



Research Article

Fault prediction of transformers using machine learning techniques

Nagendra SINGH^{1,*}, Ashutosh KASHIV², Harsh Pratap SINGH³

¹Department of Electrical Engineering, Trinity College of Engineering and Technology, Karimnagar, 505001, India

²Department of Electrical Engineering, Prestige Institute of Engineering, Management & Research, Indore, 452010, India

³Department of Computer Science Engineering, Medicaps University, Indore, 453331, India

ARTICLE INFO

Article history

Received: 30 May 2025

Revised: 11 July 2025

Accepted: 27 July 2025

Keywords:

Fault Diagnosis; Machine Learning; Power Systems; Predictive Maintenance; Transformer Fault Prediction

ABSTRACT

Transformers are a crucial part of electrical power transmission and distribution systems. Fault prediction with high accuracy reduces downtime, lowers maintenance costs, and increases the lifespan of transformers. This study presents a novel multi-featured machine learning framework for transformer fault prediction, evaluated over dual-scale datasets. Machine learning techniques Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and K Nearest Neighbour (KNN) are used to detect transformer faults using two different datasets that include historical information such as temperature, vibration, oil level, humidity, load current, oil viscosity, and power transformer age. The performance of machine learning methods is compared in terms of accuracy, precision, F1 score and recall using two test case studies. The first case considered a simple dataset of 50 thousand, and the second case study involved complex data of more than one lakh transformers. The Decision Tree and Support Vector Machine models attained the highest classification accuracy across both datasets, with SVM showing near-perfect accuracy but failing to give meaningful precision, recall, and F1-score values—indicating potential overfitting or class imbalance in the present experimental evaluation. The Random Forest algorithm has obtained the best results in recall and F1 score, which indicates that it is better to recognize the fault pattern in complex operating conditions. While logistic regression has a moderate level of recall, it suffers from low precision and predictive power. The K-Nearest Neighbour method shows moderate performance in all the metrics and performance is quite consistent as the size of the dataset's changes. The decision tree method produced a trade-off between interpretability and predictive ability that was a compromise.

These findings proved the feasibility of machine learning for transformer fault diagnostics, with notable implications for real-time health monitoring in advanced power systems.

Cite this article as: Singh N, Kashiv A, Singh HP. Fault prediction of transformers using machine learning techniques. Sigma J Eng Nat Sci 2026;44(4):2783–2796.

*Corresponding author.

*E-mail address: ernsingh007@gmail.com

This paper was recommended for publication in revised form by
Editor-in-Chief Ahmet Selim Dalkilic



INTRODUCTION

A transformer is an important component of power systems as it is used to change the voltage levels during the transmission and distribution of electrical power. They help reduce the losses of the energy system and maintain its stability [1]. Power transformers are an important part of the power system to maintain the required voltage levels at the generation, transmission and distribution points to ensure smooth operation. Early fault diagnosis of power transformers is important to avoid equipment failure, reduce downtime and ensure continuity of supply [2]. Transformers are expensive and critical, and if they malfunction, they can have serious consequences such as economical repairs and extended power failures [3]. Therefore, pre fault diagnosis with accuracy has become an important area of research in recent years. Various types of transformer faults are illustrated in Figure 1.

The stability and efficiency of power systems heavily depend on the uninterrupted operation of transformers. However, transformers are susceptible to a number of operational stresses, environmental changes, aging and unforeseen faults. Such failures can cause service interruptions, financial loss and damage to associated equipment. Routine inspection or diagnostic testing is the traditional method for fault detection. However, these methods are usually not accurate enough, especially for early and reliable prediction of faults [4].

To overcome this, automation of diagnostic methods is being considered. The use of machine learning (ML) is an effective approach to analyze complex transformer data to predict faults with increased accuracy. The historical data, like temperature, vibration, oil properties, humidity, load conditions, etc., can be used to learn meaningful patterns from it, which can then be utilized to predict potential fault conditions by using ML models [5]. With the proper use of planned maintenance strategies, these prediction algorithms can minimize the number of downtime events, reduce the operational cost of the power system, and increase the overall reliability of the system. Of particular note, ML methods can be used to continually monitor and diagnose using both historical and real-time data [6].

In this paper, a detailed implementation of machine learning methods such as Decision Tree, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) is presented to forecast the faults in the power transformers. The analysis is performed over different operational conditions such as vibration, oil level, ambient temperature, humidity, load current, oil viscosity, and age of the transformer [7]. One of the novel features of this study is the fact that the transformer age was included as a predictive variable that has not been widely explored in previous studies. A graphical abstract of the proposed methodology is shown in Figure 2.

The structure of this paper is as follows: After Introduction in section 1, section 2 provides a literature review of existing fault diagnosis methods. Section 3 details

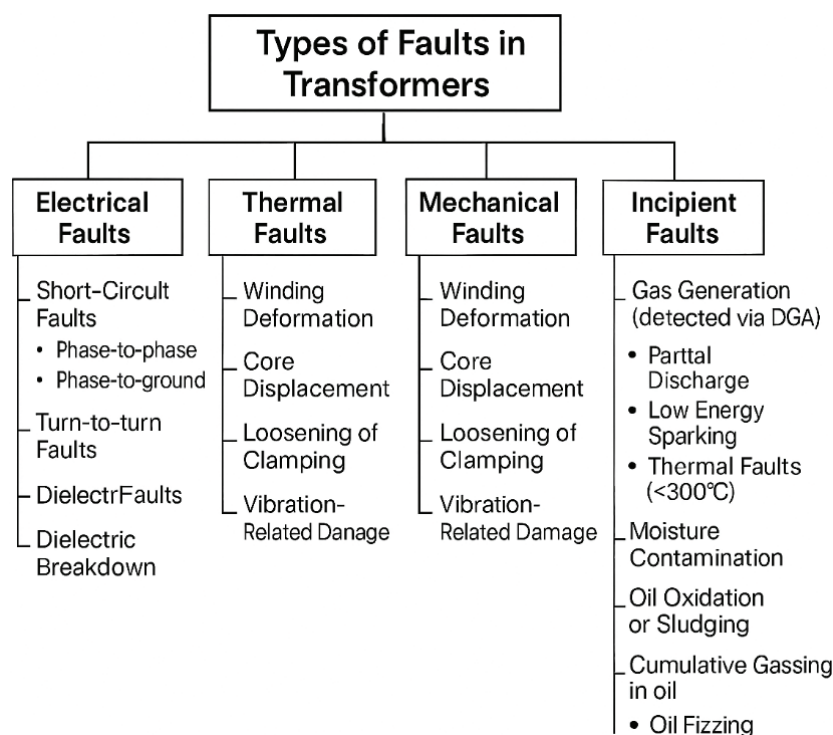


Figure 1. Types of transformer faults.

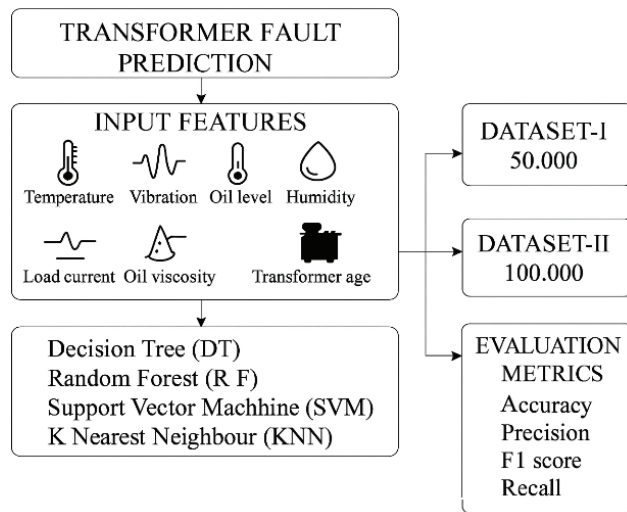


Figure 2. Graphical abstract of the present work.

the machine learning techniques applied. In section 4 and 5 problem formulation and methodology adopted are mentioned respectively. Section 6 presents the results obtained for the proposed work. In Section 7 conclusion of the present work with key findings are given.

