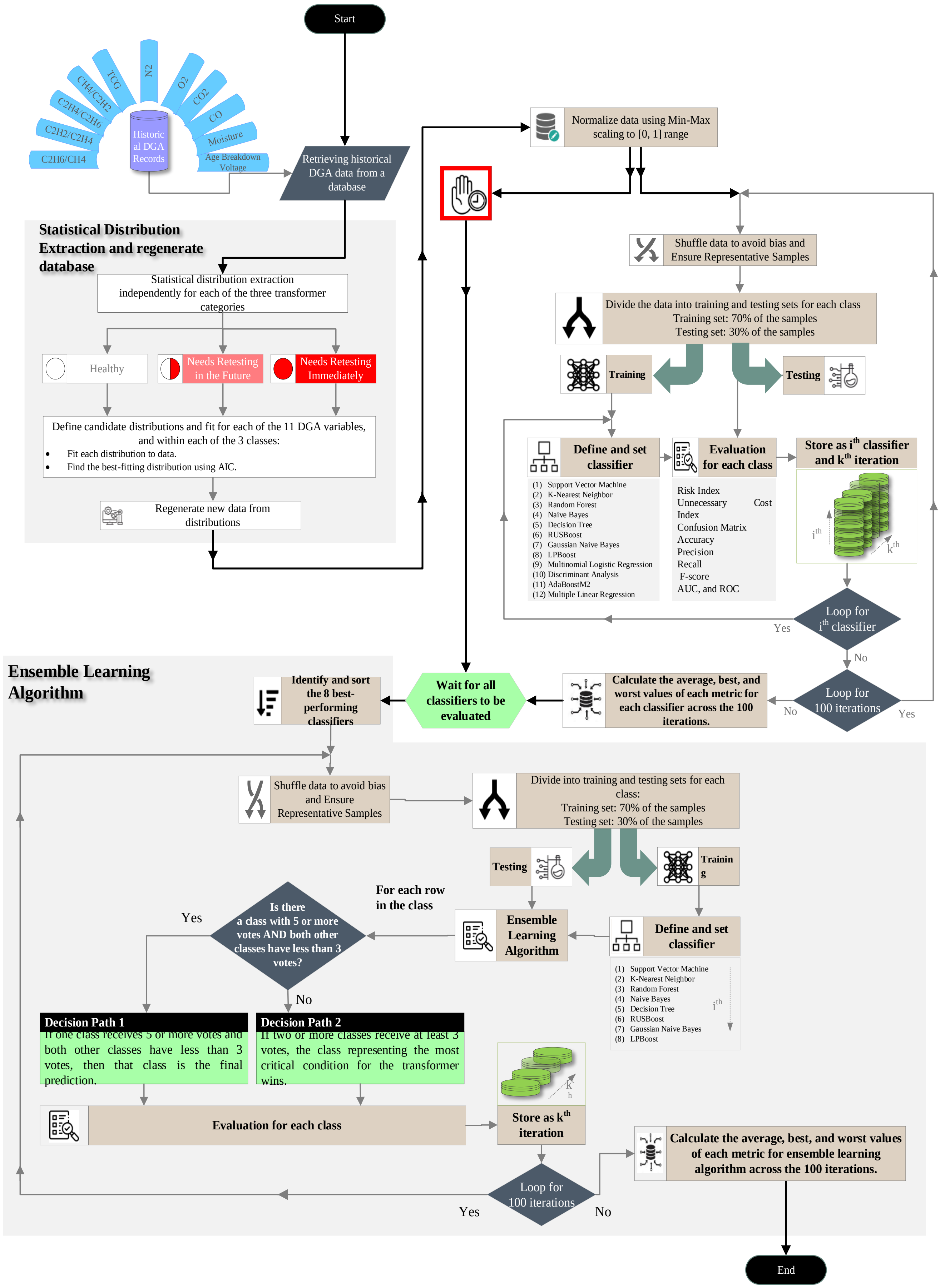

Fig. 1: Flowchart of the proposed transformer health assessment algorithm.

When paper insulation degrades, it releases CO_2 (a substance more difficult to analyze than CO). Taking into account the CO_2/CO ratio can aid in accurately determining the type and severity of the insulation issue. The oil containing oxygen poses a significant issue as it accelerates oxidation, leading to the production of acids and sludge. When oxygen enters the system, it typically indicates an air leak in the generator or conservator tank, necessitating immediate repair. While nitrogen typically serves as a background gas, it can occasionally serve as a diagnostic indicator. A high nitrogen level could mean that air is getting into the oil or that nitrogenous molecules are breaking down. Total Combustible Gases (TCG) measures the amount of combustible gas released during partial discharges, overheating, and arcing. To ascertain the nature of the mistake, the program must thus probe further into the TCG level spike. Gas ratios (CH_4/C_2H_2, C_2H_6/CH_4, C_2H_4/C_2H_6, and C_2H_2/C_2H_4) offer a detailed understanding of fault situations. Analyzing the relative concentration of different gases allows for the inference of the fault's nature (e.g., arcing, overheating, or partial discharge) and its severity. The literature frequently used these ratios in conjunction with recognized diagnostic criteria such as the Duval Triangle or IEC 60599 to assess DGA results effectively. The historical DGA data (which builds up over time and includes data from many transformers) helps learn more about how transformers work. It makes preventative maintenance easier and lowers the chance of failures, which keeps the power grid running smoothly.

B. Step 2: Statistical Distribution Extraction

In this step, the statistical distribution governing each of the 11 DGA variables is extracted separately for each of the three transformer health classes ("Healthy," " needs retesting in the future," and "needs immediate retesting "). The purpose of this procedure is to create synthetic data and increase the database's size. Synthetic data is generated to address limitations in the amount and diversity of real-world data while also reducing the cost and time required for data acquisition. With the increase in the volume of training data, various patterns are better learned by the model. As a result, the accuracy and generalizability of the model increase. Various operational conditions and different types of errors is simulated by generating synthetic data. In this way, the model is helped to perform acceptably under various conditions and to be more resilient against new data. The process of extracting statistical distributions is such that several probable statistical distributions (such as normal, log-normal, Weibull, gamma, beta, and exponential) are applied to the data of each variable in each class. Then, using the Akaike Information Criterion (AIC), the best distribution governing the data is selected. The AIC criterion, by simultaneously considering the goodness of fit of the distribution to the data and the complexity of the model, helps in selecting the best distribution. Each variable is analyzed to determine its statistical distribution within each of the three classes. This technique yields 33 unique distributions, given 11 variables and 3 classes. New DGA data is produced with the Monte Carlo approach with these 33 distributions. The data are incorporated into the primary database to augment the training dataset, hence enhancing the efficacy of machine learning models in assessing the health status of transformers.

C. Step 3: Data Normalization

The third step of the suggested algorithm is to normalize the data. This is done after getting the statistical distribution, simulating synthetic data. This step prepares data for the learning algorithm to improve detection. For this purpose, the Statnorm normalization method is used. It transforms the data based on the standard normal distribution (with a mean of zero and a standard deviation of one). Statnorm is a powerful normalization technique specifically suitable for data that do not follow a normal distribution. This method transforms the data into a standard normal distribution by using the rank transformation and then applying the inverse cumulative normal distribution function. This approach uses Statnorm's rank-based outlier elimination. This minimizes algorithm outlier sensitivity, which is especially important for DGA data (with outlier values). In addition, the data normalizing into a standard normal distribution improves the performance of some machine learning algorithms (such as SVM, k-NN, and logistic regression). These algorithms perform more effectively for data that follows a normal distribution. In order to prevent the introduction of bias, the normalization of both the training and test datasets should be same.

