Journal of Materials and Electronic Devices 1 (2025) 21-28

FYTRONIX JOURNAL OF MATERIALS AND ELECTRONIC DEVICES JMED http://dergi-fytronix.com/

An Artificial Intelligence–Based Approach for Fault Detection and Classification in Electric Power Systems

Merve PARLAK BAYDOGAN

Technical Sciences Vocational School, Computer Technologies Department, Firat University, Elazig, TURKEY

In this study, a machine learning-based approach is developed for the detection and classification of faults occurring in electric power systems. A dataset containing different fault scenarios was obtained from the Kaggle platform, and a comprehensive performance analysis was performed by applying various machine learning algorithms to this data. The models were compared within the scope of fault detection and fault classification tasks. In the analysis, Decision Tree (DT), Random Forest (RF), XGBoost, LightGBM (LGBM), Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Naive Bayes (NB), Logistic Regression (LR), and Multi-Layer Perceptron (MLP) algorithms were evaluated. The results revealed that tree-based and ensemble methods exhibited higher accuracy, reliability, and stability compared to other models. In particular, the RF and XGBoost algorithms stood out as the most successful models in terms of fault detection and classification. The study findings demonstrate that the proposed approach enables rapid, accurate, and reliable fault detection in power systems. Future studies are recommended to test the model with larger, more balanced datasets from different systems and assess its applicability in real-time conditions.

Key words: *Machine learning, Fault detection, Fault classification, Electrical power systems, Ensemble methods*

Submission Date: 16 June 2025
Acceptance Date: 23 August 2025

**Corresponding author: mpbaydogan@firat.edu.tr*

1. Introduction

Today's technological advances in electrical energy are crucial for the uninterrupted and efficient distribution of energy. The pace of industrialization in societies is directly affected by the availability of affordable energy. Electric power systems are critical infrastructure systems comprised of interacting, dynamic, and complex components such as generation, transmission, and distribution. Any disruption or malfunction in these systems, if not detected and resolved promptly, can lead to serious system instabilities, cascading failures, regional blackouts, and significant financial losses. The reliable and stable operation of these systems is crucial for minimizing negative impacts on the industrial, transportation, business, and residential sectors and preventing potential financial losses. Transmission lines, in particular, stand out as one of the most critical components within these systems. Increasing the efficiency of the production sector depends significantly on the quality of the energy transmission/distribution lines extending from the power plant to the factory. Rural power transmission lines are located in less protected areas than distribution power plants. In these regions, potential faults due to interruptions have necessitated minimizing the impact of physical events on these lines. Today, the power distribution system in some major cities is entirely carried out via underground cables. Numerous methods have been developed over time to address potential faults in these cabling systems [1-3].

One of the primary factors causing instability in electricity distribution networks is the widely known phenomenon of sudden electrical faults occurring on transmission or distribution lines. Deviations in current voltage and current values from nominal values can be caused by a variety of factors, including complex environmental conditions, natural disasters (floods, storms, lightning), equipment failures (e.g., insulation failure), sudden load changes, and overheating. Furthermore, in the event of an electrical fault, overcurrents that can damage various apparatus and equipment travel through the network. In such cases, the overall reliability of the power system can be jeopardized, resulting in increased maintenance costs. Early and accurate detection of faults in electricity distribution networks is essential to reduce equipment damage, service disruptions, and loss of human and animal life [4-7].

Faults occurring on transmission lines (TL), one of the most critical components of the EGS, can be permanent or transient. These faults are also classified as symmetrical or asymmetrical short circuit faults. Asymmetrical faults occur as single line-to-ground (SLG), double line-to-ground (LLG), and line-to-line (LL), while symmetrical faults occur as three-phase (LLL) or three-phase-to-ground (LLLG) short circuits [8]. Traditional model-based techniques are widely used to detect these faults. These techniques include impedance-based and traveling wave (TW)-based methods. Although these methods are widely used, they have serious limitations, such as being time-consuming, computationally intensive, and lacking adaptability [9]. While traditional protection systems and rule-based approaches can yield successful results under certain fault scenarios, they may be limited by variable load conditions, measurement noise, complex fault types, and new-generation system dynamics [9,10]. In this context, machine learning, and particularly deep learning techniques, offer a new and effective solution for electrical fault detection and classification. In recent years, significant research has been conducted on the development of machine learning-based approaches for detecting electrical faults [11].

