Vol. 8, No. 2 | July – December 2025

Fault Prediction and Classification in Power Distribution Transformer Using Machine Learning

Zawar Ahmed Khan[1], Muhammad Amir Raza[1],*, Ramesh Kumar[2] , Shakir Ali Soomro[1], Hizbullah Odho[1]

Abstract:

This article used Random forest (RF) classifier, decision tree (DT) classifier, K-nearest neighbors (KNN) classifiers, extra tree (ET) classifier and extreme gradient boosting (XGB) classifier for fault prediction and classification in power distribution transformer to predict the phase line voltages, line currents, line-to-line voltages, neutral current, ambient temperature indicator, oil temperature indicator, oil level indicator, and various other alarm indicators like oil temperature indicator trip oil temperature indicator alarm and magnetic oil gauge alarm. The findings demonstrate that XGB and ET classifier, provide superior predictive capabilities compared to other methods.

Keywords: *Power transformer; fault detection; machine learning; sustainability.*

1. Introduction

Power transformers are crucial components in electrical systems, but they are susceptible to various types of faults that can lead to operational failures and significant economic losses [1]. Earth fault occurs when there is an unintended connection between the transformer windings and the ground. The magnitude of the fault current depends on the grounding method and the impedance of the windings. Solid grounding typically results in higher fault currents [2]. Internal faults happen due to insulation breakdown between turns in a winding, often caused by voltage surges or aging. They can create localized heating, leading to further insulation degradation and potential transformer failure [3]. Core losses cause the insulation breakdown within the core or debris accumulation that leads to excessive eddy currents. Core faults can cause overheating and significantly affect transformer performance. Regular oil and gas analysis can help in early detection. Tank fault involves oil leakage from the transformer tank, which can lead to reduced insulation and overheating. Such faults are often visible through oil level drops or external signs of leakage [4]. Open circuit faults occur when one phase of the transformer becomes disconnected, which can lead to overheating. These faults are generally less harmful as they can be manually addressed by disconnecting the transformer from the system. Phase to phase faults are rare but serious, these faults occur between phases and can result in high current flows similar to earth faults [5]. Protection schemes must account for these faults due to their potential for severe damage. External faults occur

[1] Department of Electrical Engineering, Mehran University of Engineering and Technology, SZAB Campus Khairpur Mir's, Sindh Pakistan
[2] Senior Manager (Power Distribution), K-Electric Karachi, Sindh Pakistan
*Corresponding Author: M. Amir Raza (amirraza@muetkhp.edu.pk)

outside the transformer but affect its operation, such as short circuits on connected lines. They can cause mechanical stress and thermal damage if not cleared promptly. Tap changers regulate voltage levels but can fail due to mechanical issues or electrical arcing. Regular maintenance is critical to prevent these faults [6]. The traditional and advanced fault prediction and protection methods for power transformer are described in Table 1.

Table 1: Traditional and advanced fault prediction and protection methods of power transformer

Traditional relay systems	
Buchholz Relay [7]	A gas-operated relay that detects internal faults by monitoring gas accumulation from insulation breakdown
Overcurrent Protection [8]	Engages during excessive current flow, preventing damage from overload conditions
Differential Protection [9]	Detects differences in current between primary and secondary windings, indicating potential faults
Advanced systems	
Fuzzy Logic Systems [10]	This approach allows for quicker identification of electrical faults and mechanical failures, significantly reducing maintenance costs and improving operational reliability
Internet of Things (IoT) Integration [11]	IoT technologies are being integrated into transformer monitoring systems to enhance fault diagnosis capabilities. The use of GPRS communication networks facilitates effective data transmission between monitoring terminals and data acquisition systems, making it easier to identify faults promptly.
Machine Learning Techniques [12]	Machine learning is also applied to transformer for fault detection, leveraging historical data to predict potential failures. These methods can analyze complex patterns in transformer behavior that traditional methods might miss, thereby enhancing predictive maintenance strategies
Dissolved Gas Analysis [13]	This technique detect the faults in oil-immersed transformers. This method analyzes the concentration of dissolved gases in transformer oil, correlating specific gas ratios with known fault types
Frequency Response Analysis [14]	This is another promising technique that compares the frequency response of a transformer under healthy conditions against its response during suspected faults.

2. Research Methodology

The dataset used in this study was obtained from Kaggle: https://www.kaggle.com/datasets/sreshta140/ai-transformer-monitoring. This data is collected via IoT devices from June 25th, 2019 to April 14th, 2020 which was updated every 15 minutes. This study undertakes the parameters like line-to-line voltages (VL12-voltage line 12, VL23-voltage line 23, VL31-voltage line 31), line currents (IL1-current line 1, IL2-current line 2, IL3-current line 3), phase line voltages (VL1-phase line 1, VL2-phase line 2, VL3-phase line 3), ambient temperature indicator (ATI), oil level indicator (OLI), neutral current (INUT), oil temperature indicator (OTI), and also some alarming indicators are proposed like magnetic oil gauge alarm (MOG_A), oil temperature indicator trip (OTI_T) and oil temperature indicator alarm (OTI_A). The complete data was updated and monitored after every 20 minutes and indicate fault when appear 1 and on no fault system shows 0. The proposed method for this study is shown in Figure 1. All experiments were conducted using Python 3.9 in Google Colab with libraries including Scikit-Learn (v1.1.3), imbalanced-learn (v0.11.0), XGBoost (v1.7.5), pandas (v1.5.3), and Seaborn (v0.12.2).