As mentioned above, the method builds a synthetic dataset to enhance the transformer health detection algorithm through three steps: statistical distribution extraction, synthetic data production, and data normalization. This method improves PT health assessment models by adding training data. In addition, it improves these models by conducting sensitivity analysis and simulating different situations., useful, and reliable. Transformer condition assessment improves, resulting in fewer failures.

D. Step 4: Data Preparation

This step is very important for getting data ready for programs that use machine learning. The chosen method reduces possible errors and improves the model's performance with new data (by randomizing and dividing it up). There are two important parts to the data preparation step. In this regard, the suggested method mixes up the data in a way that gets rid of any bias that might come from the order of the original dataset. This issue guarantees that both the training and testing sets (accurately) represent the entire data distribution. Secondly, it partitions the data into training and testing sets, often according to a 70/30 ratio. All three categories—"healthy," "requiring future retesting," and "requiring immediate retesting"—utilize this procedure.

E. Step 5: Classifier Training and Evaluation

At this step, the algorithm focuses on the precise training and evaluation of various classifiers to detect the health status of the transformer. The goal of this step is to select the best algorithm for classifying DGA data and accurately diagnosing the PT health. The proposed PT health assessment system uses 12 well-known machine learning classifiers [30]. The approach optimizes classifiers via hyper-parameter tweaking. It finds optimal hyper-parameter combinations using grid search. In SVM, grid search optimizes the kernel type (linear or polynomial), penalty parameter (C), and kernel coefficient (gamma). Adaptive Boosting (AdaBoostM2), Linear Programming Boosting (LPBoost), and Random Under-sampling Boosting (RUSBoost) all have multi-class versions. The number of weak learners (decision trees) is an important hyper-parameter that is optimized using cross-validation.

The cross-validation procedure optimizes the number of nearest neighbors (k) in the KNN. In addition, the algorithm adjusts the number of decision trees in a random forest based on cross-validation to achieve a balance between model accuracy and complexity. This paper evaluates many classifiers and chooses the optimal model for transformer health status detection using a cross-validation approach with 100 random repeats. The method is accurate and reliable. The method randomly splits data into 70% training and 30% testing per iteration. The results are independent of the data separated by this approach. Using training data, each classifier is trained and the grid search technique optimizes the model's hyper-parameters. The suggested method uses testing data to evaluate the accuracy, precision, and confusion matrix of the trained model. It randomly divides data between training and testing sets in each iteration . The procedure's primary loop repeats data partitioning, training, and evaluation 100 times. The primary loop calculates the mean and standard deviation of the evaluation metrics for each classifier. This procedure is done to evaluate the system performance. The advantages of this technique encompass a reduction in the variance of model performance estimate, an evaluation of model stability against fluctuations in training and testing data, and the identification of optimal model parameters and data partitioning configurations.

This paper uses two new evaluation metrics to assess the performance of classifiers in diagnosing the health status of transformers in addition to traditional criteria.

- Risk Index (primary priority):

This metric addresses the reduction of errors that pose significant risks. In practice, misidentifying a faulty transformer as healthy is a critical error. This issue has the potential to cause catastrophic consequences. The value of this metric is derived from the sum of the following two values. First, the number of PTs that truly need immediate retesting but are mistakenly classified as healthy or needing retesting in the future is determined. Then, this number is divided by the total number of PTs that truly need immediate retesting This value (in percentage) allows the algorithm to model the percentage of high-risk errors. Second, the algorithm identifies the number of transformers incorrectly classified as "healthy" but actually requiring retesting in the future. Then, the algorithm divides this number by the total number of transformers that need retesting in the future (convert to percentage). Finally, the two obtained percentages are summed to derive the error metric.

- Unnecessary Cost Index (secondary priority):

The purpose of this criterion is to reduce unnecessary expenses. In fact, unnecessary expenses are imposed on the system when a healthy transformer is wrongly classified as defective. Then, this number is divided by the total number of healthy transformers (then converted to a percentage). This method makes it possible to determine the percentage of unnecessary expenses. Additionally, the number of transformers that actually need retesting in the future but have been mistakenly classified as needing immediate retesting is determined. Then, this number is divided by the total number of transformers that actually need to be retested in the future (then converted to a percentage). Finally, we add the two obtained percentages together to derive the cost metric. In these criteria, different errors are weighted according to their importance. For instance, the error criterion assigns more weight to errors that result in misclassifying faulty transformers as healthy. The algorithm supplements conventional metrics with these criteria to evaluate the performance of classifiers from various aspects. These two new criteria allow the algorithm to select classifiers that not only have high accuracy in detecting transformer health status but also minimize high-risk errors and unnecessary costs.

F. Ensemble Learning Algorithm

At this step, the program applies an ensemble learning strategy. This procedure is a very efficient machine learning method that applies multiple classifiers at once. In fact, it combines multiple classifiers into a single, more accurate, and reliable forecast. The proposed ensemble learning based algorithm implements eight classifiers. It trained these classifiers and optimized their parameters using grid search methods and cross-validation in the preceding phases. The following are included: SVM, KNN, Random Forest, Naive Bayes, Decision Tree, RUSBoost, Gaussian Naive Bayes, and LPBoost The proposed algorithm selected these classifiers due to their exceptional performance and diversity in the classification of DGA data. The algorithm receives DGA data as input for each case. Each of the eight classifiers independently predicts the health status of the transformer based on the input DGA data and categorizes it into one of three classes. The algorithm calculates class votes (predictions) from eight classifiers. For choosing the winning class (final transformer health diagnostic), two choice paths are considered:

- Decision path 1:

If it gets 6 out of 8 votes, a class wins and becomes a transformer.

- Decision path 2:

If both classes receive at least three votes each, the class indicating a more critical condition for the transformer will be declared the winner. In other words, the priority order will be "needs immediate retesting" followed by " needs retesting in the future" and finally "healthy".

The proposed algorithm uses various metrics to evaluate the ensemble learning performance. These include accuracy, precision, and confusion matrix. These metrics encompass error rate metrics, risk factors, and unnecessary cost indexes. The algorithm The algorithm also utilized these metrics in step 5, which involves training and evaluating classifiers. The algorithm conducts the evaluation procedure 100 times to enhance confidence in the outcomes. Subsequently, it computes the mean and standard deviation of the evaluation metrics. The proposed ensemble learning technique enhances the evaluation of PT health by merging the predictions of the top eight classifiers and accounted for more significant factors in the final decision-making process. Employing this technique can diminish maintenance expenses and prolong the longevity of PTs.

Numerical Results and Analysis

This section explores the numerical outcomes of implementing the method from Section 2 in real-world scenarios. This section demonstrates how the proposed method accurately categorizes PT health using DGA data. Statistical distribution assessment, classical classifier evaluation, and group modeling are key components of this technique. Section 3.1 will carefully evaluate the statistical distributions of critical DGA parameters for each transformer health class to find data trends. This information should be utilized to generate synthetic data. Section 2 proposes a 100-fold random cross-validation method to compare the performance of twelve different machine learning classifiers, including SVM, KNN, and Random Forest. Thus, accuracy, precision, confusion matrix, and unnecessary cost indices are used. This section will test the ensemble learning model's transformer health state classification accuracy and dependability. The ensemble learning model uses eight top classifier predictions. The proposed method is compared to existing DGA interpretation methods for advantages and disadvantages.