This study proposes a deep learning-based approach for the detection and classification of faults occurring in electrical transmission lines. In the proposed method, using time series data obtained from three-phase voltage and current signals, multiple convolutional neural networks (CNN) architectures enable both fault detection and high-accuracy fault classification. Additionally, various machine learning algorithms were applied in the study for performance comparison. In this context, the following are included: Decision Trees (DT), k-Nearest Neighbors (kNN), Naive Bayes (NB), Support Vector Machines (SVM), Random Forests (RF), Logistic Regression (LR), AdaBoost, XGBoost, LightGBM) and Multilayer Perceptron (MLP) algorithms were used. The remainder of the study is organized as follows: The second section provides a literature review and summarizes previous studies. The third section presents the dataset, feature extraction, and preprocessing steps. Furthermore, the proposed method and the machine learning and deep learning algorithms used are explained in detail. The fourth section evaluates the experimental results and compares the performance of different algorithms. The final section provides a general assessment and offers suggestions for future work.

2. Literature Review

This section reviews literature studies on the detection and classification of electrical faults and presents existing approaches to the topic.

In a study conducted by Goni et al., the Extreme Learning Machine (ELM) algorithm was proposed for fast and accurate fault detection and classification on transmission lines. This method, which offers lower computational complexity and shorter processing time compared to traditional artificial neural networks, was tested on two different TL configurations in the MATLAB Simulink environment. Data from ten different fault types were preprocessed using min-max normalization, eliminating the need for additional transformation techniques. The ELM model significantly reduced the learning time by determining the output weights in a single step. The obtained results achieved over 99% accuracy in fault classification and performed approximately 7.6 times faster than ANN in terms of prediction time [1].

Jamil and dig proposed a system for the detection and classification of faults occurring on electric power transmission lines using artificial neural networks (ANNs). In the proposed method, three-phase currents and voltages at one end are taken as inputs. The voltage and current values of the three phases (a total of six inputs) are fed to the neural network. The overall mean square error of the trained neural network is found to be 0.036043. Examining the confusion matrices prepared for the training, validation, and testing phases (Figure 8), it is seen that the selected neural network has 100% accuracy in fault detection. The ability (efficiency) of the trained neural network to distinguish between ten possible fault types is determined to be 78.1%. They report that the neural network can distinguish all ten possible fault types on the transmission line [2].

Guo and Dig developed a DL approach using the Hilbert-Huang Transform (HHT) and Convolutional Neural Network (CNN) for fault classification in power distribution systems. In this study, data obtained from the PSCAD/EMTDC model simulation and the physical system model built in the laboratory were used, and the fault conditions (type, location, resistance, and onset angle) were considered in detail. The method involves generating a time-frequency energy matrix (35x40) from seven different fault signals (three-phase voltage, three-phase current, and zero-sequence voltage) through an HHT band-pass filter and feeding this matrix as input to the CNN. This approach achieved high performance with an average classification accuracy of 99.92% for ten different fault types and has been reported to offer faster execution time (approximately 0.2125 s) compared to other methods [12]. In Roy et al.'s study, an LSTM-based deep learning approach was developed for fault detection, classification, and online fault location determination in grid-connected microgrid (MG) systems. LSTM network was used for fault detection and classification, and a combination of LSTM and Feed Forward Artificial Neural Network (FFNN) was used for location estimation. A MG model, including a solar PV, wind turbine, diesel generator, and battery, was created in the MATLAB/Simulink environment, and 11 different fault types were simulated at three different locations. The findings show that the LSTM-based method provides higher accuracy and speed compared to ANN, and that the proposed combination demonstrates successful performance in location determination [13].