2.1 Data Preprocessing

- **Normalization**

The normalization of features was ensured that each is capable to train the model. However, the standard deviation is considered to be 1 and become 0 as mean. The representation for normalization is given in Eqn. (1) [15].

$$Xscaled = \frac{X - \mu}{\sigma} \quad (1)$$

In Eqn. (1) original feature value is defined by X, features mean by μ and standard deviation by σ.

- **Fault Classification/Labelling**

By combining OTI_A, OTI_T, and MOG_A, a new column was created named fault indicator. A function was built to classify faults based on the fault indicator to provide clearer labels for the dataset. The function uses the logic as shown in Table 2.

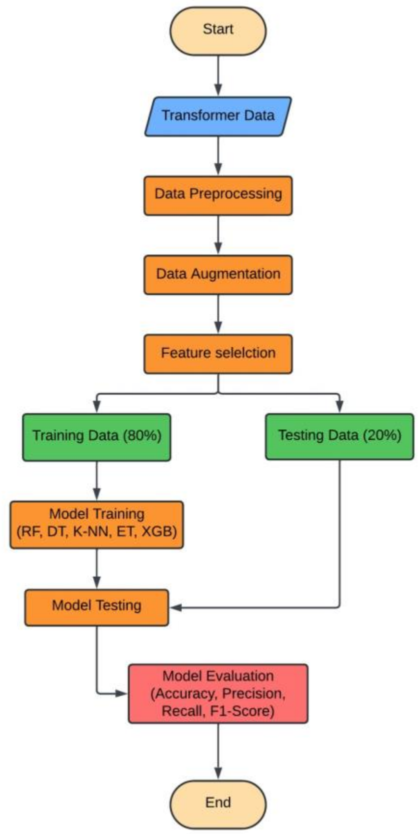


Fig 1. Research method flow diagram

Table 2: Shows faults classification based on codes

OTI_A	OTI_T	MOG_A	Fault Indicator	Fault Classification
0	0	0	000	NF
0	0	1	001	OL
1	0	0	100	OH
0	1	0	010	OH
0	1	1	011	OH+OL
1	0	1	101	OH+OL
1	1	1	111	OH+OL

The classify () function was run to each row in the data frame to generate a new column called 'fault classification'. This organized method to fault labeling and categorization ensures that future analysis and modelling are based on well-defined categories.

2.2 Data Balancing

The original dataset was highly imbalanced, with a significant majority of instances labeled as ‘No Fault’ as shown in Table 3. To address this imbalance, synthetic minority over-sampling technique (SMOTE) and random under sampling techniques were employed.

Table 3: Class distribution before sampling

Fault type	No fault (NF)	Overheating (OH)	Oil leakage (OL)	OH+OL
Distribution	12879	167	1846	207

- **Random under sampling**

Random under sampling involves reducing the number of majority class instances to match the minority class. This technique helps in balancing the dataset but can lead to loss of potentially valuable information. The process can be mathematically represented as given in Eqn. (2) [16].

$$Nsampled = min\,(Nminority, Nmajority) \quad (2)$$

Where ($N_{sampled}$) is the majority class samples under sampling, ($N_{majority}$) is the majority class samples, and ($N_{minority}$) is the minority class samples. In this study, the 'No Fault' instances were reduced from 12,879 to 4,958.

- Synthetic minority over-sampling technique

The SMOTE technique is used to address the dataset's class imbalance. SMOTE generates synthetic samples for the minority class by interpolating between existing ones. This method addresses the minority class, resulting in a more balanced dataset for model training. The SMOTE algorithm can be mathematically represented as given in Eqn. (3) [17].

$$Xnew = xi + \lambda\,(xj, xi) \quad (3)$$

The x_i and x_j are the minority class instances, and λ is considered between 0 and 1 randomly. This equation generates new synthetic samples by interpolating between existing minority class instances.

- **Tomek Links**

Tomek links are used to identify and remove overlapping instances between classes. A Tomek Link consist of a pair of instances (xa, xb), where xa is an instance from the minority class and xb is the nearest neighbor of xa from the majority class. The process can be mathematically represented as given in Eqn. (4) and Eqn. (5) [18].

$$d(xa, xb) < d(xa, xi)\ for\ all\ xi \neq xb \quad (4)$$

$$d(xa, xb) < d(xb, xj)\ for\ all\ xj \neq xa \quad (5)$$

Where d(x, y) denotes the distance between instances x and y. These conditions imply that x_a and x_b are each other’s nearest neighbors and belong to different classes. Removing Tomek linkages helps to clean the dataset by reducing borderline occurrences that are difficult to classify, hence enhancing the classifier's performance. After applying SMOTE Tomek, the fault distribution was balanced as shown in Table 4. This balanced dataset was then used to train and evaluate the machine learning models. The use of these data augmentation techniques ensures that the models are not biased towards the majority class and can accurately predict and classify faults in distribution transformers.

