© March 2025 | IJIRT | Volume 11 Issue 10 | ISSN: 2349-6002

Power Transformer Health Monitoring using Machine Learning

Priya Dhule[1], Priyanka Pawar[2], Tanvi Girhe[3], Shital Gavhale[4], Snehal Ingle[5], Prof. Nilesh G. Bundhe[6]
[1,2,3,4,5] *Padm. V. B. Kolte College of Engineering, Malkapur, India*
[6]*V. B. Kolte College of Engineering, Malkapur, India*

***Abstract*—It is essential for power transformers, which are crucial parts of electrical networks, to function dependably in order to keep the system stable. It is possible for transformer faults to result in significant disruptions and costly repairs if they are not accurately diagnosed and classified. The purpose of this study is to present a machine learning algorithm for transformer fault classification that makes use of failure history data and advanced pattern recognition algorithms. This approach to transformer operating scenario classification makes use of decision trees, support vector machines, and latent differential equations. In order to improve the accuracy of categorisation, the data is preprocessed. A 5-fold validation validates the performance of the model. The findings demonstrate that machine learning improves predictive maintenance, grid dependability, and transformer health monitoring compared to traditional methods.**

***Index Terms*—Transformer Fault Diagnosis, Machine Learning, Support Vector Machine.**

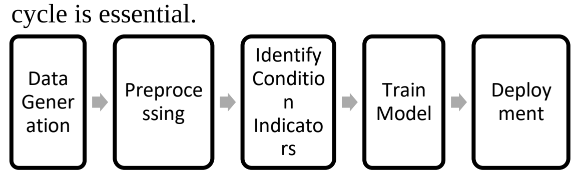
I. INTRODUCTION

Transformers facilitate the transmission and distribution of electricity across voltage levels in an effective manner inside power networks. They are dependent on substations, and as a result, their dependability has an effect on the performance of the system. In order to preserve power and extend the life of equipment, transformers require failure protection on both the inside and the outside. Short circuits, overheating, partial discharges, and insulation that is failing are all examples of internal issues that could potentially compromise the reliability of their operation. The detection of problems in a timely and correct manner helps to extend the life of transformers, decrease the costs of maintenance, and prevent catastrophic failures. Despite the fact that traditional diagnostic procedures such as FRA and DGA give valuable information, there are frequently problems with accuracy and the ability to apply them in real time. Recent developments in machine learning and signal processing have given rise to the usefulness of data-driven fault classification and detection. In the absence of differential protection, through faults are external faults that result in an excessive amount of current flowing through the transformer. It is possible for these failures to be caused by short circuits or disturbances in the phase-to-ground network. feeder lines were connected. Transformers have the ability to withstand such loads within the short-circuit tolerance limits; nevertheless, prolonged exposure leads to insulation and winding degradation for the transformer.

The use of machine learning was investigated by K. Premalatha and colleagues for the purpose of transformer failure categorisation and diagnosis [1]. It has been demonstrated that conventional methods of defect identification are both speedy and inaccurate. Both historical and current data are utilised by the authors in order to classify transformer issues through the application of machine learning. Wavelet transform-based feature extraction for equipment failure diagnostics was proven to be effective in managing complex and non-stationary data by A. Vyshnavi and colleagues [2]. Comparative analysis of wavelet-based fault classification methods that make use of machine learning is presented in this work. For the purpose of accurate fault diagnosis, they placed a priority on modern signal processing. For the purpose of enhancing fault diagnosis, Hailong Ma and colleagues devised a new method for the identification of hidden faults in power transformers [3]. It was suggested that a fault detection system may be developed by utilising an IPSO algorithm with a BP neural network. In order to prevent premature convergence and improve optimisation in traditional PSO, they made use of dynamic inertia weight modulation and nonlinear learning factor modification. Using machine learning, A. Balan and colleagues [4] are able to diagnose defects in transformers. Methods that have been traditionally used to diagnose transformer breakdown have been demonstrated to be ineffective and erroneous. In the study, the use of data-driven machine learning resulted in improvements to fault prediction and categorisation. H. Hadiki, F. S. Hasnaoui, and S. Georges conducted research on the application of machine learning to the prediction of transformer defects [5]. Precision, reaction time, and the cost of maintenance were evaluated for traditional defect detecting devices based on their performance. In order to forecast transformer failures, a prediction system that is based on machine learning and makes use of transformer data was proposed. T.-H. Han and colleagues came up with a Superconducting Fault Current Limiter (SFCL) that is a transformer-type three-phase device for the purpose of limiting fault currents in power systems [6]. The primary objective of their research is to find ways to effectively limit three-phase ground fault currents in order to enhance the safety and stability of electrical network systems. A novel approach to fault current mitigation makes use of two superconducting modules (SCMs) that are incorporated into the secondary windings of a three-phase transformer [6]. Through the application of the any-shot learning problem, Yue Zhuo and colleagues were able to address industrial fault detection problems that involved defect samples that were either scarce or unobtainable [7]. Using Generative Adversarial Networks was the solution that the specialists recommended for avoiding this problem. A valid diagnostic model was constructed with the assistance of this by gathering a large number of samples for a variety of illnesses. A new method for evaluating the failure of transformers was created by J. H. Estrada and colleagues [8]. This method makes use of magnetic flux entropy to identify distortions, effects of ageing, overloading, and harmonics. Using concepts related to entropy, the research investigated the non-invasive measurement of thermal and magnetic fluxes in transformer prototypes.