Dasgupta et al. proposed an approach using the K-Nearest Neighbor (k-NN) algorithm with cross-correlation-based feature extraction for transmission line fault detection and classification. The method extracts features using ECG-like cross-correlograms to measure the relationship between post-fault and pre-fault current signals and analyzes these data with a k-NN classifier. The synthetically generated dataset was obtained from half-cycle pre- and post-fault current signals. In experiments conducted with a dataset containing a total of 11 fault types, an accuracy rate of 99.67% was achieved; Furthermore, it has been reported that the k-NN algorithm performs similarly or even better with lower computational overhead compared to methods such as DWT+ANN and SVM [14].

In this study, the proposed system provides a comprehensive comparison of different machine learning and deep learning algorithms for the multi-class fault classification problem. In this context, Decision Tree (DT), k-Nearest Neighbor (kNN), Naive Bayes (NB), Support Vector Machines (SVM), Random Forests (RF), Logistic Regression (LR), XGBoost, LightGBM (LGBM), AdaBoost, and Multilayer Perceptron (MLP) algorithms were evaluated. In this respect, the study offers a competitive solution compared to existing methods in the literature in terms of both classification accuracy and computational efficiency.

3. Materials and Methods

In this section, detailed information about the dataset used in the study and the proposed deep learning-based model is presented.

3.1. Materials

In this study, the dataset was sourced from the publicly available Kaggle platform [15]. This dataset was obtained through simulation to train and test artificial neural networks for the detection and classification of faults in a three-phase electric power transmission system. A typical $400\times10^3 V$ three-phase transmission line system with generators at each end was used to create the dataset. The transmission line length was 300 km and was modeled using distributed-type parameters to obtain more accurate results. The power system model was simulated using the SimPowerSystems toolbox in Simulink in the MATLAB® environment. Using the three-phase V-I measurement block, relevant three-phase voltage and current samples were measured at terminal A of the transmission line. The three-phase currents and three-phase voltages of the relevant three phases (A, B, C) were taken as input to the artificial neural networks.

In the proposed system, the dataset used for fault classification contains a total of 7,861 samples and 10 feature columns. Input variables include three-phase current (Ia, Ib, Ic) and voltage (Va, Vb, Vc). Target variables are columns A, B, C, and G, which indicate, as a binary (0 or 1), whether the fault occurs at Phase A, Phase B, Phase C, or ground, respectively. The training set contains a total of 7,861 inputs across six input and four output combinations. Table 3.1 lists some sample output labels and their corresponding fault types.

Table 3.1. Examples of output labels and their associated fault types

Labels (G C B A)	Fault Type	Description
[0 0 0 0]	No Fault	Normal Condition
[1 0 0 1]	LG Fault	Phase A-to-Ground
[0 0 1 1]	LL Fault	Phase A-to-Phase B Fault
[1 0 1 1]	LLG Fault	Phase A–B-to-Ground Fault
[0 1 1 1]	LLL Fault	Three-Phase Fault (without ground)
[1 1 1 1]	LLLG Fault	hree-Phase-to-Ground Fault (Symmetrical Fault)

3.2. Methods

This study proposes a deep learning-based system for the detection and classification of faults occurring in the electric power transmission system. A comprehensive comparative analysis is conducted by applying various machine learning algorithms to the multi-label fault classification problem. The aim is to identify models that are both highly accurate in classification and computationally efficient. In this context, the following algorithms were used: DT, kNN, NB, SVM, RF, LR, XGBoost, LGBM, AdaBoost, and MLP. The working scheme of the proposed system is given in Figure 3.1.