LITERATURE REVIEW

This section presents a comparative review analysis of the different machine learning techniques used by the researchers in their study for the prediction of faults in transformer (Table 1) [1-21].

The majority of the studies presented in this paper are using dissolved gas analysis (DGA) or gas as the input for ML, but the increased use of ML in the diagnostics of power transformers shows up. Although gas data can be indicative of fault conditions, it does not capture the full operational profile of a power transformer. Additionally, many techniques particularly those using deep learning or hybrid models are often computationally expensive and lack real-time applicability due to high training and processing overhead [12]. A critical limitation identified across the literature is the narrow scope of input features. Most of the studies do not take into account factors such as ambient temperature, humidity, oil viscosity, vibration, load current, age of the transformer, etc., which are also important in the real operation and environment. The factors can serve as context and increase the robustness of ML models when included in forecasting models. The present study addresses these gaps by incorporating a wider and more comprehensive set of condition monitoring features into a single ML

Table 1. Literature Survey of application of Machine learning models used to detect fault in Transformers

Ref No.	Technique	Advantages	Limitations
[1]	Sweep Frequency Response Analysis	Effective fault prediction	Limited to supervised datasets
[2]	ANN and Genetic Algorithm	Good accuracy	Uses only gas data
[3]	Data Mining	Improved accuracy	Complex modeling
[4]	Dissolved Gas Analysis	Reliable performance	Relies solely on gas samples
[5]	Genetic Neural Computing	Low failure rate	High computation time
[6]	Random Forest	Better accuracy than traditional and AI-based methods	Based on DGA
[7]	Support Vector Machine	Outperforms ANN in some cases	Gas data dependency
[8]	ML Technique	High scalability and accuracy	Complex modelling process
[9]	ML and IoT	Enables real-time monitoring	High infrastructure cost
[10]	Standard ML and AutoML	High accuracy with unbalanced data	Long computation time
[11]	Fuzzy Logic and Neural Networks	High accuracy	Dependent on data quality; slow training
[13]	Multivariable, Multi-step ML	High prediction performance	Requires large memory
[14]	Fuzzy Logic and Neural Networks	Good results	Complex and data-dependent
[15]	Optical Spectroscopy and Data Mining	High accuracy	Influenced by oil sample quality
[16]	Hidden Markov Model	Dynamic prediction using DGA	Needs extensive historical data
[17]	LSTM Network	High prediction accuracy	Large data requirements
[18]	Bi-level ML	Improved prediction accuracy	Complex and oil-dependent
[19]	K-Nearest Neighbors	Strong performance on unseen data	Model selection is challenging
[20]	IoT and Ensemble ML	Enables real-time analysis	High memory cost and data security concerns
[21]	RNN , Bi-GRU and LSTM	Early fault detection	Requires large datasets and high computation power

framework. Compared to previous studies, it consists of both operational and environmental variables, which for a more comprehensive modelling of transformer behavior under different conditions. Moreover, this study examines the performance of the models on two different types of case studies: a mid-scale data set containing 50,000 samples and a larger-scale data set with more than 100,000 samples, to investigate ML scalability and generalization capabilities. In this research, the four supervised learning models, decision tree, random forest, support vector machine (SVM), and KNN are also systematically compared to each other. In this integrative and comparative study, it extends the previous methodologies, and a more feasible, scalable solution for transformer fault diagnosis is presented.

ML Techniques

In this section, the ML models used in this work are described, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR) and K-Nearest Neighbors (KNN). In this section, the theory behind the ML models used in this work are described: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR) and K-Nearest Neighbors (KNN).

Decision Tree (DT)

Due to simplicity and the ability to handle complex historical data, the transformer decision tree gives highly accurate predictions of faults. This technique predicts the faults as per the input data and classifies the types of faults [22]. According to historical data of transformers, this method can predict which parameters will cause a fault to arise. The tree consists of (i) Root node: It represents the entire dataset of the transformer and chooses the best splitting feature. (ii) Internal Nodes: This node represents decision points and divides the data based on feature thresholds. (iii) Leaf Nodes: These nodes provide the final output prediction, which can be a fault state or the normal operation of the transformer.

For the fault prediction of transformer used the following steps. (i) Data Collection: It collects the transformer parameters such as oil temperature, load current, humidity etc. (ii) Feature Selection & Preprocessing: Features are chosen based on their relevance to transformer faults. Missing values and outliers are handled to improve model performance. (iii) Tree Construction: The dataset is split using a splitting criterion until a stopping criterion is met (e.g., maximum depth, minimum samples per leaf) (iv) Fault Classification: Once trained, the Decision Tree classifies new data into faulty or not faulty condition of transformer (v) Model Evaluation: In terms of accuracy, precision, recall, and F1-score, the performance of the model is assessed. A confusion matrix is used to analyze misclassifications.

Random Forest (RF)

A RF is a powerful ensemble learning method that improves transformer fault prediction by combining multiple Decision Trees [23]. It reduces over fitting, improves accuracy,

and handles noisy or imbalanced data effectively. A Random Forest consists of multiple Decision Trees, where using random subset of transformer data each tree of RF techniques is trained. Each random subset of data is dividing in a tree in such a way to reduce the correlation. The final prediction is obtained through classification or averaging. Following steps flowed by RF technique to identify the fault in transformer. (i) First historical data of transformer is collected. (ii) Feature Engineering & Preprocessing: Handle missing values, normalize data, and select relevant features. (iii) Model Training: Create multiple decision trees using different feature subsets and data samples. (iv) State of Fault Classification: The model predicts faulty or Normal Condition. (v) Model Evaluation: Proposed model is evaluated in terms of accuracy, precision, recall, and F1-score are measured. A confusion matrix helps identify misclassification errors.

Support Vector Machine (SVM)

SVM is a powerful supervised machine learning techniques can be used for transformer fault prediction. It works well with small datasets, can handle non-linearly separable data, and is robust against over fitting. SVM separates data points using a hyper plane in the number of features [24]. SVM technique has the following Key Concepts. Hyperplane – A decision boundary that separates fault and non-fault conditions. (ii) Support Vectors – Data factors closest to the hyperplane, defining the margin. (iii) Kernel Trick– it adjusts the transformer data that is not linearly separable to a higher dimension so that it can be separated. When data is linearly separable, it uses the linear kernel. If data is complex, the polynomial kernel is used, and if nonlinear data is available, the radial basis function kernel is used. Following steps followed by SVM technique for Transformer Fault Prediction.

Logistic Regression (LR)

Logistic regression is a popular statistical and machine learning technique for classification when the outcome of interest is binary. It can forecast the conditions and tell us whether a transformer is likely to fail or not, according to whether it offers the input features of temperature fluctuation and others. It provides a probability score, helping operators make risk-based decisions [25]. Unlike complex deep learning models, logistic regression is computationally inexpensive and can work in real-time fault monitoring systems. The Following Key Features have the LR techniques: The predictive capability of logistic regression depends on selecting relevant features of the transformer. which may include: Thermal and electrical parameters, Operating Conditions such as voltage, load variations, environmental conditions and transformer age. The LR technique considers the following steps for model training and fault prediction in transformers. (i) Collection of historical data of the transformer. (ii) Collected data analysis and split into training and testing sets. (iii) For obtaining the optimum solution of the collected data LR model is trained using maximum

likelihood estimation (MLE). (iv) After find the optimum solution of the LR model we can assess the performance in terms of accuracy, precision, recall, F1-score, and ROC-AUC.