II. MACHINE LEARNING CLASSIFICATION

A machine learning workflow is a methodical process that provides a framework for the development and deployment of machine learning models. In most cases, it starts with the acquisition of data and the preprocessing of that data, which involves gathering and cleaning the necessary data [9]. The following phase is involved in feature engineering, which is the process of selecting or creating the variables that are the most informative in order to train the model. After the data has been further segmented into training and testing sets, the machine learning algorithm that has been chosen should then be trained on the training dataset [10]. The performance of the model on data after it has been trained is subjected to evaluation in order to determine its accuracy and generalisation capabilities. In the event that the outcomes do not satisfy the requirements, the process will iterate with more parameter tuning or method selection. If the outcomes do not meet the requirements, the model will be developed for practical application. It is essential to continuously check and maintain the performance of the model in order to guarantee that it is able to adjust to shifting data patterns and preserve its accuracy [11]. When it comes to utilising the potential of machine learning for tasks such as predictive analytics, classification, regression, and other similar activities, this iterative cycle is essential.



III. SUPPORT VECTOR MACHINE

It is possible to resolve the binary problem by employing the SVM classifier model [12], which also maps the linear datasets from low-dimensional to excessive-dimensional function space. The appropriate hyperplane is comprised of a distinct subset of samples that are located between different classes. Given a dataset D = {(x1, y1), (x2, y2), …, (xn, yn)}, the schooling is represented by the variable x, where x is the input feature and y is the result of the selection label for {-1, 1}. There is a possibility that the hyperplane that corresponds to the choice to divide the samples could be either a positive or a negative hyperplane. This is the formula:

$$y_i(W^T X_i + b) \geq 1, \forall i \quad (1)$$

The offset vector for a scalar threshold, which represents the margins, is denoted by the letter b, while the load, which indicates the direction of the hyperplane, is denoted by the letter w. One can calculate each sample x by calculating the distance between the hyperplane (w,b) and the margin that is the greatest for that particular sample. Every single sample ought to be able to fulfil the requirements for the input feature vectors for both classes. To solve the optimisation problem, the equation is as follows:

$$\min \frac{1}{2}\|w\|^2 + C \sum_i \xi^i$$

Subject to

$$y_i(W^T X_i + b) \geq 1 - \xi^i, \xi^i \geq 0, \forall i \quad (2)$$

Given that ξ represents the slack variable of the proscribing boundary and C represents the manage that corresponds to the misclassification penalty among the greatest and minimum margins, the following equation can be obtained. The twin problem can be solved by transforming the linear class problem into a nonlinear category problem with a dimensional feature space and employing the kernel feature of linear SVM algorithms. This will allow for the twin problem to be solved. By utilising this method, the two times of mapping, which are referred to as mapping from low-dimensional to high-dimensional features, are represented.

$$K(x, x_i) = e^{\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)} \quad (3)$$

where the symbol σ represents the standard deviation and is an effective true cost expression. It is possible to use the kernel feature of equation (3) to map the features in order to solve the twin problem of linear support vector machine analysis for the purpose of simplifying the Lagrange equation of equation (2). It is possible to obtain the answer by employing a nonlinear category in the manner that is defined under:

If αi is the multiplier of the Lagrange equation, then

$$f(x) = sign\{\sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + b\} \quad (4)$$

should meet the criterion range of 0 < αi < C.