Table 4: Class distribution after sampling

Fault Type	NF	OH	OL	OH+OL
Distribution	4958	4625	4459	5000

2.3 Feature Selection

The feature selection procedure includes identifying and maintaining the most relevant features to efficiently classify transformer fault types. The feature set, was chosen based on its ability to capture significant variations across different fault conditions as shown in Eqn. (6) and Eqn. (7) [19].

$$X = [ATI, OTI, OLI, VL1, VL2, VL3, IL1, IL2, IL3, INUT, VL12, VL23, VL31] \quad (6)$$

$$Y = [NF, OH, OL, OH + OL] \quad (7)$$

To justify and refine the feature set, a comprehensive feature importance analysis was conducted using SHapley Additive exPlanations (SHAP) values and permutation importance. SHAP values provided clear interpretability by quantifying each feature's contribution to model predictions, highlighting which voltage and current indicators carried the most predictive value, and revealing any redundant or negligible features that could dilute the model's effectiveness. Permutation importance by measuring the drop in predictive performance when a feature's values are randomly shuffled corroborated these rankings, further validating the retained features. To address the risk of multicollinearity, especially with overlapping voltage and current measurements, variance inflation factors (VIFs) and correlation matrices were evaluated, ensuring that no predictors exhibited excessive linear dependence, which could destabilize coefficient estimates and lead to unreliable inferences. Given observed correlations, principal component analysis (PCA) was also performed to reduce dimensionality, combining correlated features into uncorrelated principal components that preserved high data variance while mitigating redundancy and improving model robustness. This multi-pronged approach ensured that the final feature set was not only highly informative and non-redundant but also less susceptible to instability from multicollinearity.

- Current measurement

These current-related features are crucial in detecting anomalies such as overloading and neutral imbalance. The line plot (Figure 2) illustrates that the current in each of the three phases (IL1, IL2, IL3) sharply increases during (OL) faults, declining noticeably during (OH) and (OH + OL). Even though the neutral current (INUT) has a similar tendency, suggesting that it may be useful for fault identification. This pattern indicates that phase and neutral currents are important for fault identification.

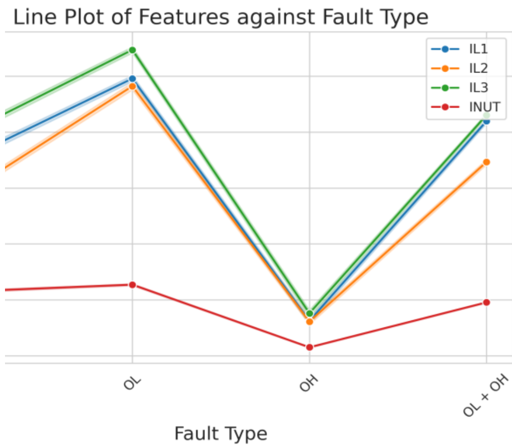


Fig 2. Line plot of 3 phase currents vs Fault Types

- Voltage measurement

Phase and line-to-line voltage imbalances that occur during failures are measured using voltage. Line-to-line voltages (VL12, VL23, and VL31) exhibit a dramatic spike during OL faults and a large reduction during OH and combined fault circumstances, as depicted in Figure 3. Although they still exhibit some minor fluctuations, phase voltages (VL1, VL2, and VL3) stay comparatively steady.

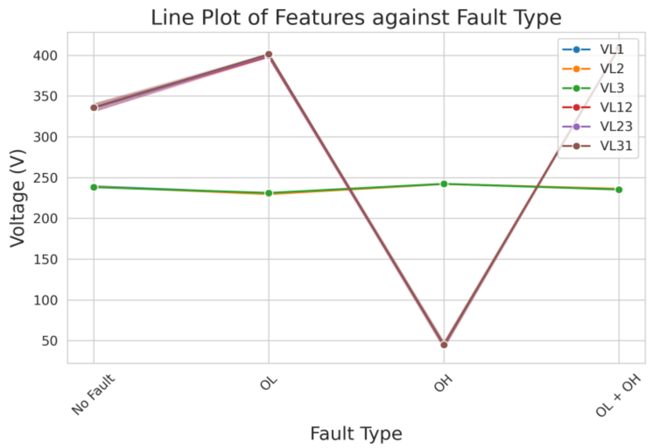


Fig 3. Line plot of 3 phase voltages vs Fault Types

- Temperature and OLI measurement

OTI increases continuously during OH and OH + OL. Then under the oil leakage condition, OLI drops at greater level with the purpose to detect oil leakage. However, negligible variance in fault types was noticed with the implication of reduced predictive value in ATI. To detect the transformer failure by using the feature selection, OLI and OTI are the most significant factors. These selected features interpret transformer behavior under various fault scenarios, contributing to fault detection and classification in the machine learning model.