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a non-parametric, instance-based learning algorithm that operates by comparing a new data point with existing instances in the training dataset [26]. In this study, KNN is employed to classify power transformers as either healthy or faulty using sensor-based diagnostic data. The input features include oil temperature, load current, voltage, vibration, moisture content, ambient temperature, and several time-domain statistical features such as mean, standard deviation, RMS, skewness, and kurtosis—derived from current and vibration signals. Since KNN is a distance-based algorithm, all continuous features were normalized to ensure that variables with larger scales did not dominate the distance calculations. The optimal value of k (number of neighbors) was selected by cross-validation in a trade-off between model complexity and generalization capability. The class label of a new transformer instance is predicted through the majority vote of its k nearest neighbors in the feature space. This approach enables effective fault recognition based on patterns of similar past operating conditions.

Problem Formulation

For the reliable operation and proper maintenance, it is required to identify the early fault detection in the transformer. Traditional diagnostic methods are often time-consuming, labor-intensive, and may not capture complex patterns in multivariate sensor data. This work proposes a data-driven machine learning approach for predicting Power Transformer faults using historical sensor measurements.

Objectives

The primary objective of this study is to develop and evaluate machine learning models capable of accurately classifying Transformer health conditions as faulty or non-faulty based on sensor data. The models are trained to learn from both simple (less number of variables) and complex datasets and are assessed using classification metrics such as accuracy, precision, recall, F-1 and confusion matrices.

Problem definition

Formally, let $X \in \mathbb{R}^{n \times m}$ represents the feature matrix, where n is the number of samples and m is the number of features. The target variable $y \in \{0, 1\}^n$ represents the class labels for each instance, where 0 denotes a healthy power transformer and 1 denotes a faulty condition [27]. Supervised binary classification used to define the problem statement.

$$\text{Given: } X = \{x_1, x_2, \dots, x_n\}, \quad (1)$$

$$y = \{y_1, y_2, \dots, y_n\} \text{ with } y_i \in \{0, 1\} \quad (2)$$

$$\text{Find: } f: \mathbb{R}^m \rightarrow \{0, 1\} \text{ such that } f(x_i) \approx y_i \quad (3)$$

The goal is to train a predictive model f that can classify a Transformer instance $x' \in \mathbb{R}^m$ as either faulty or non-faulty with high accuracy and reliability.

MATERIALS AND METHODS

Here, a full machine learning pipeline is applied for classification of Power Transformer health status using the sensor data. The methodology includes preprocessing of data, dimensionality reduction, model training using cross-validation, performance evaluation and comparative analysis of several classification algorithms. The workflow is designed to ensure the reliability, generalizability, and robustness of the models developed. Figure 3 presents the ML Pipeline. This visual illustration outlines the entire workflow from data acquisition and preprocessing to model evaluation and selection for better clarity.

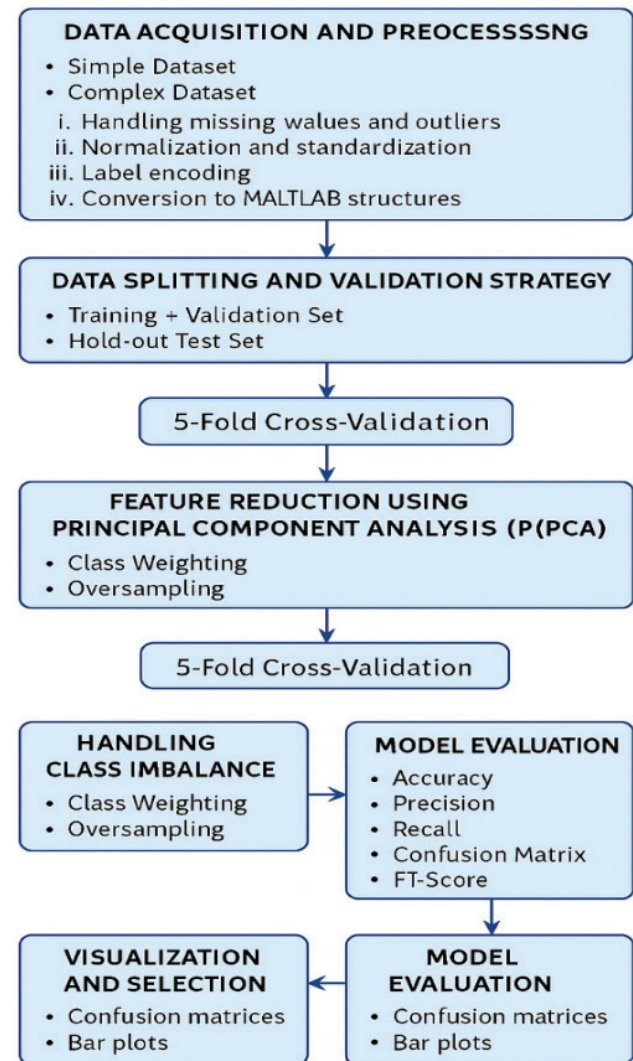


Figure 3. The machine learning implementation pipeline for the proposed work.

Table 2. Sensors and equipment used for data

Parameter	Sensor / Equipment Type
1. Vibration	Accelerometer (e.g., piezoelectric), Vibration sensor modules
2. Oil Level	Magnetic float level sensor- Ultrasonic level sensor- Capacitive level sensor
3. Ambient Temperature	Digital temperature sensor (e.g., LM35, DS18B20)- Thermistor- RTD (Resistance Temperature Detector)
4. Humidity	Digital humidity sensor & Capacitive humidity sensors
5. Load Current	Hall-effect current sensor(CT)
6. Oil Viscosity	Viscosity sensor (Inline viscosity transmitters)
7. Transformer Age	Calculated from commissioning date

Data Acquisition and Preprocessing

The sensors and equipment used for the measurement of respective parameters are mentioned in Table 2.

The data used for this study is a collection of different substations in Madhya Pradesh, India. However, we would like to clarify that due to the difficulty in accessing real-world data that is consistent and complete, some part of the dataset used in our work is synthetic or manipulated to simulate realistic operating conditions and cover a wider spectrum of fault scenarios.

Data Splitting and Validation Strategy

To ensure fair and unbiased evaluation of the ML algorithms testing dataset divided into three parts: a) training, b) validation, and c) testing. The full data set was first shuffled to randomize the distribution of the sample and to prevent any ordering bias. The data is split up as follows: i. 60% of the data is used to train the models and learn the internal patterns and representations from the input features. ii. 20% used during model development for evaluation of intermediate performance and to help tune hyper parameters or compare model variants. No complex tuning was done in this study, but this set allowed a clean separation between fitting the model and final evaluation. iii. In the final model only 20% test set kept for evaluation to publish unbiased accuracy, precision, and recall results. The split was chosen to strike a balance between sufficient training and generalizability. The validation set is important because it allows monitoring the model and possibly to optimize it without leaking any information from the test set, which thus preserves the integrity of the final performance assessment.