A. Evaluation of Data Preprocessing Steps

This section will address the numerical evaluation of the first three steps of the proposed algorithm, which include "statistical distribution extraction," "synthetic data generation," and "data normalization." The goal of this step is to identify an appropriate statistical distribution for each of the 11 DGA parameters (including breakdown voltage, moisture, gas concentrations, and their ratios) within each of the three transformer health classes. The information regarding the transformer under study is presented as follows. Fig. 2 to Fig. 5 present the statistical distribution of the key features of the studied transformers. Fig. 2 shows the statistical distribution of transformer lifespans. As can be seen, most transformers have a lifespan of between 5 to 10 years (with a frequency of about 38%). Additionally, there are a few transformers with a lifespan of over 35 years. This indicates that the data includes transformers with a variety of ages, from new to old. Fig. 3 shows the statistical distribution of transformer capacities. The majority of transformers (about 70%) have a capacity between 30 to 40 MVA. A few transformers with lower capacities (10-20 MVA) and (20-30 MVA) are also present in the data. Fig. 4 shows the statistical distribution of transformer oil weights. Transformers with an oil weight of 12 to 16 tons (approximately 48%) exhibit the highest frequency. Additionally, transformers with lower oil weights (4-8 tons) and 8-12 tons are also present in the test samples. Fig. 5 shows the statistical distribution of the type of oil used in transformers. Most transformers (over 90%) use IEC-296- type oil. Only a small percentage of transformers (less than 10%) use Nynas oil. These figures show that the transformers under study cover a wide range of age, capacity, oil weight, and type of oil. It is also worth mentioning that all the transformers studied in this research have a voltage level of 63/20 kV.

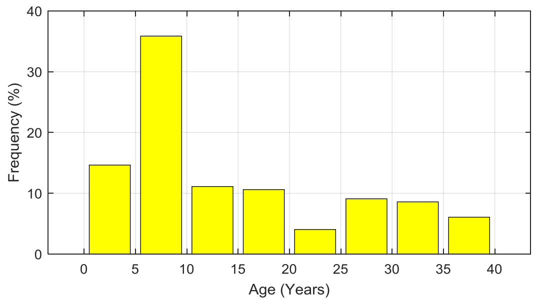

Fig. 2: Histogram of PT ages, showing that most transformers are between 5 and 10 years old.

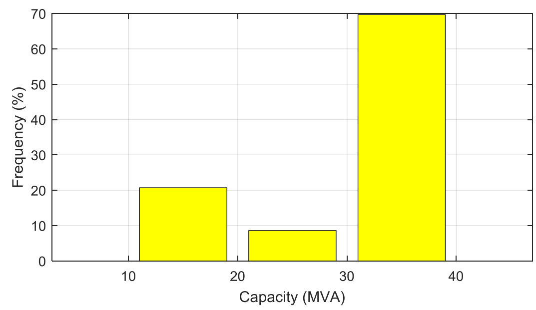

Fig. 3: Histogram of PT capacities, showing that the majority of transformers have a capacity between 30 and 40 MVA.

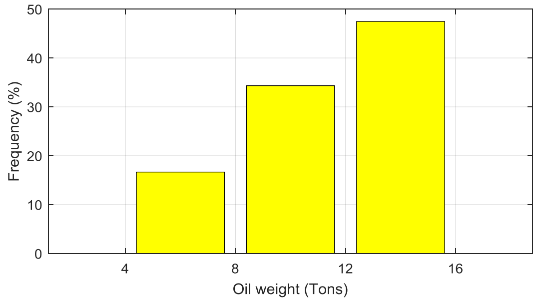

Fig. 4: Histogram of PT oil weights, showing that most transformers have an oil weight between 12 and 16 tons.

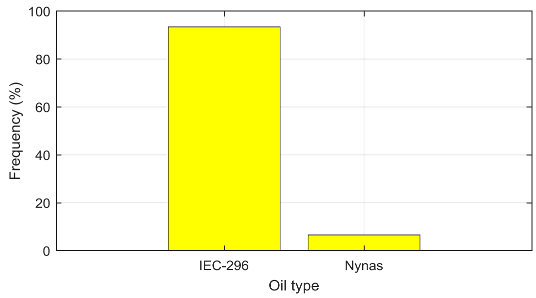

Fig. 5: Bar chart of PT oil types, showing that over 90% of transformers use IEC-296 oil.

Fig. 6 to Fig. 10 show scatter plots and histograms of DGA parameters for the three health classes of transformers so that you can look at how the data is spread out and check how well the statistical distribution extraction is working. The distributions of breakdown voltage and humidity, CO and CO_2, O_2 and N_2, and TCG and are shown in Fig. 6 to Fig. 9 for the three classes. Fig. 10 also displays the three classes' gas ratios (CH_4/C_2H_2, C_2H_6/CH_4, and C_2H_2/C_2H_4). Parameter transformer classes have varied DGA parameter distributions, as shown in these figures. Fig. 6 demonstrates a different distribution of "breakdown voltage" in the "healthy" class compared to the other two classes. Fig. 7 illustrates the distinct relationship between the "CO concentration" and "CO_2 concentration" across the three classes. Accurate modeling of key data requires identifying the appropriate statistical distribution for each. Each DGA parameter is analyzed across three transformer health classes. This ensures the synthetic data accurately represents the different health conditions of PTs.

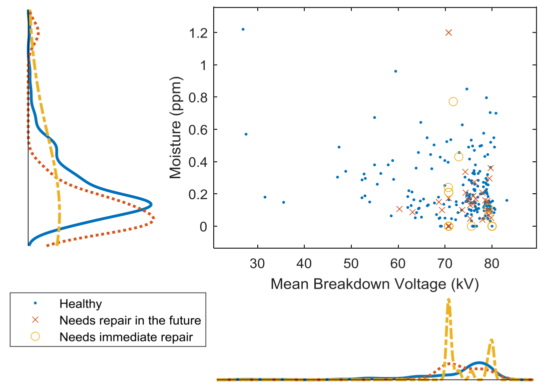

Fig. 6: Visualization of DGA data for breakdown voltage and moisture across three PT health Class.

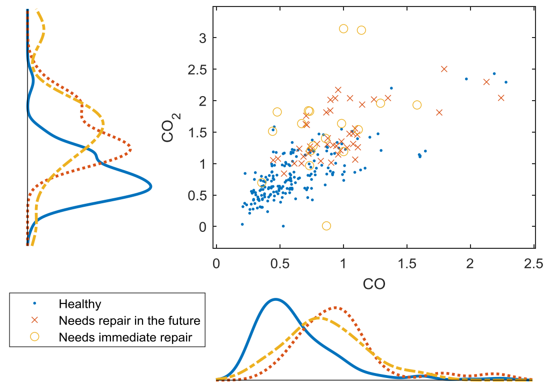

Fig. 7: Distribution of CO and CO_2 concentrations according across three PT health Class.

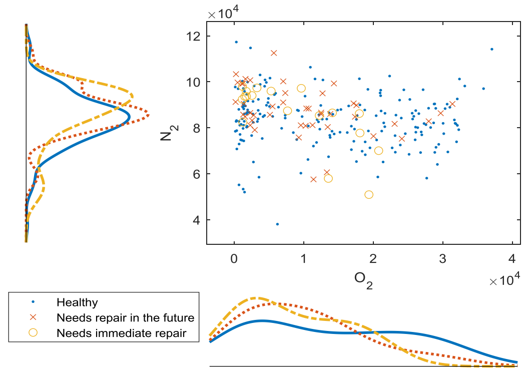

Fig. 8: Scatter plot and histogram depicting O_2 and N_2 concentrations across three PT health Class.