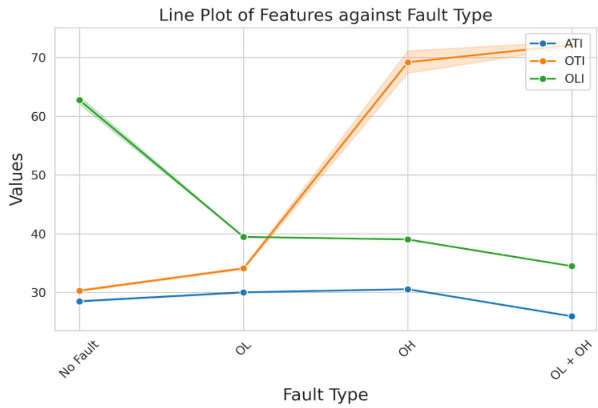


Fig 4. Line plot of ATI, OTI, and OLI vs Fault Types

2.4 Model Selection

The models chosen for this study included the RF classifier, DT classifier, K-NN classifier, ET classifier, and XGB classifier [20-23].

- Random forest classifier

The RF classifier constructs decision trees in a multiple way during training. The final output is determined by the majority voting mechanism among the individual trees. This method is renowned for its accuracy and robustness, particularly in managing large datasets with high dimensionality.

- Decision tree classifier

This algorithm creates a model that predicts the target variable's value using a set of simple decision rules built from data attributes. While decision trees are simple to understand, they are prone to overfitting if not properly trimmed.

- K-nearest neighbors classifiers

KNN operates on the principle that similar data points are located close to each other in a feature space. It is straight forward to implement and performs well on smaller datasets but may struggle with larger datasets due to its computational requirements.

- Extra tree classifier

This classifier also makes trees with bit different approach as compared with the RF classifier. ET classifier offers reduced training time and has greater accuracy.

- Extreme gradient boosting

XGB is used normally for greater speed performance and also it has greater flexibility in handling large amount of data with high accuracy.

2.5 Model Training and Testing

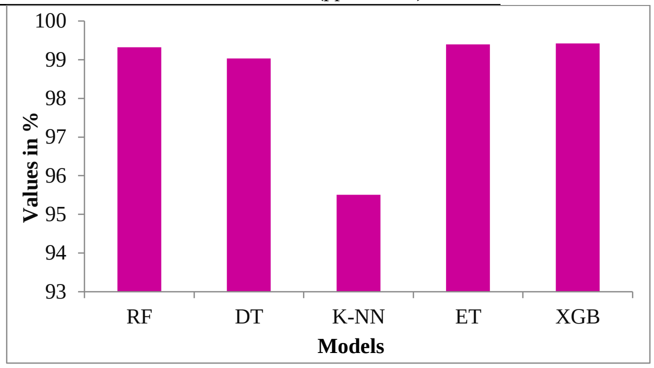
Training of the model was done on 80% resampled dataset and underlying the parameters which are connected with the fault condition. Then the model was tested at 20% dataset. The complete description for dataset is given in Table 5. We conducted stratified k-fold cross-validation with k=10, ensuring equal class proportions in each fold. We have now included the mean and standard deviation of accuracy, precision, recall, and F1-score across folds to report variance and reduce the risk of overfitting. Additionally, to ensure generalizability and mitigate overfitting, we evaluated the models on a completely independent 15% final test set that was not involved in any data balancing or training processes. To evaluate generalizability, we tested the models on a test set (15%) that was taken from the original imbalanced dataset (without resampling). The results show strong performance, validating that the model does not merely overfit to the balanced dataset.

Table 5: Shows the shape of the train and test set

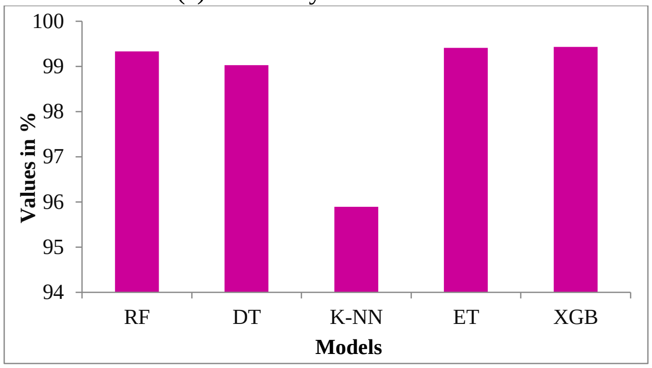
Total Samples	Train Set Samples	Test Set Samples
19042	15233	3809

2.6 Model Evaluation

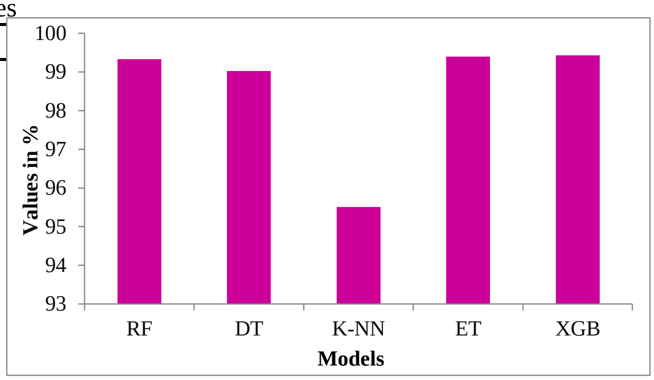
The evaluation of model was done using the key parameters like F1-score, recall, precision and accuracy. After the cross validation, the performance evaluation results are briefly described in Figure 5(a-d) and Figure 6(a-d). The designed model has the capability to classify the fault effectively in power transformer. Furthermore, the model performance of each model was evaluated and compared with possible benefits and weaknesses.