Cross-Validation and Model Evaluation Strategy

To ensure robust and generalizable model performance, this work adopted a two-stage evaluation process comprising 5-fold cross-validation followed by a final test set evaluation. At the starting, the total dataset was split into two part a) Training and Validation Set keeping 80% of the data, and b) Hold-out Test Set where keeping 20% of the data. The training and validation set was then used for 5-fold cross-validation. In this process, the data was partitioned into five equal-sized folds. In each of the iterations four folds were used for training and the remaining one fold was used for validation. This was

repeated five times ensuring that each fold was a validation set once. The performance indicators accuracy, precision and recall were averaged for each fold to obtain an overall score for each machine learning algorithm. The machine learning algorithms considered in this paper are: Decision Tree Random Forest, SVM, Logistic Regression, KNN. The best performing model was chosen based on the average cross validation accuracy. The model was then re-trained on the full training+validation set, and its performance was evaluated on the independent hold-out test set (20%) that was not used in any part of model training or validation. This evaluation pipeline is useful to reduce over fitting and also to ensure that the performance reported is an indication of the model's ability to generalize to unseen data.

Feature Reduction using Principal Component Analysis (PCA)

This work used Principal Component Analysis (PCA) to reduce the input features' dimensionality and improve the model efficiency. Before PCA, the features were standardized to have zero mean and unit variance. The PCA transform was performed with the minimum number of principal components required to explain at least 95% of the variance in the dataset. This dimensionality reduction helped to reduce multi-collinearity, reduce noise and improve model generalization while preserving most of the original information. The reduced feature set was then used as input for the machine learning algorithms. To prevent data leakage, PCA transform was applied only on the training data inside each cross-validation fold.

Handling Class Imbalance

In practical fault detection situations, data are often imbalanced, meaning that the number of non-faulty (majority) data points is much larger than that of faulty (minority) data points. To address this problem, our study used two main approaches to reduce the effect of class imbalance and improve the model reliability:

Class Weighting via Cost-Sensitive Learning

A custom cost matrix was used to introduce class-specific misclassification penalties into the training. This matrix penalizes the misclassification of minority class

instances (fault cases) more. This makes the model more sensitive to these critical examples. This was done by using the cost parameter of MATLAB in models such as Decision Tree and SVM: “The cost matrix was constructed inversely proportional to class frequencies, so that the minority class got proportionally more importance in training. This reduces false negatives, which are particularly problematic for fault detection systems.

Oversampling of Minority Class

Besides cost-sensitive learning, a simple random oversampling technique was employed to further balance the dataset. This was done by duplicating the occurrences of the minority class, so that both classes had the same number of samples. In the case of over-sampling, instances of the minority class were added to the training data, without losing any information. This method was particularly effective when used with models like K-Nearest Neighbors (KNN) and Logistic Regression, which are sensitive to class distributions.

Model Evaluation

Each trained model was evaluated using the following metrics: Accuracy: Overall correctness, Precision: Correctness of positive predictions. Recall (Sensitivity): Ability to detect actual faults, Confusion Matrix: To visualize classification performance. F1-score: Harmonic mean of precision and recall. Performance metrics were computed for each fold, and average values were used for comparative analysis.

Visualization and Selection

Confusion matrices were generated individually for each method. Bar plots were used to visualize accuracy and

precision across methods. The best performing model was selected based on the highest accuracy and a favorable balance of precision and recall.

RESULTS AND DISCUSSION

This section presents the comparative analysis of different machine learning models for fault prediction using sensor data. Two different datasets are used to check the effectiveness and efficiency of different models consider in this article.

Details and Comparative Analysis of Data Sets

The two data sets which differ significantly in complexity and parameters are used for the prediction of fault are shown in Table 3.

Hence the first data is comparatively simple, generating only three sensor readings and using fixed threshold-based fault conditions. It is suitable for basic anomaly detection. The second dataset introduces more parameters, external influences (humidity, load, aging), and real-world imperfections (label noise, class imbalance), making it more realistic but complex.

Results Obtained from First Data Set (Simpler Data)

Performance metrics of machine learning models from first data set (simpler, data size 50,000)

Different machine learning models are tested for the given data set. The models were evaluated based on accuracy, precision, and recall as shown in Table 4. Following Table 4 summarizes the classification performance of each model.

Table 3. Data set used for the proposed work

Feature	First data set (Simpler) Size -50000	Second data set (More Complex) Size- 100000
Number of Variables	3 (Temperature, Vibration, Oil Level)	7 (Ambient Temperature, Humidity, Load Current, Oil Viscosity, Power Transformer Age)
Class Imbalance	Evenly distributed faults	Faults occur in about 10% of the dataset (more realistic)
Label Noise	No noise added	5% of labels are flipped to introduce real world uncertainty
External Factors	None	Includes environmental factors like humidity, ambient temperature, and Power Transformer aging

Table 4. Performance analysis of each model for first data set (less- 50,000 and simpler data)

Method	Accuracy	Precision	Recall	F1 Score
Decision Tree	98.7867	2.2989	2.0202	2.1505
Random Forest	96.12	4.3478	23.2323	7.3248
SVM	98.9067	0	0	0
Logistic Regression	57.62	0.8791	56.5657	1.7313
KNN	97.8867	1.3393	3.0303	1.8576

From Table 4, Support Vector Machine (SVM) achieved the highest accuracy (98.91%), but recorded zero precision, recall, and F1 score. This shows the inability of the SVM to detect faults due to imbalance in the classes in the simpler dataset. This shows that the SVM only predicted the majority class which are the healthy transformers. The Decision Tree (DT) also obtained high Accuracy (98.79%) with moderate Precision (2.30%) and F1 score (2.15%) with better fault detection than SVM while being interpretable and computationally efficient. Although the Random Forest (RF) model had a slightly lower accuracy (96.12%), it outperformed the other models in terms of recall (23.23%)

and F1 score (7.32%), which indicates its strong ability to detect minority class data points (faults) with a slightly lower overall accuracy. Such trade-off is especially useful for safety-critical applications such as transformer monitoring. Logistic Regression (LR) had very low accuracy (57.62%) but surprisingly high recall (56.56%) indicating most of the faults were identified but at the expense of very poor precision (0.87%) making it untrustworthy for real world use. The K-Nearest Neighbor (KNN) was a balanced model, but with low performance, with accuracy (97.88%), recall (3.03%) and F1 score (1.86%). It performed equally across the data complexities, but not as well as the tree-based models.

Confusion matrix analysis for first data set (simpler data)

The performance of classification problems can be measured by a confusion matrix. This allows us to evaluate how well a machine learning model is performing by comparing actual labels with predicted labels. Figure 4 shows the confusion matrix for the Decision Tree model. The confusion matrix for the Decision Tree model indicates a high true negative detection but a very poor true positive identification (only 2 out of 99 faults) exhibiting a strong bias toward the majority class. The model accuracy is high but the recall is low which reflects the effect of class imbalance on the detection of faults.

The confusion matrix for the random forest model is illustrated in Figure 5. The confusion matrix for the Random Forest model shows improved fault detection, correctly identifying 23 true positives out of 99 fault cases. The model gets a better recall and F1 score, thus being more effective to identify faults in imbalanced datasets, although it generates more false positives (506). Similarly, the confusion matrix for the SVM model in Figure 6 indicates no

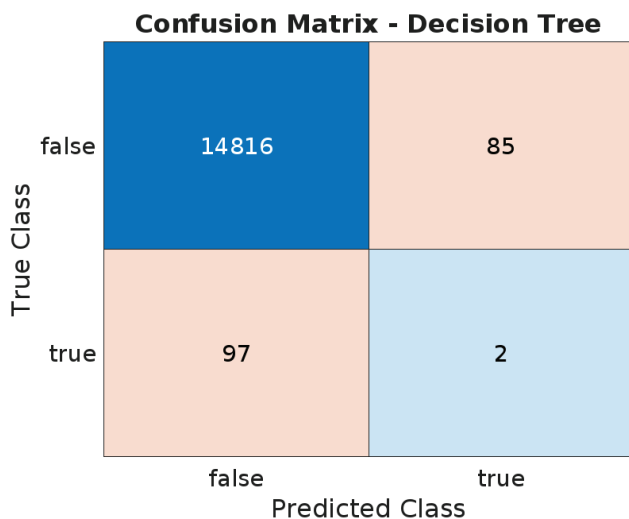


Figure 4. Confusion matrix by Decision Tree method for simpler data.