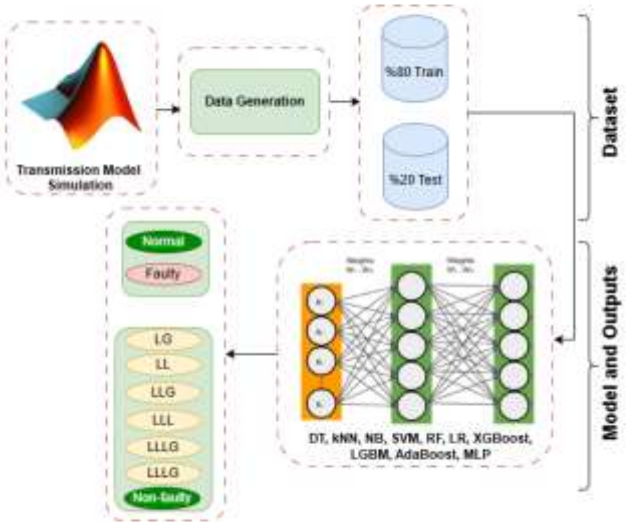


Figure 3.1. Pipeline of the Proposed System

Figure 3.2. shows that the proposed model is structured to detect and classify faults occurring on three-phase electrical transmission lines. The system begins with a simulation-based data generation process in the first phase. This process involves modeling different fault scenarios (e.g., LG, LL, LLG, LLL, LLLG) in the MATLAB/Simulink environment to create a comprehensive dataset. The resulting dataset is divided into two parts: 80% training and 20% test, and these data are used to train various machine learning algorithms. Each algorithm aims not only to determine whether a fault exists but also to accurately classify its type (e.g., LG: phase-to-ground, LL: phase-to-phase, LLG: double phase-to-ground, LLL: three-phase, LLLG: three-phase-to-ground). Consequently, the proposed system offers a flexible and modular structure capable of performing fault detection and detailed fault type classification with high accuracy. Each proposed algorithm was evaluated based on key performance metrics such as accuracy, precision, recall, F1-score, and micro-averaging AUC (AUC-micro). The results were compared to determine the most effective classifier for fault diagnosis [16]. Information regarding the confusion matrix used to calculate the performance evaluation metrics is provided in Table 2. Figure 1 shows that the proposed model is structured to detect and classify faults occurring on three-phase electrical transmission lines. The system begins with a simulation-based data generation process in the first phase. This process involves modeling different fault scenarios (e.g., LG, LL, LLG, LLL, LLLG) in the MATLAB/Simulink environment to create a comprehensive dataset. The resulting dataset is divided into two parts: 80% training and 20% test, and these data are used to train various machine learning algorithms. Each algorithm aims not only to determine whether a fault exists but also to accurately classify its type (e.g., LG: phase-to-ground, LL: phase-to-phase, LLG: double phase-to-ground, LLL: three-phase, LLLG: three-phase-to-ground). Consequently, the proposed system offers a flexible and modular structure capable of performing fault detection and detailed fault type classification with high accuracy. Each proposed algorithm was evaluated based on key performance metrics such as accuracy, precision, recall, F1-score, and micro-averaging AUC (AUC-micro). The results were compared to determine the most effective classifier for fault diagnosis [16]. Information regarding the confusion matrix used to calculate the performance evaluation metrics is provided in Table 3.2.

Table 3.2. Confusion Matrix for Model Evaluation

- TP: Predicted correctly when actually true
- FN: Predicted incorrectly when actually true
- FP: Predicted correctly when actually false
- TN: Predicted incorrectly when actually false

		Actual Values	
		Positive	**Negative**
Predicted Values	**Postive**	**TP**	**FP**
	Negative	**FN**	**TN**

The equations corresponding to the performance metrics are presented below (Equations 1–4).

$$Accuracy = \left(\frac{(TP + TN)}{(TP + FP + FN + TN)}\right) \tag{1}$$

$$Precision = \left(\frac{TP}{(TP + FP)}\right) \tag{2}$$

$$Recall = \left(\frac{TP}{(TP + FN)}\right) \tag{3}$$

$$F1 - score = 2\,x\left(\frac{Precision\ x\ Recall}{Precision + \ Recall}\right) \tag{4}$$

4. Experimental Results and Discussion

This section presents the findings of experimental analyses conducted on the performance of the proposed system. Machine learning algorithms trained using data from various fault scenarios were comprehensively evaluated for both fault detection and fault classification tasks. During the experiments, the effectiveness of each model was compared using common performance metrics such as accuracy, precision, recall, F1-score, and AUC (Area Under the Curve). Thus, the applicability and reliability of the proposed model in real-time systems were empirically tested. The obtained results show that the methods used provide high accuracy and low error rates in fault diagnosis. Table 4.1 presents the results for fault classification.