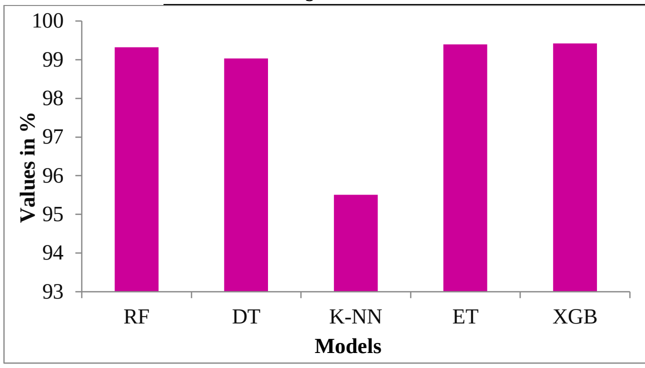
(a): Accuracy



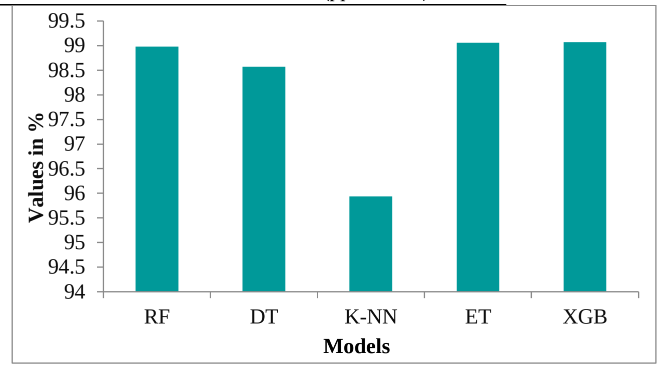
(b): Precision



(c): Recall

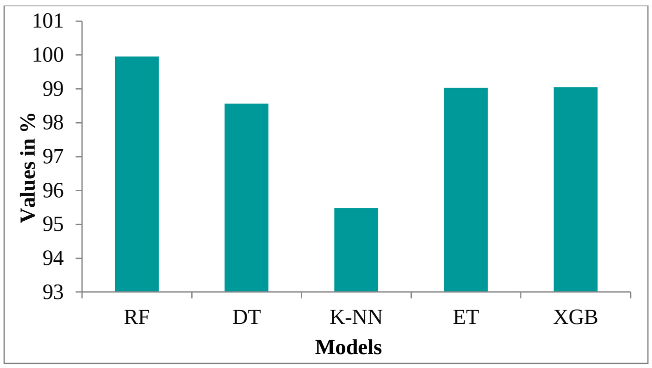


(d): F1 Score

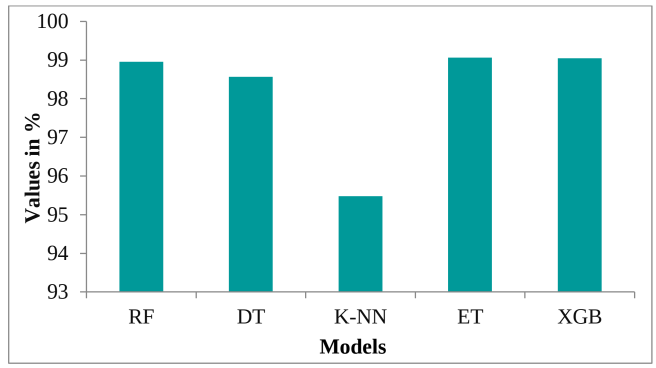


(b): Precision

Fig 5. Performance evaluation on test data

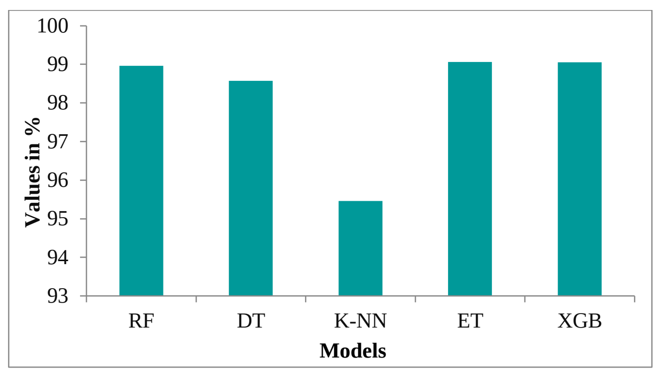


(a): Accuracy



(c): Recall

Further, the transparency in hyperparameter tuning. Each classifier's hyperparameters were optimized using GridSearchCV with 5-fold stratified cross-validation to ensure robustness. Additionally, manual tuning was performed through a hit-and-trial approach. The final tuned parameters are now included in a new section dedicated to hyperparameter tuning.



(d): F1 Score

Fig 6. Performance evaluation after cross validation

3. Results and Discussion

The machine learning models like RF classifier, DT classifier, K-NN classifier, ET classifier and XGB classifier were used in this study for the fault prediction and classification in power distribution transformer. The results of performance metrics on the test data is given in Table 6.