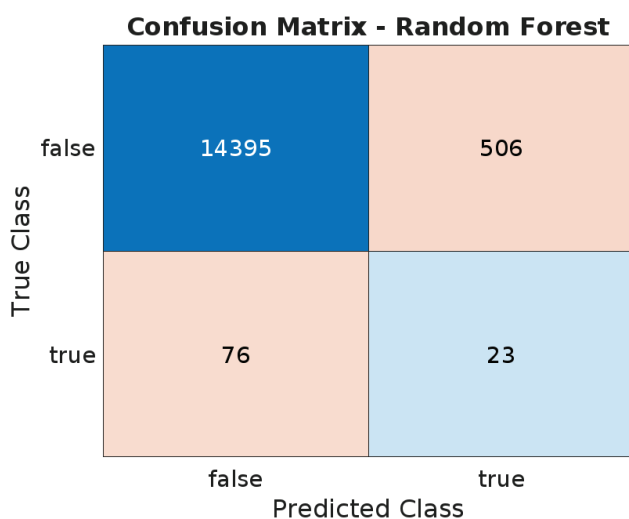


Figure 5. Confusion matrix by Random Forest method for simpler data.

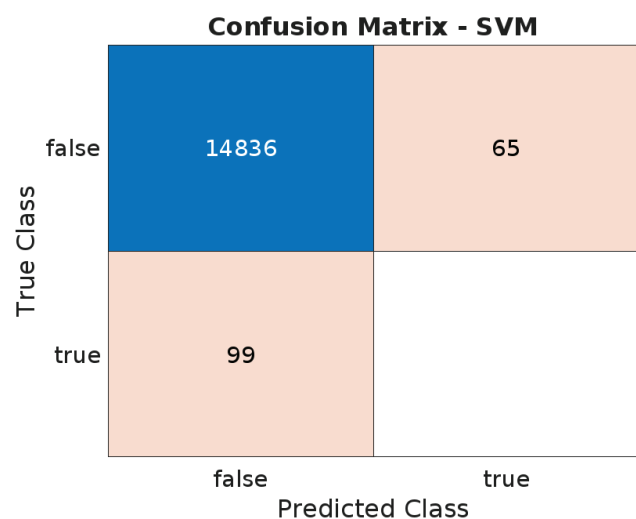


Figure 6. Confusion matrix by SVM method for simpler data.

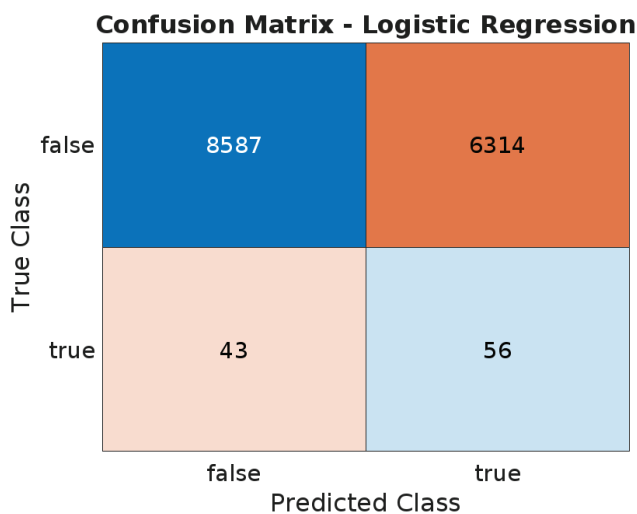


Figure 7. Confusion matrix by Logistic Regression method for simpler data.

true positives, i.e., it did not detect any faulty transformers. Despite high overall accuracy, the model is completely biased toward the majority class, resulting in zero recall and F1 score, and is thus ineffective for imbalanced fault prediction tasks.

Figure 7 shows the confusion matrix for the logistic regression model. The confusion matrix for Logistic Regression shows moderate fault detection, correctly identifying 56 out of 99 faulty cases (high recall), but with a large number of false positives (6,314). This results in low precision, making the model prone to false alarms despite its sensitivity to faults.

The confusion matrix for the KNN model, as illustrated in Figure 8 shows very limited fault detection, correctly identifying only 3 out of 99 faulty cases, with 96 false negatives. The overall accuracy is high, but the low recall and F1 score indicate that the model performs poorly in fault detection when the classes are imbalanced.

In all the confusion matrices most models like Decision Tree and SVM achieved high accuracy but poor fault detection because of the class imbalance. Random Forest achieved the best balance, detecting more faults (higher true positives) while maintaining reasonable false positive rates. for data set 1.

Graphical analysis for first data set (simpler data)

Figure 9 shows a bar graph comparison of accuracy, precision, recall and F-1 score across models.

In conclusion, for simpler data while Random Forest and Decision Tree deliver the best results in terms of absolute accuracy, Random Forest is preferable due to its enhanced robustness. For more interpretable models, Logistic Regression or SVM are also strong contenders, delivering near-perfect results with marginally lower accuracy.

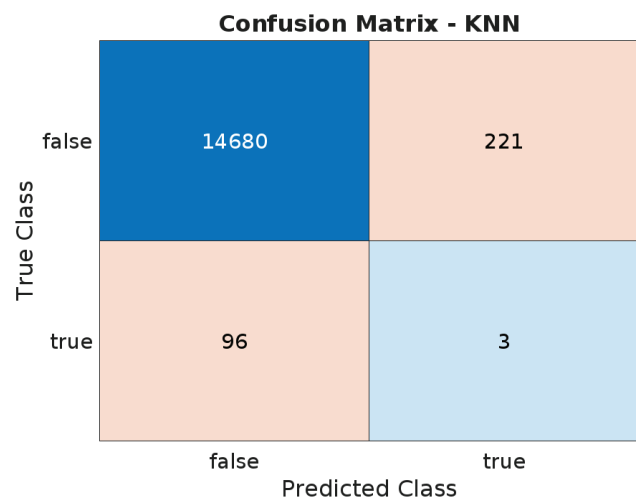


Figure 8. Confusion matrix by KNN method for simpler data.

Results Obtained from Second Data Set (Complex Data)

Performance metrics of machine learning models obtained for second data set (complex data)

Different machine learning models are tested for the given data set 2 which is comparatively more complex and having more number of input variables. The models were evaluated again for this complex data based on accuracy, precision, and recall. Table 5 summarizes the classification performance of each model.

From Table 5, The results show that while SVM achieved the highest accuracy (99.03%), but its precision, recall and F-1 score is poor. Random Forest demonstrated the best overall fault detection performance with the highest recall (23.30%) and F1 score (8.70%), making it the most reliable model for imbalanced data.

Confusion matrix analysis obtained for second data set (complex data)

Figures 10 to 14 present the confusion matrices of Decision Tree, Random Forest, SVM, Logical Regression and KNN model respectively.

Figure 10 shows the decision tree model confusion matrix shown correctly identified only 6 faulty cases while misclassifying 97, indicating low recall and limited effectiveness in detecting minority class instances. Random Forest confusion matrix present in Figure 11, showed improved fault detection with 24 true positives and relatively fewer false negatives, suggesting a better balance between sensitivity and accuracy.