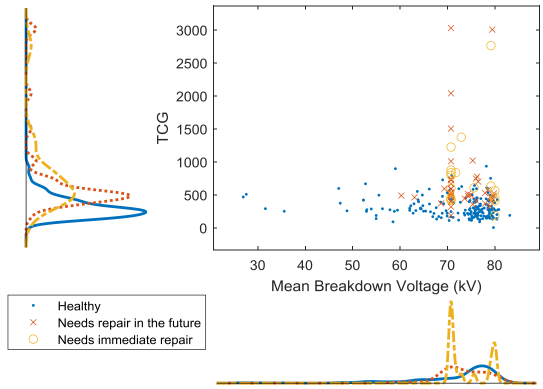

Fig. 9: Visualization of DGA data for TCG and breakdown voltage across three PT health Class.

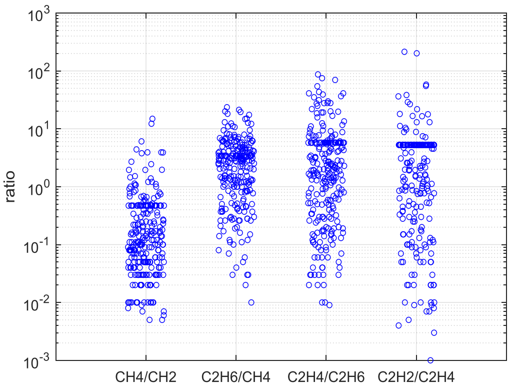

Fig. 10: Distribution of gas ratios (CH_4/C_2H_2, C_2H_6/CH_4, C_2H_4/C_2H_6, C_2H_2/C_2H_4) in DGA data.

These statistical distributions provide synthetic data and enable statistical analysis. The systematic technique encompassed the subsequent methods to ascertain the statistical distribution of each parameter: The normal distribution, log-normal distribution, Weibull distribution, gamma distribution, beta distribution, and exponential distribution were used to model the DGA data. Each class parameter's data was fitted to one of these distributions. The maximum likelihood estimation technique is used to estimate parameters. Each distribution's AIC is obtained after fitting to the data.

The exponential distribution is the best distribution for the "average breakdown voltage" parameter since it has the lowest AIC value in all classes. Similar analysis was done on the remaining other DGA parameters, identifying the optimum statistical distribution. Table 1 shows that DGA parameters in different classes have varied statistical distributions. The "Parameter" column of Table 1 indicates the 11 DGA metrics employed to evaluate transformer health. The parameters encompass breakdown voltage, moisture levels, concentrations of CO, CO2, O2, and N2 gases, along with the ratios of CH4/C2H2, C2H6/CH4, C2H4/C2H6, and C2H2/C2H4 gases ratios. The "Class" column indicates the three health classes of the transformer, which are: Healthy (H): The transformer is in a healthy condition. Needs retesting in the future (FR): The transformer is currently healthy, but it requires retesting in the future. Immediate retest required (IR): The transformer is in a critical condition and requires an immediate retest. The AIC criterion selects the best statistical distribution for each parameter in each class, as shown in the "Distribution" column. Using the Monte Carlo method, parameters for each distribution, like A (lower bound), B (upper bound), μ (mean), and π (standard deviation), are used to make fake data. As observed in the table, the statistical distribution of DGA parameters varies across different transformer health classes. This indicates that the health status of the transformer affects the statistical distribution of DGA parameters. As mentioned, this paper illustrates the impact of transformer health on the statistical distribution of DGA parameters. The parameter "CO concentration" has a log-normal distribution in the "healthy" class and a gamma distribution in the "needs immediate retesting" class. Table 1 shows that exponential and log-normal distributions are the most common for DGA parameters. The following are the optimal distributions for seven parameters in a variety of classes. Additionally, this implies that these distributions may be capable of modeling DGA data. "Data normalization" and "synthetic data generation" steps will be implemented subsequent to this.

B. Evaluation of Classifiers

This section evaluates all 12 classifiers utilized in Fig. 1. All classifiers were trained and tested with preprocessed DGA data during the evaluation process, utilizing various metrics to measure the performance of each classifier. To ensure the reliability and robustness of the assessment, a 100-fold randomized cross-validation technique was utilized.

Table 1: Statistical distributions of data across three transformer health classes

Oil Property/ Characteristic	Class	Best-Fit Distribution	Estimated Parameters of Best-Fit Distributions	
Average breakdown voltage 1 to 6	1	Weibull	A=75.33	B=12.1104
	2	Weibull	A=75.3733	B=19.4471
	3	Lognormal	μ= 4.30501	σ=0.0560327
Moisture	1	Exponential	μ=0.233556	
	2	Exponential	μ=0.145379	
	3	Exponential	μ=0.113236	
Average concentration of CO	1	Lognormal	μ=-0.576039	σ=0.463825
	2	Lognormal	μ=-0.0754555	σ= 0.341141
	3	Gamma	α=8.02887	β=0.107086
Average concentration of CO_2	1	Weibull	A=0.925048	B=2.19319
	2	Lognormal	μ=0.365075	σ =0.270727
	3	Normal	μ=1.59517	σ=0.74024
Average concentration of O_2	1	Exponential	μ=13764.9	
	2	Exponential	μ=9888.07	
	3	Exponential	μ=8486.44	
Average concentration of N_2	1	Normal	μ=83619.9	σ=11446.2
	2	Normal	μ=87500.9	σ=10307.7
	3	Weibull	A=90734.2	B=9.71992
Average TCG	1	Gamma	α=4.18118	β=73.8387
	2	Lognormal	μ=6.36514	σ=0.538263
	3	Lognormal	μ=6.46998	σ=0.588093
CH_4/CH_2 present in the oil	1	Exponential	μ=0.540598	
	2	Lognormal	μ=-2.66182	σ=1.63765
	3	Exponential	μ=0.12083	
C_2H_6/CH_4 ratio in the oil	1	Weibull	A=2.63218	B=0.853263
	2	Lognormal	μ= 1.4266	σ=0.596508
	3	Exponential	μ=5.73942	
C_2H_4/C_2H_6 ratio in the oil	1	Exponential	μ=7.13276	
	2	Exponential	μ=1.45367	
	3	Exponential	μ=1.31425	
C_2H_2/C_2H_4 ratio in the oil	1	Exponential	μ=6.48753	
	2	Exponential	μ=1.28539	
	3	Exponential	μ=1.68514	

The classifier was trained and evaluated on the training and testing sets by employing random data splitting for each cross-validation cycle. To obtain a reliable estimate of the classifier's performance, the results were subsequently averaged over 100 iterations. Each classifier's performance was measured using a variety of metrics, including accuracy, precision, and the confusion matrix. In addition to these conventional metrics, two new metrics were developed to address the importance of misclassifications in transformer fault diagnosis: the Risk Index and the Unnecessary Cost Index. The evaluation results are given in Table 2. This table shows the average and standard deviation of each assessment measure across 100 cross-validation iterations for each of the 12 classifiers. In the next subsections, we do a thorough examination of each classifier's performance, assessing its strengths and limitations across many evaluation measures. This investigation will provide useful insights into each classifier's suitability for the task of transformer defect diagnostics, as well as guidance in selecting the most successful algorithm for this application.