IV. RESULTS AND DISCUSSION

An 11th generation Intel Core i3 machine with a computational speed of 3 GHz is used to execute all of the calculations and model training. MATLAB 2017a is used for this calculation and training. Condition indicators were added after all of the features had been extracted, and they were based on the description of the dataset. The entire set of data is then divided into two parts: seventy percent of the data is retained for the purpose of training the model, and thirty percent of the data is retained for the purpose of validation. Table I displays the levels of efficiency achieved by models that have been successfully trained.

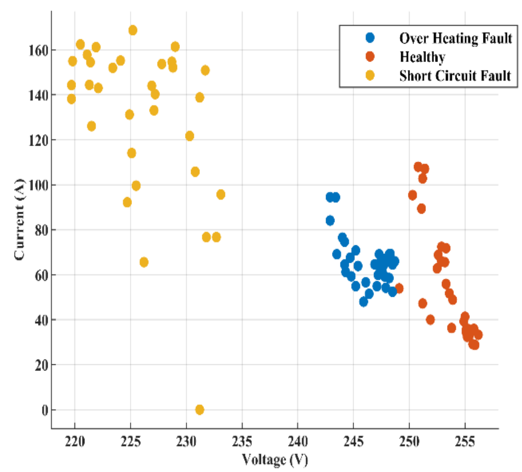


Figure 1: Scatter Plot

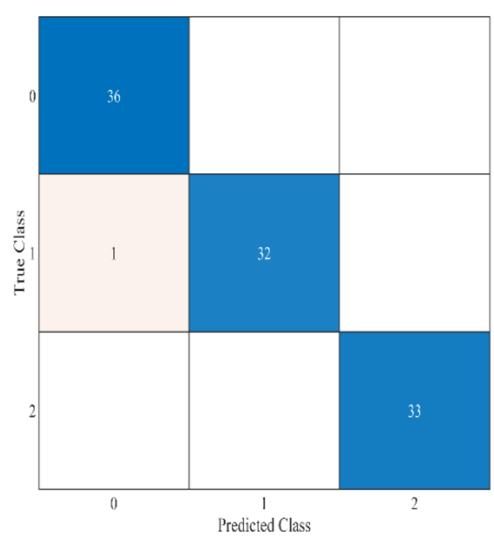


Figure 2: Confusion Matrix for Decision Tree

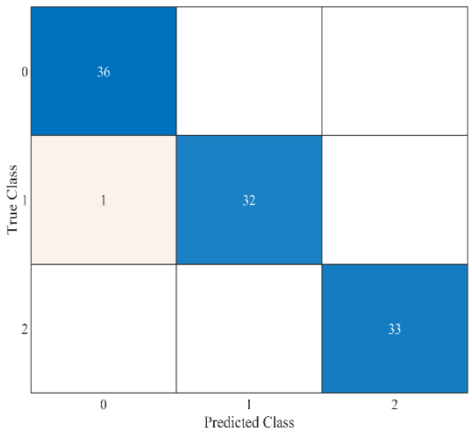


Figure 3: Confusion Matrix for Naive Bayes

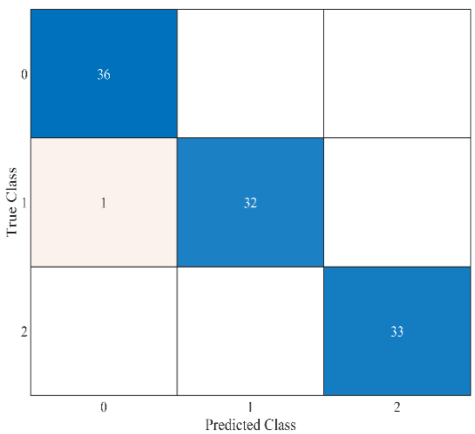


Figure 4: Confusion Matrix for SVM

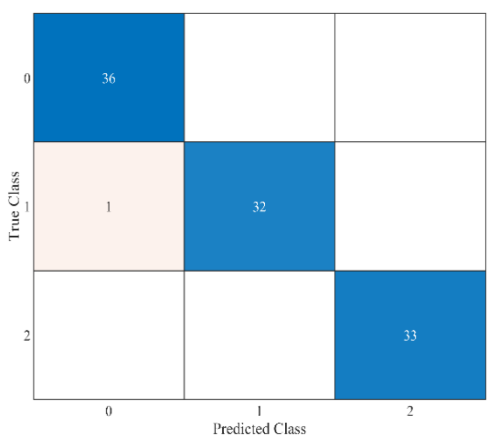


Figure 4: Confusion Matrix for SVM after hyperparameter tuning

Table 1: Comparison of classification accuracy

Algorithm	Accuracy
Decision Tree	99 %
Naïve Bayes	99 %
Support Vector Machine	99 %
SVM with Hyperparameter tuning	100 %

As can be observed from the confusion matrices, the classification accuracy achieved by Decision Tree, Nave Bayes, and Support Vector Machines was 99%. As can be seen from the scatter plot, the data points are quite distinguishable; hence, the hyperparameter adjustment is done in order to improve the accuracy of the classification. As can be seen from the confusion matrix, the authors then proceeded to modify the hyperparameters in order to attain a precision of one hundred percent.

V. CONCLUSION

The purpose of this research is to propose a framework that can properly classify the operating characteristics of transformers as either healthy or malfunctioning. Through the application of wavelet transform, the current and voltage values of the transformer are utilised in order to extract significant characteristics. Following the extraction of features, a number of machine learning algorithms are trained using the extracted data. A total of five different methods have been evaluated for the problem that is being discussed. Out of all of them, the LDA algorithm was found to have the best performance, achieving a fault classification accuracy of one hundred percent. There is a possibility that the authors will use time-frequency features in the future to train a deep learning network, which could lead to the development of alternative approaches in this field.