Figure 12 presents SVM confusion matrix, failed to detect any faults, with zero true positives and high false negatives, despite achieving the highest accuracy, which reflects its poor handling of class imbalance.

Figure 13 present the confusion matrix of the Logistic Regression, achieved the highest recall among the models by identifying 52 faults, but at the cost of a very high number of false positives (5,912), thus compromising its overall precision.

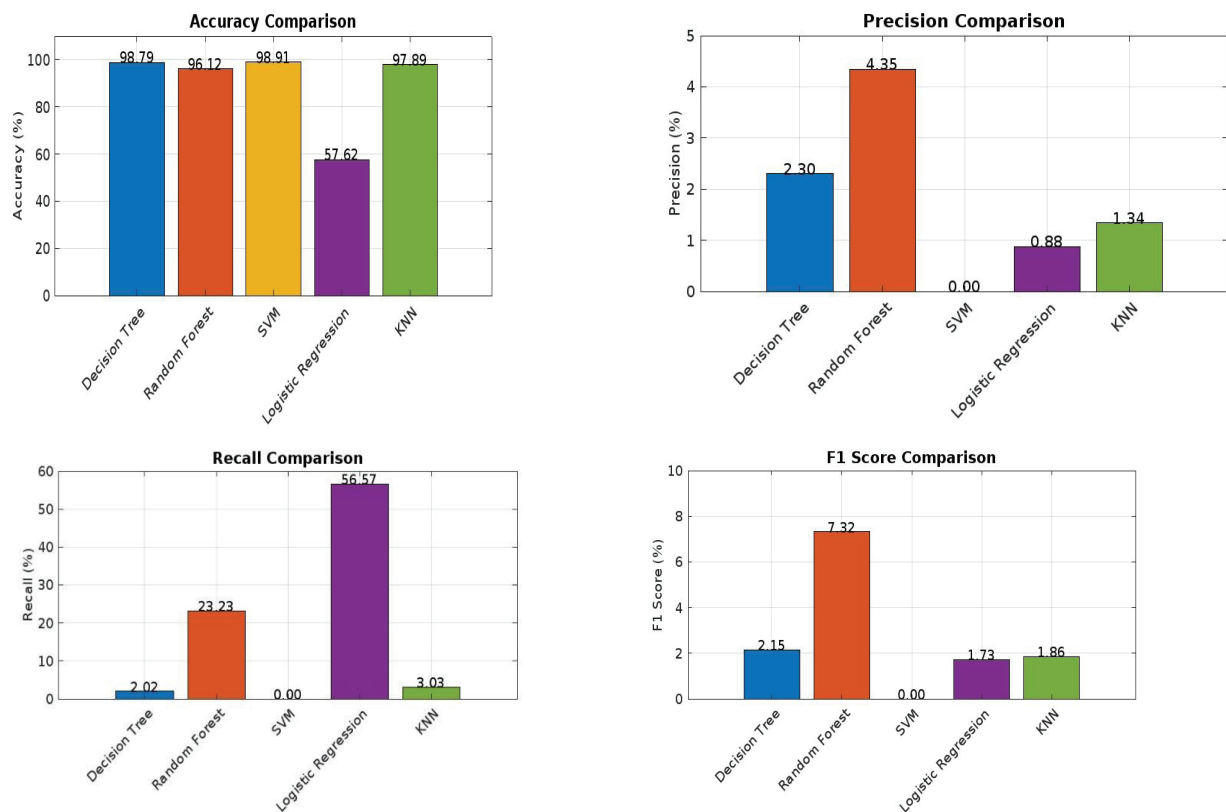


Figure 9. Accuracy, Precision, Recall F1 comparison graph obtained by different methods for first data set.

Table 5. Classification performance of each model for **second data set (high-100000, complex data)**

Method	Accuracy	Precision	Recall	F1 Score
Decision Tree	98.8467	7.3171	5.8252	6.4865
Random Forest	96.64	5.3452	23.301	8.6957
SVM	99.0333	0	0	0
Logistic Regression	60.2467	0.8719	50.4854	1.7142
KNN	97.9733	1.9139	3.8835	2.5641

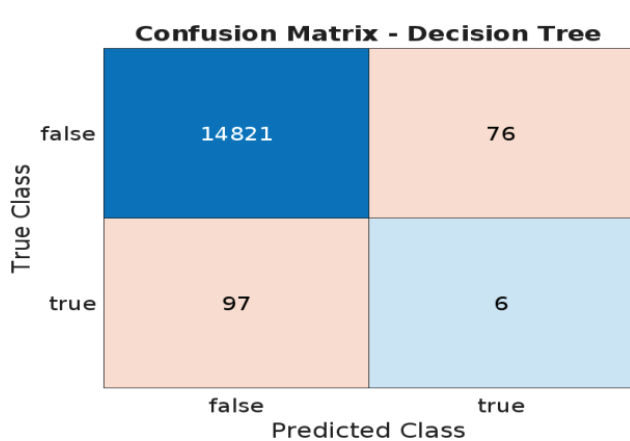


Figure 10. Confusion matrix by Decision Tree for second data set (complex data).

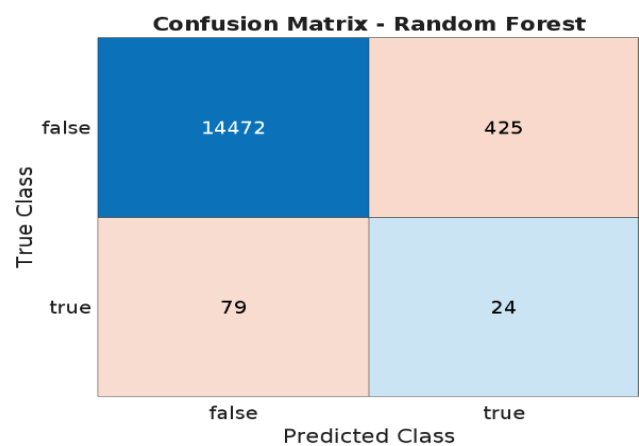


Figure 11. Confusion Matrix by Random Forest for second data set (complex data).

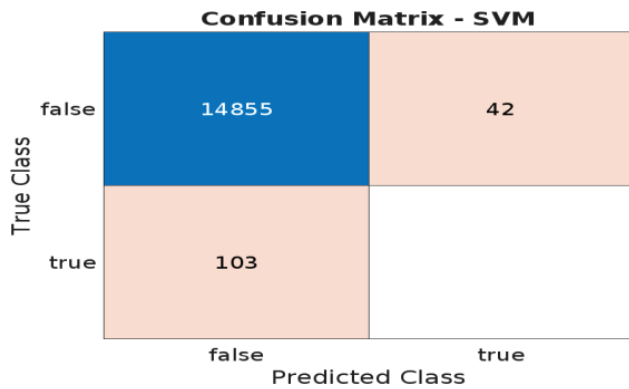


Figure 12. Confusion Matrix by SVM for second data set (complex data).

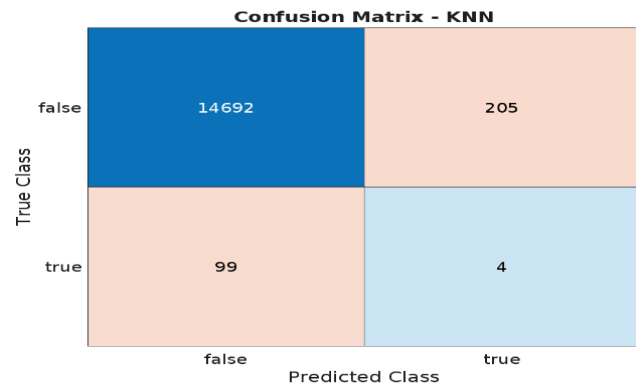


Figure 14. Confusion Matrix by KNN for second data set (complex data).