Table 2: Statistical distributions of data across three transformer health classes

Algorithm		Avg (%)		Max (%)		Min (%)		Avg Accuracy (%)	Confusion Matrix									Class	Average per class (%)	
		Risk	Cost	Risk	Cost	Risk	Cost		Avg			worst case			Best case				Accuracy	Precision
Ensemble Learning	Used in ensemble learning?	16	6	26	13	7	2	97	582	4	2	583	3	2	583	3	2	1	99	99
									3	122	7	5	118	9	5	118	9	2	92	92
									2	6	46	3	9	42	3	9	42	3	83	86
Support Vector Machine	Yes	23	6	34	11	13	2	97	582	4	2	585	2	1	585	2	1	1	99	99
									4	121	6	6	119	7	6	119	7	2	92	92
									3	7	44	7	9	38	7	9	38	3	85	81
K-Nearest Neighbor	Yes	24	7	34	12	13	3	96	577	9	2	579	9	0	579	9	0	1	99	98
									5	120	7	6	118	8	6	118	8	2	88	91
									3	8	43	6	10	38	6	10	38	3	83	79
Random Forest	Yes	24	2	33	5	14	0	97	586	1	1	587	1	0	587	1	0	1	99	100
									6	124	2	7	124	1	7	124	1	2	94	94
									3	8	43	5	10	39	5	10	39	3	94	80
Naive Bayes	Yes	38	35	61	53	22	18	87	552	20	16	557	17	14	557	17	14	1	98	94
									10	84	38	17	89	26	17	89	26	2	71	64
									3	14	37	5	21	28	5	21	28	3	42	69
Gaussian Naive Bayes	Yes	38	35	61	53	22	18	87	552	20	16	557	17	14	557	17	14	1	98	94
									10	84	38	17	89	26	17	89	26	2	71	64
									3	14	37	5	21	28	5	21	28	3	42	69
Decision Tree	Yes	42	9	62	18	22	4	93	575	8	5	572	13	3	572	13	3	1	97	98
									10	113	9	18	108	6	18	108	6	2	85	85
									6	13	35	8	18	28	8	18	28	3	72	66
RUSBoost	Yes	43	47	63	64	28	28	86	571	6	10	570	10	8	570	10	8	1	97	97
									14	60	58	20	63	49	20	63	49	2	77	45
									6	12	36	6	20	28	6	20	28	3	35	67
LPBoost	Yes	59	47	194	196	0	0	76	476	52	60	588	0	0	588	0	0	1	97	81
									10	86	36	124	8	0	124	8	0	2	NaN	65
									5	23	26	45	9	0	45	9	0	3	NaN	48
Multinomial Logistic Regression	No	83	9	95	15	68	4	90	576	10	2	578	8	2	578	8	2	1	97	98
									15	109	9	13	115	4	13	115	4	2	71	82
									5	34	15	7	39	8	7	39	8	3	60	28
Discriminant Analysis	No	86	14	101	21	71	7	89	577	10	1	580	8	0	580	8	0	1	95	98
									25	92	16	33	84	15	33	84	15	2	71	70
									8	28	18	6	35	13	6	35	13	3	53	33
Multiple Linear Regression	No	93	10	106	16	83	5	88	559	29	0	562	26	0	562	26	0	1	98	95
									10	115	7	15	111	6	15	111	6	2	62	87
									3	43	8	3	48	3	3	48	3	3	52	14
AdaBoostM2	No	119	1	130	2	112	0	89	583	5	0	583	5	0	583	5	0	1	94	99
									25	107	0	39	93	0	39	93	0	2	70	81
									12	42	0	9	45	0	9	45	0	3	NaN	0

SVM:

The aim of the SVM approach is to determine the best decision boundary that differentiates between various classes. The Radial Basis Function (RBF) kernel is used in this model. This method is a non-linear kernel that allows the model to learn non-linear decision boundaries. Also, Bayesian optimization modifies hyper-parameters of the model via the penalty parameter and kernel coefficient. In this manner, the optimal values for these parameters are selected to reduce the classification error on the validation data. Also, the data are normalized (given a zero mean and unit variance) before the SVM model is trained to lessen the effect of feature scale on model performance. A binary classifier is trained for each pair of classes using the "one-vs-one" method. In reality, the coding method is also regarded as a hyperparameter for issues involving more than two classes. Consequently, the aforementioned information are employed to train the algorithm on the training data. In 100 iterations, the SVM model obtained an average of 23% risk and 6% cost, as demonstrated by the Table 2 available. The risk associated with this method is 13% in the best-case scenario and 34% in the worst-case scenario. Additionally, the accuracy of this method is 99%, 92%, and 85% for the three classes, respectively. This procedure outperforms other methods in terms of accuracy and risk, as indicated by these findings. As shown by the average confusion matrix of this method, 44 of the 54 cases classified in class 3 were correctly identified, while 7 and 3 cases were classified in lesser classes, respectively. This demonstrates that the likelihood of this method misdiagnosing class 3 samples as classes 1 and 2 is relatively low. In addition, this method has a low risk of misdiagnosing class 2 samples as class 1, as only four of the 132 cases classified in class 2 were classified in class 1. The SVM method's overall risk and accuracy are 23% and 97% respectively, which are acceptable.

KNN:

The KNN model is being developed using training data. This function places training samples in the feature space. In order to forecast the class of the new sample, K samples in the feature space that are close to it are referenced. The distance (such as the Euclidean distance) is computed between the new sample and the training samples. Presumably, the prevailing class among the K neighbors is the expected label for the new sample. For this model, a range of options for k were assessed; cross-validation was used to identify the ideal value. For this model, various values for k were tested, and the best value was obtained using cross-validation. It is obvious from the Fig. 11 that the KNN method achieves the lowest level of risk when the number of neighbors (k) is equal to 1. So, the optimal number of neighbors in this method was determined to be k = 1.

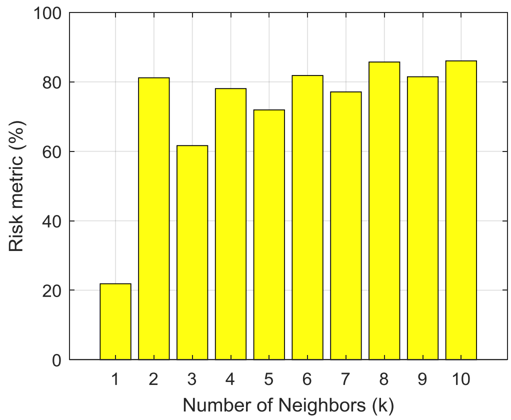

Fig. 11: Risk assessment of the KNN algorithm with different numbers of neighbors.

That is, the class of a new sample is predicted solely by utilizing its adjacent neighbor in the feature space. According to Table 2, this model has attained an average accuracy of 96%, a risk of 24%, and a cost of 7% over 100 iterations. This method carries a risk of 13 % in the best-case scenario and 34% in the worst-case scenario. In addition, the accuracy of this method for the three classes is 99%, 88%, and 83%, respectively. Based on these findings, this procedure outperforms other methods in terms of accuracy and risk. Based on the average confusion matrix of this method, 43 of the 54 cases in class 3 were correctly identified, while 8 and 3 cases were classified into lesser classes, respectively. This matter implies that the likelihood of this method misclassifying class 3 samples as classes 1 and 2 is relatively low. In addition, the method's minimal risk of misclassifying clas s 2 samples as class 1 is supported by the fact that only 5 of the 132 cases in class 2 were classified as class 1. The KNN method with k = 1 has an overall risk of 24%, which is deemed satisfactory.