Table 6: Performance metrics on test data

Model	F1-Score	Accuracy	Recall	Precision
	99.32	99.32	99.32	99.33
DT	99.03	99.03	99.03	99.03
K-NN	95.54	95.51	95.51	95.89
ET	99.40	99.40	99.40	99.41
XGB	99.42	99.42	99.42	99.43

3.1 Model performance analysis

- **Random Forest Classifier**

Overall, 99.32% accuracy was achieved by the RF classifier which shows the robustness in handling dataset. The recall and precision also helps to minimize the false negative and false positives. Further, 99.32% F1-score were achieved that helps to create a balance between recall and precision. The confusion matrix for RF classifier is shown in Figure 7.

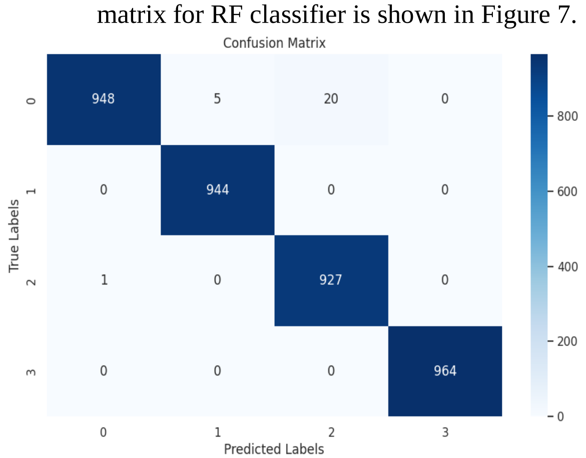


Fig 7. Confusion matrix for random forest classifier

- Decision tree classifier

The DT classifier demonstrated a high accuracy of 99.03%, with corresponding precision and recall metrics of 99.03%. This indicates its strong performance; however, it may still be prone to overfitting when compared to ensemble methods. The confusion matrix for DT classifier is shown in Figure 8.

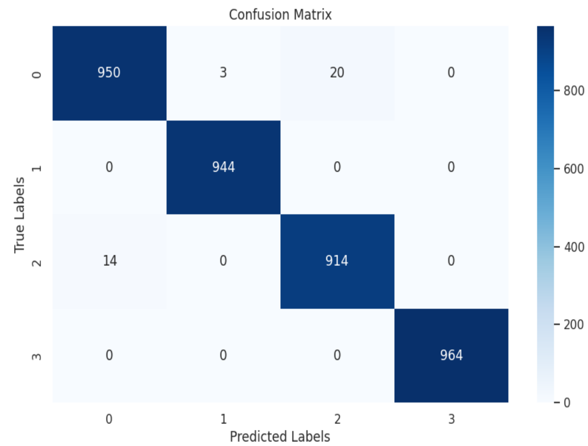


Fig 8. Confusion matrix for decision tree classifier

- K-nearest neighbors classifiers

The KNN classifier recorded an accuracy of 95.51%, supported by relatively high precision and recall metrics. The F1-score of 95.54% shows that while it performs adequately, it is less effective than the ensemble methods in handling the complexity of the dataset. The confusion matrix for KNN classifier is shown in Figure 9.

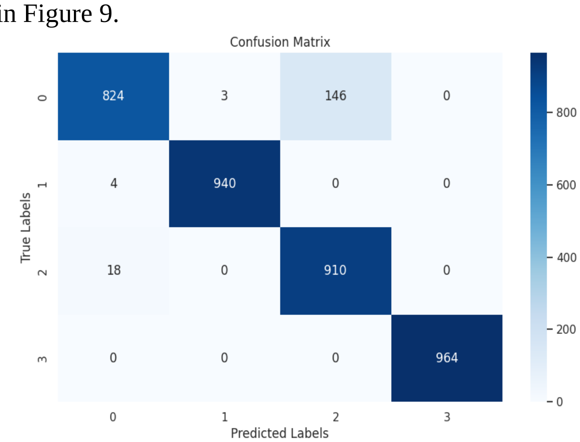


Fig 9. Confusion matrix for KNN classifier

- Extra tree classifier

The ET classifier outperformed most models with an accuracy of 99.40%, showcasing impressive precision and recall scores. The

F1-score of 99.40% highlights its capability to maintain a balance between precision and recall, resulting in robust fault prediction. The confusion matrix for ET classifier is shown in Figure 10.

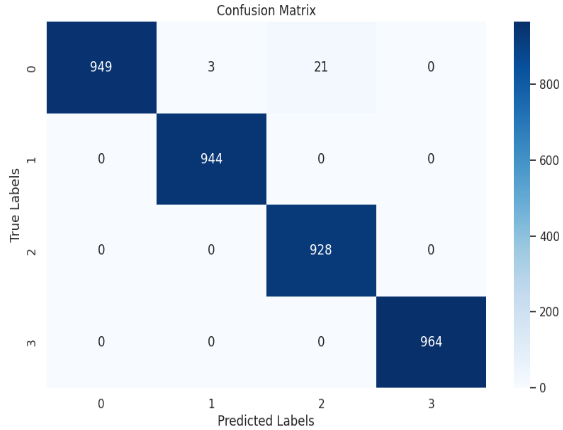


Fig 10. Confusion matrix for extra tree classifier

- Extreme gradient boosting

The XGB model emerged as the best performer with an accuracy of 99.42%, demonstrating superior precision, recall, and F1-score metrics. This underscores its robustness in predicting fault types, showcasing its effectiveness in managing the intricacies of the dataset. The confusion matrix for XGB classifier is shown in Figure 11.