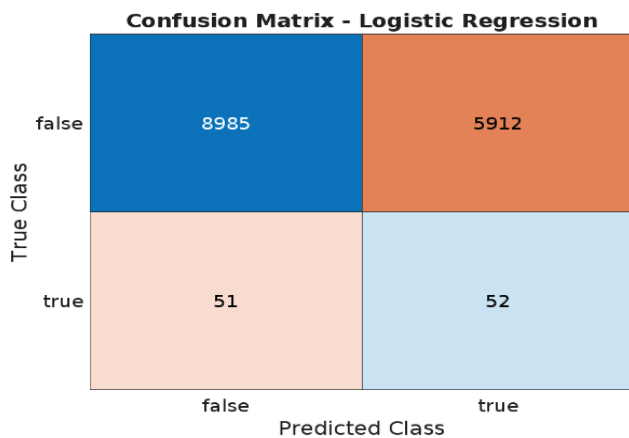


Figure 13. Confusion Matrix by Logical Regression for second data set (complex data).

Figure 14 presents the confusion matrix of KNN, detected only 4 true faults and missed 99, showing weak fault identification capability similar to the Decision Tree and SVM models.

Graphical analysis for second data set (complex data)

Figure 15 shows a bar graph comparison of accuracy and precision of ML models for complex data.

In conclusion, Random Forest provided the most balanced and effective fault detection performance among all the evaluated models.

Comparison of Obtained Result with Conventional Method

To evaluate transformer fault diagnosis methods, both conventional and machine learning-based approaches are considered. The Table 6 [24, 32] summarizes and compares these methods based on key performance indicators.

Table 6. Comaritive analysis of different methods

Method	Accuracy	Training Time	Interpretability	Real-Time Capability	Remarks
ML in present work – Random Forest, Decision Tree, SVM, LR, KNN	96 % (Random Forest) 99.03 % (SVM) Very high	Fast	Medium	Moderate (batch-base)	Effective on structured tabular data, easy to implement and training is fast.
Deep Learning Approaches [24, 25]	95–98% Very high	High	High	Moderate	High accuracy, but needs large annotated datasets and has low transparency
Convolution Neural Network [26]	77 %	Continuous	Low	Very High	fault features at different time scales
Dissolve Gas Analysis Method [28, 29]	High	Not Applicable	Medium	No spatial fault info, slow to react	Detecting chemical degradation and arcing
Partial Discharge method [30, 31]	High	Not Applicable	Fast	Moderate	Affected by electrical noise; suitable early fault detection
Thermal imaging [32]	Very low	Not Applicable	Low	Low , Only detects surface-level issues	Detecting hotspots and external component issues
FRA [27]	Medium	Not Applicable	Medium	No real time capability	Suitable for Structural diagnosis

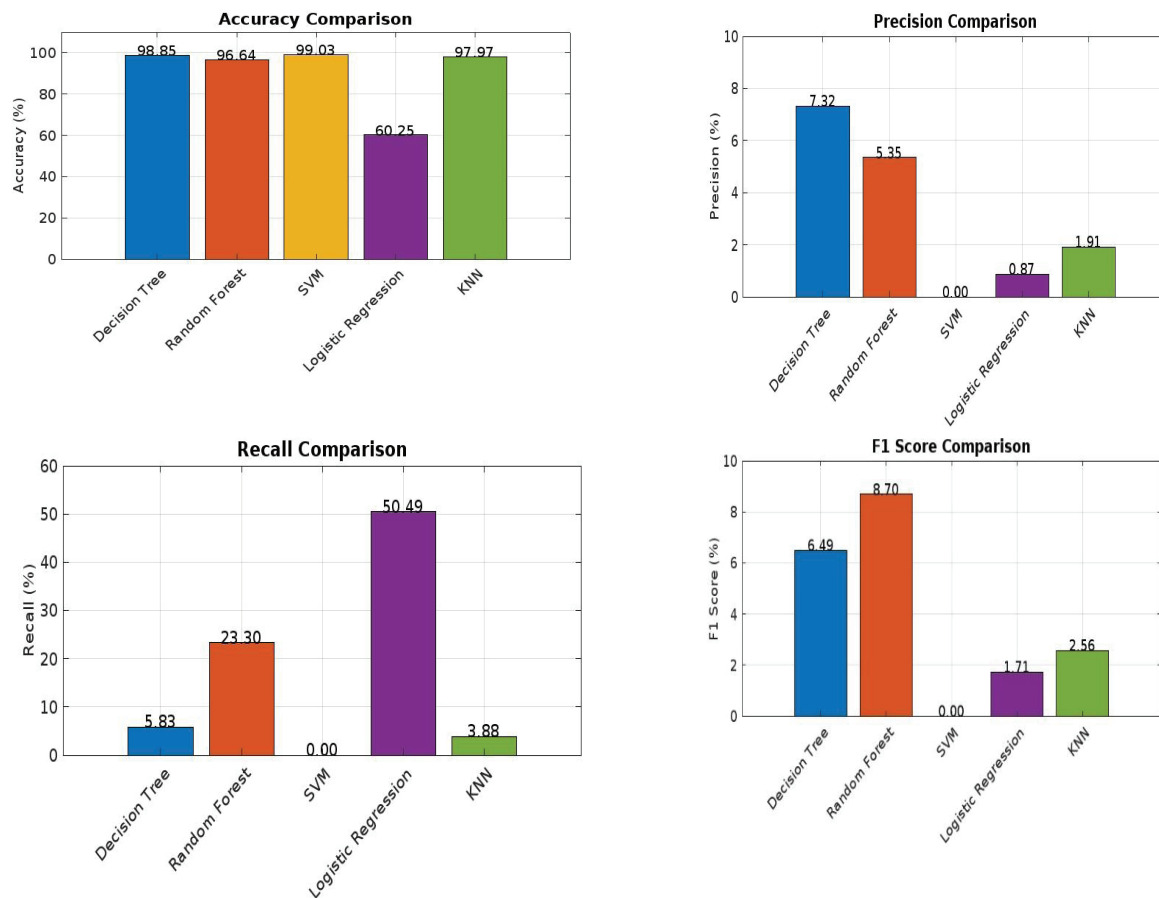


Figure 15. Accuracy, Precision, Recall, F1 comparison graph obtained by different methods for first data set.

Conventional methods such as DGA, partial discharge detection, and FRA offer established diagnostic reliability, while modern ML and deep learning techniques demonstrate higher accuracy as well as predictive capabilities.

CONCLUSION

The study demonstrated the potential of machine learning models in predicting transformer faults using multi-featured datasets. Decision Tree and SVM were the most accurate among all models for simple as well as complex datasets. However, the performance of SVM was affected by the zero precision, recall and F1-score which shows the effect of class imbalance. The most robust model was Random Forest, achieving the best compromise between recall and F1-score, particularly in complex operational settings. Logistic Regression and KNN performed moderate to poor. Logistic regression had higher recall but at the cost of very low precision. This studies overall concluded that the proposed tree model improves the interpretability and predictive strength. The obtained results confirm that the viability of machine learning for reliable transformer fault detection, enabling proactive maintenance and enhancing grid reliability.

ABBREVIATIONS

DGA	Dissolved gas analysis (DGA) techniques
GA	Genetic Algorithm
KNN	K-Nearest Neighbors (KNN)
ML	Machine Learning
SVM	Support Vector Machine (SVM)
RF	Random Forest

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

Not applicable.