Random Forest:

This study employs the random forest method to classify PT health. In the training phase, the algorithm generates many decision trees and determines a class based on the average of these classifications or the mean prediction (regression) of the individual trees. The classifier is trained with a random forest model including 100 decision trees. Each decision tree is trained using random features and a random subset of the training data. Bagging is the technique employed to do this. This technique enhances tree diversity and mitigates overfitting. hence, it strengthens the model and increases its applicability across diverse scenarios. According to Fig. 1, a 100-fold randomized cross-validation method is used to evaluate the random forest model performance. The results, shown in the Table 2, show that the model was accurate 97% of the time, with a risk index of 24% and an unnecessary cost index of 2%. The model showed that the worst-case risk index is 33%, and the best-case risk index is 14%. The model is 99% accurate for the "Healthy" class, 94% accurate for the " needs retesting in the future" class, and 94% accurate for the "Needs Immediate Retesting" class. Due to its poor performance compared to SVM and KNN, this method placed third on the Risk Index. The average confusion matrix shows that out of 54 cases in Class 3, 43 were correctly identified as being in that class, 8 were mistakenly put in Class 2, and 3 were mistakenly put in Class 1. This means that the system did a pretty good job of finding major faults (Class 3). Also, out of 132 cases in Class 2, 124 were correctly classified. Six were wrongly labeled as Class 1, and two were wrongly labeled as Class 3. This indicates a minimal likelihood of erroneously classifying transformers that require retesting in the future. The Random Forest model achieves an adequate equilibrium among accuracy, risk, and cost. Nonetheless, its comparatively elevated risk score compared to SVM and KNN indicates that it has the potential to more effectively identify defective transformers.

Other classifiers:

In this part, we look at how well the classifiers used in this study worked. These are Naive Bayes, Decision Tree, RUSBoost, Gaussian Naive Bayes, LPBoost, Multinomial Logistic Regression, Discriminant Analysis, AdaBoostM2, and Multiple Linear Regression. By examining the Table 2, we can analyze the performance of these classifiers in relation to the risk index, the unnecessary cost index, and their corresponding confusion matrices. For instance, the risk index of the Naive Bayes classifier is 38%, whereas the unnecessary cost index is 35%. On the other hand, the decision tree classifier displays an unnecessary cost index of 9% and a danger index of 42%. Furthermore, a review of the confusion matrices reveals that certain classifiers do better in accurately categorizing transformers as belonging to class 3. The SVM, KNN and Random Forest classifiers notably identify a greater number of Class 3 instances than the Naive Bayes classifier. The classifier selection relies on the application and the relative relevance of accuracy, risk, and cost. For example:

- A classifier with a lower risk index should be preferred to reduce the danger of misclassifying damaged transformers as healthy.
- Choose a classifier with a lower unnecessary cost index to reduce costs from misclassifying healthy transformers as faulty.

The following points can be made based on the data of Table 2.

- Risk:
  - Risk classifiers between 25% and 50%:
    - Naive Bayes (38%)
    - Gaussian Naive Bayes (38%)
    - Decision Tree (42%)
    - RUSBoost (43%)
  - Risk classifiers between 50% and 75%:
    - LPBoost (59%)
  - Classifiers with a risk greater than 75%:
    - Multinomial Logistic Regression (83%)
    - Discriminant Analysis (86%)
    - Multiple Linear Regression (93%)
    - AdaBoostM2 (119%)
- Cost:
  - Classifiers with less than 10% cost:
    - Random Forest (2%)
    - Decision Tree (9%)
    - Multinomial Logistic Regression (9%)
    - Multiple Linear Regression (10%)
  - Classifiers with more than 10% cost:
    - Discriminant Analysis (14%)
    - Gaussian Naive Bayes (35%)
    - Naive Bayes (35%)
    - LPBoost (47%)
    - RUSBoost (47%)
- Confusion matrix (for class 3):
  - Naive Bayes:
    - Out of 54 samples in class 3, 37 are correctly identified, and 14 are classified in class 2 and 3 in class 1.
  - Decision Tree:
    - Out of 54 samples in class 3, 35 items are correctly identified, 13 items are classified in class 2, and 6 items are classified in class 1.
  - RUSBoost:
    - Out of 54 samples in class 3, 36 are correctly identified, and 12 are classified in class 2 and 6 in class 1.
  - LPBoost:
    - Out of 54 samples in class 3, 26 are correctly identified, and 23 are classified in class 2 and 5 in class 1.
  - Multinomial Logistic Regression:
    - Out of 54 samples in class 3, 15 cases are correctly identified, 34 cases are classified in class 2, and 5 cases are in class 1.
  - Discriminant Analysis:
    - Out of 54 samples in class 3, 18 items are correctly identified, 28 items are classified in class 2, and 8 items are classified in class 1.
  - AdaBoostM2:
    - Out of 54 samples in class 3, 0 items are correctly identified, 42 items are classified in class 2, and 12 items are classified in class 1.
  - Multiple Linear Regression:
    - Out of 54 samples in class 3, 8 cases are correctly identified, 43 cases are classified in class 2, and 3 cases are in class 1.

Ensemble learning based classifiers:

This section shows the performance of the suggested ensemble learning model, which uses the top eight classifiers. This model combines classifier predictions to improve detection. This approach uses SVM, KNN with k=1, Random Forest, Naive Bayes, Decision Tree, RUSBoost, Gaussian Naive Bayes, and LPBoost. The selection of these models was based on their exceptional performance and diversity in categorizing DGA data.

In the first subfig. 1 of Fig. 12, the risk indices of the 12 used classifiers are shown. This figure shows a significant break in the risk chart after the eighth classifier, indicating an increase in risk values to high and unacceptable levels. In other words, the first eight classifiers have significantly lower risk compared to the next four classifiers. For this reason, the top 8 classifiers have been selected as the best classifiers and have been used in the ensemble learning algorithm. So, it is possible to create an accurate and reliable detection model by combining the predictions of these low-risk classifiers. Moreover, eliminating four high-risk classifiers mitigates the excessive escalation of model complexity and preserves its speed and efficiency. Table 2 shows that the model reduced risk by 7% relative to the best single classifier (23%) over 100 iterations, achieving a risk level of 16%. The proposed method exhibits risk index in range of 7% in the best-case scenario to 26% in the worst-case scenario. The proportion of unnecessary expenses is observed to range from 2% to 13%, yielding an average value of 6%. In addition to the fact that the ensemble method offers less risk compared to the results of the individual method, it achieves an average accuracy of 97%. The accuracy attained in class 3 of the ensemble technique exceeds that of the individual classifier methods. This issue stems from prioritizing method selection intended to reduce the risk index. In the typical confusion matrix for this approach, 54 occurrences in class 3 were assessed, leading to 46 instances being correctly identified. In contrast, 6 cases were wrongly put in class 2, and 2 cases were wrongly put in class 1. This proves that the suggested method works very well at finding samples that belong to class 3. Only 3 of the 132 class 2 cases were misclassified; the other 121 were correctly classified. This demonstrates that using this method, there is a very low risk of incorrectly identifying class 2 samples. The ensemble learning method has the lowest risk, at 16% (14% for class 3 and 2% for class 2), of all the methods this study looked at. In addition, the unnecessary cost index that is related to this method works better than other methods, except for Random Forest and Decision Tree. Because this method has a low-risk rating and doesn't cost too much, it can be assumed that it is a trustworthy way to check the health of transformers. Fig. 12 shows how well the suggested ensemble learning model works compared to different algorithms. Looking at other methods, the ensemble learning approach has a relatively lower risk level. While reducing the chance of high-risk mistakes, this study shows that the suggested method is both efficient and effective at accurately figuring out the health status of PT.