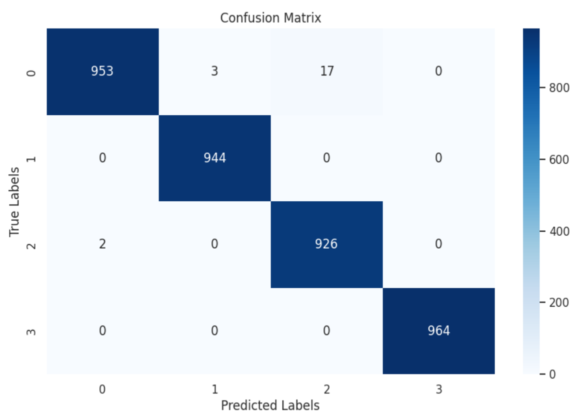


Fig 11. Confusion matrix for XGBoost classifier

4. Comparative analysis of model performance analysis

The comparative analysis indicates that while all models performed exceptionally well, XGBoost and Extra Trees Classifiers exhibited significantly higher performance across all metrics. This performance can be linked to their ensemble nature, which leverages multiple decision trees to enhance accuracy and generalization.

The use of SMOTE Tomek technique effectively addressed the initial class imbalance within the dataset. The resampled dataset allowed for improved learning and evaluation, providing a fair comparison across all models. The models demonstrated increased resilience against bias toward the majority class, which is crucial in fault detection scenarios. Furthermore, the performance comparison of proposed model with the existing models in the literature is presented in Table 7.

Table 7: Performance comparison of machine learning models for fault detection

Models	**Accuracy**
This study used RF, DT, KNN, ET and XGB	99.99%
Support Vector Machine [24]	99.90%
Multilayer Perceptron [25]	95.3%
Convolutional Neural Network [26]	86.82%
Artificial Neural Network [27]	97%

We have validated our simulation results (model predictions) by performing the following:

- Using an independent final test set (15% of the original data) for final evaluation.
- Cross-validation (5-fold stratified) to assess performance variance.
- Comparing results against recent literature to ensure consistency and competitive accuracy. This multi-pronged validation provides confidence

that the simulation (model) results reflect real-world predictive capability.

5. Limitations and Future Work

Despite the promising results, certain limitations were noted in this study. The models were trained on a specific dataset, and their performance may vary with other datasets or under different operational conditions. Future work could involve exploring advanced ensemble techniques, hybrid models, or deep learning approaches to further enhance fault prediction accuracy.

6. Conclusion

In summary, this study successfully applied various machine learning models for the prediction and classification of faults in distribution transformers. The findings demonstrate that ensemble methods, particularly XGB and ET classifier, provide superior predictive capabilities compared to traditional models. These insights pave the way for improved fault management strategies in distribution transformer maintenance. My approach supports sustainable maintenance practices and grid resilience by leveraging predictive maintenance technologies that enable proactive, data-driven management of grid assets. Through advanced sensors, real-time monitoring, and AI-powered analytics, potential equipment failures are anticipated and addressed before causing disruptions. This minimizes unnecessary maintenance activities and reduces resource consumption, thereby extending asset lifespans and lowering environmental impacts.

Acknowledgement

The authors are highly thankful to the Mehran University of Engineering and Technology, SZAB Campus, Khairpur Mir's Pakistan.

References

[1] V. A. Thiviyanathan, P. J. Ker, Y. S. Leong, F. Abdullah, A. Ismail, and M. Z. Jamaludin, "Power transformer insulation system: A review on the reactions, fault detection, challenges and future prospects," *Alexandria Engineering Journal*, vol. 61, pp. 7697-7713, 2022.

[2] L. Jin, D. Kim, A. Abu-Siada, and S. Kumar, "Oil-immersed power transformer condition monitoring methodologies: A review," *Energies*, vol. 15, p. 3379, 2022.

[3] M. Khan, M. A. Raza, M. Faheem, S. A. Sarang, M. Panhwar, and T. A. Jumani, "Conventional and artificial intelligence based maximum power point tracking techniques for efficient solar power generation," *Engineering Reports*, p. e12963, 2024.

[4] I. B. Taha, S. Ibrahim, and D.-E. A. Mansour, "Power transformer fault diagnosis based on DGA using a convolutional neural network with noise in measurements," *IEEE Access*, vol. 9, pp. 111162-111170, 2021.

[5] M. S. Ali, A. Omar, A. S. A. Jaafar, and S. H. Mohamed, "Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review," *Electric Power Systems Research*, vol. 216, p. 109064, 2023.

[6] M. Khan, M. A. Raza, T. A. Jumani, S. Mirsaeidi, G. Abbas, E.-D. Touti, *et al.*, "Modeling Of Intelligent Controllers for Solar Photovoltaic System Under Varying Irradiation Condition," *Frontiers in Energy Research*, vol. 11, p. 1288486, 2023.