CONFLICT OF INTEREST

Not applicable.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

- [1] Bjelić M, Brković B, Žarković M, Miljković T. Machine learning for power transformer SFRA based fault detection. *Int J Electr Power Energy Syst* 2024;156:109779. [\[CrossRef\]](#)
- [2] Velásquez RM, Lara JV. Root cause analysis improved with machine learning for failure analysis in power transformers. *Eng Fail Anal* 2020;115:104684. [\[CrossRef\]](#)
- [3] Ravi NN, Drus SM, Krishnan PS. Data mining techniques for power transformer failure prediction model: A systematic literature review. In: 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE); 2019. p. 305–309. IEEE. [\[CrossRef\]](#)
- [4] Demirci ME, Gözde H, Taplamacıoğlu MÜ. Fault diagnosis of power transformers with machine learning methods using traditional methods data. *Int J Tech Phys Probl Eng* 2021;13(49).
- [5] Al-Janabi S, Rawat S, Patel A, Al-Shourbaji I. Design and evaluation of a hybrid system for detection and prediction of faults in electrical power transformers. *Int J Electr Power Energy Syst* 2015;67:324–335. [\[CrossRef\]](#)
- [6] Al-Sakini SR, Bilal GA, Sadiq AT, Al-Maliki WA. Dissolved gas analysis for fault prediction in power transformers using machine learning techniques. *Appl Sci* 2024;15:118. [\[CrossRef\]](#)
- [7] Singh HP, Singh KN, Songare LS, Singh N, Arya N. Energy management systems using artificial intelligence: Future trends, challenges and solutions. In: *Power system management*. CRC Press; 2025.
- [8] Vita V, Fotis G, Chobanov V, Pavlatos C, Mladenov V. Predictive maintenance for distribution system operators in increasing power transformers' reliability. *Electronics* 2023;12:1356. [\[CrossRef\]](#)
- [9] Ramesh J, Shahriar S, Al-Ali AR, Osman A, Shaaban MF. Machine learning approach for smart distribution power transformers load monitoring and management system. *Energies* 2022;15:7981. [\[CrossRef\]](#)
- [10] Santamaria-Bonfil G, Arroyo-Figueroa G, Zuniga-Garcia MA, Azcarraga CG, Bassam A. Power transformer fault detection: A comparison of standard machine learning and AutoML approaches. *Energies* 2023;17:77. [\[CrossRef\]](#)
- [11] Mateus BC, Farinha JT, Mendes M. Fault detection and prediction for power transformers using fuzzy logic and neural networks. *Energies* 2024;17:296. [\[CrossRef\]](#)
- [12] Niranjane V. Machine learning for fault identification in power transformers. *Adv Nonlinear Var Inequal* 2025;28:540–558. [\[CrossRef\]](#)
- [13] Bai Y, Zhao J. A novel power transformer-based multi-variable multi-step prediction method for chemical process fault prognosis. *Process Saf Environ Prot* 2023;169:937–947. [\[CrossRef\]](#)
- [14] Mateus BC, Farinha JT, Mendes M. Fault detection and prediction for power transformers using fuzzy logic and neural networks. *Energies* 2024;17:296. [\[CrossRef\]](#)
- [15] Fauzi NA, Ali NN, Ker PJ, Thiviyathan VA, Leong YS, Sabry AH, et al. Fault prediction for power transformer using optical spectrum of power transformer oil and data mining analysis. *IEEE Access* 2020;8:136374–136381. [\[CrossRef\]](#)
- [16] Jiang J, Chen R, Chen M, Wang W, Zhang C. Dynamic fault prediction of power transformers based on hidden Markov model of dissolved gases analysis. *IEEE Trans Power Deliv* 2019;34:1393–1400. [\[CrossRef\]](#)
- [17] Ma X, Hu H, Shang Y. A new method for power transformer fault prediction based on multifeature enhancement and refined long short-term memory. *IEEE Trans Instrum Meas* 2021;70:3098383. [\[CrossRef\]](#)
- [18] Zhang D, Li C, Shahidehpour M, Wu Q, Zhou B, Zhang C, et al. A bi-level machine learning method for fault diagnosis of oil-immersed power transformers with feature explainability. *Int J Electr Power Energy Syst* 2022;134:107356. [\[CrossRef\]](#)
- [19] Taha IB, Mansour DA. Novel power transformer fault diagnosis using optimized machine learning methods. *Intell Autom Soft Comput* 2021;28:739–752. [\[CrossRef\]](#)
- [20] Zhang C, He Y, Du B, Yuan L, Li B, Jiang S. Power transformer fault diagnosis method using IoT based monitoring system and ensemble machine learning. *Future Gener Comput Syst* 2020;108:533–545. [\[CrossRef\]](#)
- [21] Zollanvari A, Kunanbayev K, Bitagsir SA, Bagheri M. Power transformer fault prognosis using deep recurrent neural network over vibration signals. *IEEE Trans Instrum Meas* 2020;70:3026497. [\[CrossRef\]](#)
- [22] Gil ES, Sivilla JEM, González AS, et al. Fuzzy logic applied to the diagnosis of technical conditions of distribution power transformers. *J Eng Appl Sci* 2023;70:129. [\[CrossRef\]](#)
- [23] Xie X, Huang F, He C, et al. A design of power prediction algorithm based on health assessment optimized neural network. *J Eng Appl Sci* 2024;71:96. [\[CrossRef\]](#)
- [24] Rucconi V, De Maria L, Garatti S, Bartalesi D, Valecillos B, Bittanti S. Deep learning for fault detection in transformers using vibration data. *IFAC-PapersOnLine* 2021;54:262–267. [\[CrossRef\]](#)
- [25] Singh N, Kumar Y, Pahariya Y, Valsove V, Singh HP. *Power system management: Advances and application*. CRC Press; 2025. p. 1–285. [\[CrossRef\]](#)

- [26] Li L, Wang S, Zhao Y, Wang F. Fault diagnosis for single-phase grounding in distribution network based on Hilbert-Huang transform and Siamese convolution neural network. In: 2023 35th Chinese Control and Decision Conference (CCDC); 2023. p. 3571–3576. [\[CrossRef\]](#)
- [27] Khalili Senobari R, Sadeh J, Borsi H. Frequency response analysis (FRA) of transformers as a tool for fault detection and location: A review. *Electr Power Syst Res* 2018;155:172–183. [\[CrossRef\]](#)
- [28] Singh N, Tamrakar S, Mewada A, Gupta SK. Artificial intelligence techniques in power systems operations and analysis. CRC (Auerbach) Publications; 2023. p. 1–72. [\[CrossRef\]](#)
- [29] Al-Sakini SR, Bilal GA, Sadiq AT, Al-Maliki WAK. Dissolved gas analysis for fault prediction in power transformers using machine learning techniques. *Appl Sci* 2025;15:118. [\[CrossRef\]](#)
- [30] Faizol Z, Zubir F, Saman NM, Ahmad MH, Rahim MKA, Ayop O, et al. Detection method of partial discharge on transformer and gas-insulated switchgear: A review. *Appl Sci* 2023;13:9605. [\[CrossRef\]](#)
- [31] Singh N, Chandanan AK, Wadhawan P, Volynets I, Mishra A, Gupta SK. Application of AI in fault detection and grid optimization. In: *Power system management*. CRC Press; 2025. [\[CrossRef\]](#)
- [32] Fanchiang KH, Huang YC, Kuo CC. Power electric transformer fault diagnosis based on infrared thermal images using Wasserstein generative adversarial networks and deep learning classifier. *Electronics* 2021;10:1161. [\[CrossRef\]](#)