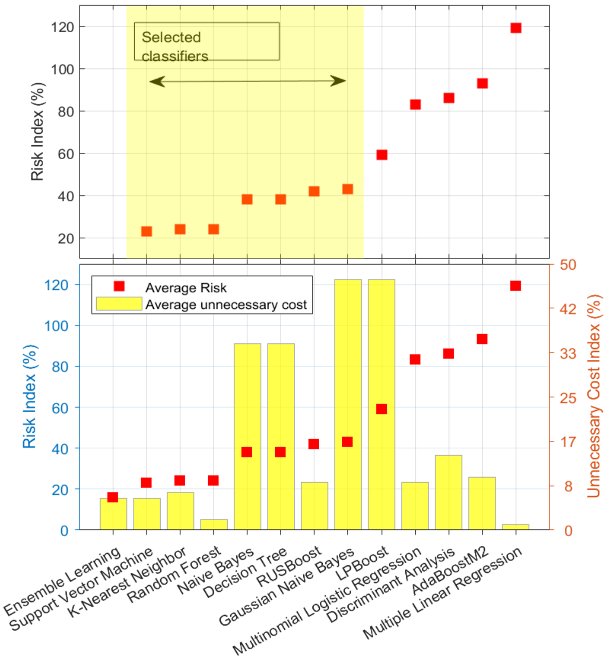

Fig. 12: Performance evaluation of the ensemble learning model compared to other methods.

Furthermore, the unnecessary cost index associated with this method is maintained at a level that is deemed acceptable when compared to the majority of alternative methods. This outcome demonstrates its efficacy in minimizing excessive expenditures. Overall, the ensemble learning model demonstrates commendable performance regarding the risk index and unnecessary cost index, indicating its efficacy in diagnosing the health status of the transformer.

Conclusion

A new ensemble learning technique for PT health assessment (dividing them into three groups: "healthy," "needs retesting in the future," and "needs immediate retesting.") using DGA data has been proposed in this paper. The algorithm combines the performance of eight classifiers (such as: SVM, KNN, Random Forest, Naive Bayes, Decision Tree, RUSBoost, Gaussian Naive Bayes, and LPBoost) and considers errors and costs to create a transformer health diagnosis model with high accuracy and reliability. The proposed approach is evaluated by the risk index, unnecessary cost index, and accuracy, which demonstrate superior performance compared to previous techniques across these metrics. The risk index has decreased as a result. This finding suggests that it can identify defective transformers more accurately. This model can minimize transformer maintenance expenses and extends their lifespan. This study found overall accuracy 97%, average risk index 16%, unnecessary cost index 6%, best-case risk 7%, and worst-case risk 26%. The findings show that the suggested approach efficiently evaluates transformer health, reducing the expenses and dangers of incorrect diagnosis. The optimal result of single classifier (SVM), is an average risk index of 23. The comparison of these values with the average risk index of 16% in the proposed ensemble learning technique demonstrates its efficiency and effectiveness in reducing misdiagnosis. In addition, the method improved the accuracy in the worst-case class. This improvement is particularly significant due to the importance of this class. It achieved an accuracy of 86%, compared to the SVM and KNN techniques (with recorded accuracy of 81% and 79%, respectively). Hence, the results show that the ensemble learning method works better at finding healthy transformers than single-classifier methods. This approach can be further improved through the incorporation of additional data and the application of sophisticated machine learning techniques. This methodology is capable of identifying additional transformer and electrical network anomalies.

Future Work

This study provides a number of directions for further investigation. The authors suggest that the following should be accounted for in future work:

- In this paper, a simple classification method with three classes was used. In future works, hierarchical classification methods can be used for a more precise categorization of transformer health status. For example, the class "Need for Retesting" can be divided into two subclasses: "Need for Retesting in the Near Future" and "Need for Retesting in the Distant Future".
- The DGA data used in this paper includes a part of the total technical information regarding transformers. In the future works, data can be collected with more detailed technical information, such as the type of transformer, the age of the transformer, and environmental conditions. This work can help improve the accuracy of machine learning models.
- In the proposed method, 11 DGA parameters were used to train machine learning models. Accordingly, the impact of feature selection on parameters can be examined for the future works. For example, methods such as PCA or mutual information-based approaches can be used to select the most important features.
- The proposed method is established based on traditional machine learning algorithms such as SVM and KNN. The deep learning algorithms, such as convolutional neural networks (CNN), can be used for fault detection in transformers in future works. These algorithms can improve fault detection accuracy (due to their high capability in learning complex features from data).
- The proposed method in this paper is established based on the certainty and accuracy of the initial data and the correct evaluation by the expert. In future works, the impact of uncertainty in DGA data on the performance of machine learning models can be examined. Additionally, methods based on uncertainty learning can be used to improve the accuracy of models.

Author Contributions

K. Gorgani Firouzjah and J. Ghasemi implemented the methods and evaluated their performance. Also, they wrote the paper, coordinated the study and contributed to the analysis of the results. They edited the manuscript. In addition, authors revised and discussed the results and approved the final manuscript.

Acknowledgment

This work is completely self-supporting, thereby no any financial agency's role is available.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

AdaBoostM2	Adaptive Boosting
AIC	Akaike Information Criterion
CH_4	Methane
C_2H_2	Acetylene
C_2H_4	Ethylene
C_2H_6	Ethane
CNN	Convolutional Neural Networks
CO_2	Carbon Dioxide
CO	Carbon Monoxide
DGA	Dissolved Gas Analysis
H_2	Hydrogen
KGM	Key Gas Method
KNN	K-Nearest Neighbors
LPBoost	Linear Programming Boosting
NN	Neural Network
O_2	Oxygen
PT	Power Transformer
RBF	Radial Basis Function
RUSBoost	Under-Sampling Boosting
SVM	Support Vector Machines
TCG	Total Combustible Gases

References

[1] Y. Wang, S. Gong, S. Grzybowski, "Reliability evaluation method for oil-paper insulation in power transformers," Energies, 4(9): 1362-1375, 2011.

[2] D. G. T. Da Silva, H. J. B. Da Silva, F. P. Marafão, H. K. M. Paredes, F. A. S. Gonçalves, "Enhanced health index for power transformers diagnosis," Eng. Fail. Anal., 126: 105427, 2021.

[3] C. Xia, M. Ren, B. Wang, M. Dong, G. Xu, J. Xie, C. Zhang, "Infrared thermography-based diagnostics on power equipment: State-of-the-art," High Volt., 6(3): 387-407, 2021.

[4] F. M. Laburú, T. W. Cabral, F. V. Gomes, E. R. de Lima, J. C. S. S. Filho, L. G. P. Meloni, " New insights into gas-in-oil-based fault diagnosis of power transformers," Energies , 17(12): 2889, 2024.

[5] S. Yu, D. Zhao, W. Chen, H. Hou, "Oil-immersed power transformer internal fault diagnosis research based on probabilistic neural network," Procedia Comput. Sci., 83: 1327-1331, 2016.

[6] L. Cheng, T. Yu, "Dissolved gas analysis principle-based intelligent approaches to fault diagnosis and decision making for large oil-immersed power transformers: A survey," Energies, 11(4): 913, 2018.