[7] P. D. Singh, P. Borah, A. Perme, H. Allo, A. Chatterjee, and R. K. Lohar, "IoT-based Online Condition Monitoring of Power Transformer," in *2024 6th International Conference on Energy, Power and Environment (ICEPE)*, 2024, pp. 1-6.

[8] J. F. Piñeros, D. A. Echeverri, and L. M. Benavides, "Automatic fast calculation and verification method

for transformer overcurrent protection settings," *International Journal of Electrical Power & Energy Systems,* vol. 157, p. 109802, 2024.

[9] N. Chothani, M. Raichura, and D. Patel, "An Overview of the Protection of Power Transformers," *Advancement in Power Transformer Infrastructure and Digital Protection,* pp. 27-69, 2023.

[10] S. A. Sarang, M. A. Raza, M. Panhwar, M. Khan, G. Abbas, E. Touti, *et al.*, "Maximizing solar power generation through conventional and digital MPPT techniques: a comparative analysis," *Scientific Reports,* vol. 14, p. 8944, 2024.

[11] M. N. Ali, M. Amer, and M. Elsisi, "Reliable IoT paradigm with ensemble machine learning for faults diagnosis of power transformers considering adversarial attacks," *IEEE Transactions on Instrumentation and Measurement,* 2023.

[12] R. M. A. Velásquez and J. V. M. Lara, "Root cause analysis improved with machine learning for failure analysis in power transformers," *Engineering Failure Analysis,* vol. 115, p. 104684, 2020.

[13] D. Yang, J. Qin, Y. Pang, and T. Huang, "A novel double-stacked autoencoder for power transformers DGA signals with an imbalanced data structure," *IEEE Transactions on Industrial Electronics,* vol. 69, pp. 1977-1987, 2021.

[14] S. M. A. N. Al-Ameri, M. S. Kamarudin, M. F. M. Yousof, A. A. Salem, F. A. Banakhr, M. I. Mosaad, *et al.*, "Understanding the influence of power transformer faults on the frequency response signature using simulation analysis and statistical indicators," *IEEe Access,* vol. 9, pp. 70935-70947, 2021.

[15] M. A. Raza, K. L. Khatri, and A. Hussain, "Transition from fossilized to defossilized energy system in Pakistan," *Renewable Energy,* vol. 190, pp. 19-29, 2022.

[16] M. A. Raza, M. M. Aman, N. A. Tunio, S. A. Soomro, M. Shahid, D. Ara, *et al.*, "Energy transition through bioelectricity in Pakistan: Implications for limiting global mean temperature below 1.5 C," *Environmental Progress & Sustainable Energy,* vol. 42, p. e14189, 2023.

[17] S. A. Gamel, S. S. Ghoneim, and Y. A. Sultan, "Improving the accuracy of diagnostic predictions for power transformers by employing a hybrid approach combining SMOTE and DNN," *Computers and Electrical Engineering,* vol. 117, p. 109232, 2024.

[18] Z. Ahmed and S. Das, "A Novel Hybrid Resampling Approach to Address Class-Imbalanced Issues," *SN Computer Science*, vol. 5, p. 865, 2024.

[19] S. S. Ghoneim, T. A. Farrag, A. A. Rashed, E.-S. M. El-Kenawy, and A. Ibrahim, "Adaptive dynamic meta-heuristics for feature selection and classification in diagnostic accuracy of transformer faults," *Ieee Access*, vol. 9, pp. 78324-78340, 2021.

[20] N. Haque, A. Jamshed, K. Chatterjee, and S. Chatterjee, "Accurate sensing of power transformer faults from dissolved gas data using random forest classifier aided by data clustering method," *IEEE Sensors Journal,* vol. 22, pp. 5902-5910, 2022.

[21] Z. Wu, M. Zhou, Z. Lin, X. Chen, and Y. Huang, "Improved genetic algorithm and xgboost classifier for power transformer fault diagnosis,"

Frontiers in Energy Research, vol. 9, p. 745744, 2021.

[22] Y. Benmahamed, O. Kherif, M. Teguar, A. Boubakeur, and S. S. Ghoneim, "Accuracy improvement of transformer faults diagnostic based on DGA data using SVM-BA classifier," *Energies*, vol. 14, p. 2970, 2021.

[23] M. A. Raza, M. Aman, G. Abbas, S. A. Soomro, A. Yousef, E. Touti, *et al.*, "Managing the low carbon transition pathways through solid waste electricity," *Scientific Reports*, vol. 14, p. 5490, 2024.

[24] G. Ciaburro, "Machine fault detection methods based on machine learning algorithms: A review," *Mathematical Biosciences and Engineering*, vol. 19, pp. 11453-11490, 2022.

[25] K. M. Mishra and K. J. Huhtala, "Fault detection of elevator systems using multilayer perceptron neural network," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2019, pp. 904-909.

[26] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, *et al.*, "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331-345, 2016.

[27] A. Et-Taleby, Y. Chaibi, M. Benslimane, and M. Boussetta, "Applications of machine learning algorithms for photovoltaic fault detection: a review," *Statistics, Optimization & Information Computing*, vol. 11, pp. 168-177, 2023.