REFERENCES

[1] K. Premalatha, R. Janaranjani, S. Bhuvaneswari and M. Karthik, "Transformer Fault Identification and Classification of Using Machine Learning," 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, 2024

[2] A. Vyshnavi, M. P. Mali, S. Hemelatha, M. Gupta, M. Sindhu and A. Singla, "Wavelet Transform Based Feature Extraction for Fault Diagnosis in Machinery," 2024 15th

International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024

[3] H. Ma, H. Song, C. Meng and R. Wang, "Fault Diagnosis of Power Transformer Based on Improved Neural Network," 2024 5th International Conference on Computer Engineering and Application (ICCEA), Hangzhou, China, 2024

[4] A. Balan, S. T. L, M. P. V and K. Deepa, "Detection and Analysis of Faults in Transformer using Machine Learning," 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2023

[5] H. Hadiki, F. S. Hasnaoui and S. Georges, "Transformers Faults Prediction Using Machine Learning Approach," 2023 Fifth International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Zouk Mosbeh, Lebanon, 2023

[6] T. -H. Han, S. -C. Ko and S. -H. Lim, "Analysis on Three-Phase Ground Fault Current Limiting Operations of Three-Phase Transformer Type SFCL Using Two Superconducting Modules," in IEEE Transactions on Applied Superconductivity, vol. 32, no. 6, pp. 1-7, Sept. 2022

[7] Y. Zhuo and Z. Ge, "Auxiliary Information-Guided Industrial Data Augmentation for AnyShot Fault Learning and Diagnosis," in IEEE Transactions on Industrial Informatics, vol. 17, no. 11, pp. 7535-7545, Nov. 2021

[8] J. H. Estrada, S. V. Ramı´rez, C. L. Cortés and E. A. C. Plata, "Magnetic Flux Entropy as Tool to Predict Transformer's Failures," in IEEE Transactions on Magnetics, vol. 49, no. 8, pp. 4729-4732, Aug. 201

[9] H. Tan, “Machine Learning Algorithm for Classification,” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1994/1/012016.

[10] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Computer Science*, vol. 2, no. 3. Springer, May 01, 2021. doi: 10.1007/s42979-021-00592-x.

[11] A. Stetco *et al.*, “Machine learning methods for wind turbine condition monitoring: A review,” *Renewable Energy*, vol. 133. Elsevier Ltd, pp. 620–635, Apr. 01, 2019. doi: 10.1016/j.renen,m

[12] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20 (1995): 273-297