[7] Z. A. A. Etman, D. E. A. Mansour, N. H. El-Amary, "Performance evaluation of dissolved gas analysis techniques against measurement errors," in Proc. 2017 IEEE 19th International Conference on Dielectric Liquids, 2017.

[8] V. Tra, B. P. Duong, J. M. Kim, "Improving diagnostic performance of a power transformer using an adaptive over-sampling method for imbalanced data," IEEE Trans. Dielectr. Electr. Insul., 26(4): 1325-1333, 2019.

[9] H. C. Sun, Y. C. Huang, C. M. Huang, "A review of dissolved gas analysis in power transformers," Energy Procedia, 14: 1220-1225, 2012.

[10] M. S. Ali, A. H. Abu Bakar, A. Omar, A. S. Abdul Jaafar, S. H. Mohamed, "Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review," Electr. Power Syst. Res., 16: 109064, 2023.

[11] M. Duval, "The duval triangle for load tap changers, non-mineral oils and low temperature faults in transformers," IEEE Electr. Insul. Mag., 24(6): 22-29, 2008.

[12] N. Poonnoy, C. Suwanasri, T. Suwanasri, "Fuzzy logic approach to dissolved gas analysis for power transformer failure index and fault identification," Energies, 4(1): 36, 2021.

[13] O. E. Gouda, S. H. El-Hoshy, S. S. M. Ghoneim, "Enhancing the diagnostic accuracy of DGA techniques based on IEC-TC10 and related databases," IEEE Access, 9: 118031-118041, 2021.

[14] A. M. Aciu, C. I. Nicola, M. Nicola, M. C. Nițu, "Complementary analysis for dga based on duval methods and furan compounds using artificial neural networks," Energies, 14(3): 588, 2021.

[15] Hartono, Y. Muharni, C. Adipura, W. Martiningsih, M. Otong, M. Irvan, "Analysis of Power Transformator Conditions Using Dga Method Using Artificial Neural Network in Krakatau Electrical Power Company," Int. J. Eng. Technol. Manag. Res., 7(6): 77-88, 2020.

[16] D. A. Barkas, S. D. Kaminaris, K. K. Kalkanis, G. C. Ioannidis, C. S. Psomopoulos, "Condition assessment of power transformers through DGA measurements evaluation using adaptive algorithms and deep learning," Energies, 16(1): 54, 2023.

[17] S. J. T. Shahrabad, V. Ghods, M. T. Askari, "Power transformer fault diagnosis using DGA and artificial intelligence," Recent Adv. Comput. Sci. Commun., 3(4): 579-587, 2019.

[18] J. Li, Q. Zhang, K. Wang, J. Wang, T. Zhou, Y. Zhang, "Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine," IEEE Trans. Dielectr. Electr. Insul., 23(2): 1198-1206, 2016.

[19] S. Genc, S. Karagol, "Fuzzy logic application in DGA methods to classify fault type in power transformer," in Proc. IEEE 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA 2020), 2020.

[20] A. Abu-Siada, S. Hmood, S. Islam, "A new fuzzy logic approach for consistent interpretation of dissolved gas-in-oil analysis," IEEE Trans. Dielectr. Electr. Insul., 20(6): 2343-2349, 2013.

[21] Y. Zhang, Y. Tang, Y. Liu, Z. Liang, "Fault diagnosis of transformer using artificial intelligence: A review," Front. Energy Res., 10: 1006474, 2022.

[22] H. Zheng, Y. Zhang, J. Liu, H. Wei, J. Zhao, R. Liao, "A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers," Electr. Power Syst. Res., 155: 196-205, 2018.

[23] S. Souahlia, K. Bacha, A. Chaari, "SVM-based decision for power transformers fault diagnosis using Rogers and Doernenburg ratios DGA," in Proc. IEEE 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13), 2013.

[24] O. Kherif, Y. Benmahamed, M. Teguar, A. Boubakeur, S. S. M. Ghoneim, "Accuracy improvement of power transformer faults diagnostic using KNN classifier with decision tree principle," IEEE Access, 9: 81693-81701, 2021.

[25] Z. B. Sahri, R. B. Yusof, "Support vector machine-based fault diagnosis of power transformer using k nearest-neighbor imputed DGA dataset," J. Comput. Commun., 2(9): 22-31, 2014.

[26] Y. Benmahamed, O. Kherif, M. Teguar, A. Boubakeur, S. S. M. Ghoneim, "Accuracy improvement of transformer faults diagnostic based on DGA data using SVM-BA classifier," Energies, 14(10): 2970, 2021.

[27] U. M. Rao, I. Fofana, K. N. V. P. S. Rajesh, P. Picher, "Identification and application of machine learning algorithms for transformer dissolved gas analysis," IEEE Trans. Dielectr. Electr. Insul., 28(5): 1828-1835, 2021.

[28] N. Haque, A. Jamshed, K. Chatterjee, S. Chatterjee, "Accurate sensing of power transformer faults from dissolved gas data using random forest classifier aided by data clustering method," IEEE Sens. J., 22(6): 5902-5910, 2022.

[29] Y. Benmahamed, M. Teguar, A. Boubakeur, "Diagnosis of power transformer oil using PSO-SVM and KNN classifiers," in Proc. IEEE 3rd International Conference on Electrical Sciences and Technologies in Maghreb, CISTEM 2018, 2018.

[30] H. Yousefpour, J. Ghasemi, "Ensemble-based detection and classification of liver diseases caused by hepatitis c," Contrib. Sci. Technol. Eng., 1(1): 33-43, 2024.

Biographies

Khalil Gorgani Firouzjah was born in Babolsar, Mazandaran, Iran. He received his B.S. degree in Electrical Engineering from Ferdowsi University of Mashhad in 2005, and his M.S. and Ph.D. degrees in Electrical Engineering from the University of Mazandaran in 2008 and 2013, respectively. He served as a Researcher at Mazandaran Regional Electric Company from 2009 to 2011. He then joined the University of Mazandaran in 2014 as an Assistant Professor in the Electrical Engineering Department, and was promoted to Associate Professor in the same department in 2018. He continues to serve as an Associate Professor at the University of Mazandaran. His research focuses on power systems, energy management, and the application of machine learning and optimization algorithms within the power systems.

- Email: k.gorgani@umz.ac.ir
- ORCID: 0000-0002-7132-6759
- Web of Science Researcher ID: AAN-9973-2020
- Scopus Author ID: 55364934100
- Homepage: https://rms.umz.ac.ir/~kgorgani/

Jamal Ghasemi received his Bachelor's, Master's, and Ph.D. degrees in Electrical Engineering from the University of Mazandaran in 2004, 2008, and 2012, respectively. Since 2012, he has been a faculty member at the University of Mazandaran, where he currently serves as an Associate Professor. His research interests include fuzzy theory, evidence theory, image processing, and pattern recognition.

- Email: j.ghasemi@umz.ac.ir
- ORCID: 0000-0002-9573-0107
- Web of Science Researcher ID: AAN-1682-2020
- Scopus Author ID: 26032012000
- Homepage: https://rms.umz.ac.ir/~jghasemi/

How to cite this paper:
F K. Gorgani Firouzjah, J. Ghasemi, "Ensemble learning algorithm for power transformer health assessment using dissolved gas analysis," J. Electr. Comput. Eng. Innovations, 13(2): 387-402, 2025.

DOI: 10.22061/jecei.2025.11450.805

URL: https://jecei.sru.ac.ir/article_2279.html