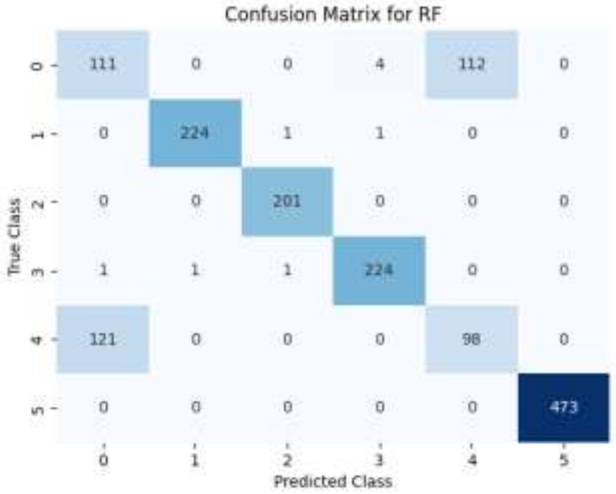
Table 4.1. Fault Type Classification Results

Algorithms	Performance evaluation metrics				
	Accuracy	Precision	Recall	F-Score	AUC (micro)
DT	**0.8538**	**0.8248**	**0.8283**	**0.8265**	**0.9123**
kNN	0.7826	0.7550	0.7509	0.7527	0.8695
NB	0.3865	0.3267	0.2921	0.2705	0.6319
SVM	0.6046	0.5703	0.5501	0.5460	0.7627
RF	0.8462	0.8178	0.8191	0.8184	0.9077
LR	0.3039	0.2539	0.2320	0.2028	0.5823
XGBoost	0.8258	0.7951	0.7951	0.7951	0.8856
LGBM	0.8347	0.8051	0.8056	0.8053	0.9008
AdaBoost	0.5188	0.4598	0.4289	0.3635	0.7113
MLP	0.4240	0.4348	0.4312	0.4027	0.6544

When the results presented in Table 4.1 are examined, it is seen that the DT algorithm demonstrated the highest performance with an accuracy of 85.38% and an AUC of 0.9123, thus classifying the fault types most accurately in the proposed system. The RF and LGBM models are also notable for their similarly high accuracy rates (84.62% and 83.47%) and strong AUC values (0.9077 and 0.9008). This demonstrates that tree-based and ensemble methods can effectively capture complex fault patterns. The XGBoost model also produced a competitive result with an accuracy of 82.58%. In contrast, the Naive-NB, LR, and MLP models struggled to accurately distinguish fault classes due to their low accuracy and sensitivity values. The kNN and SVM algorithms, on the other hand, demonstrated moderate performance.

Overall, the findings demonstrate that the proposed system performs with high accuracy in fault detection and classification tasks and can be reliably used in real-time applications. In particular, DT and RF models stand out in terms of their ability to successfully distinguish possible fault types of the system. The confusion matrix of these models is given in Figure 4.1.

When Figure 4.1. is examined, it is seen that both models exhibit high performance in fault detection and classification tasks when the confusion matrices of the RF (a) and DT (b) models are examined. These models represent the categories Class 1 (LG fault), Class 2 (LL fault), Class 3 (LLG fault), and Class 5 (LLLG fault). Both models classified the LG, LL, LLG, and LLLG fault types with high accuracy. The RF model demonstrated strong generalization ability by making nearly error-free predictions in these classes. Similarly, the DT model achieved a high level of success with 220, 201, 218, and 473 correct predictions.

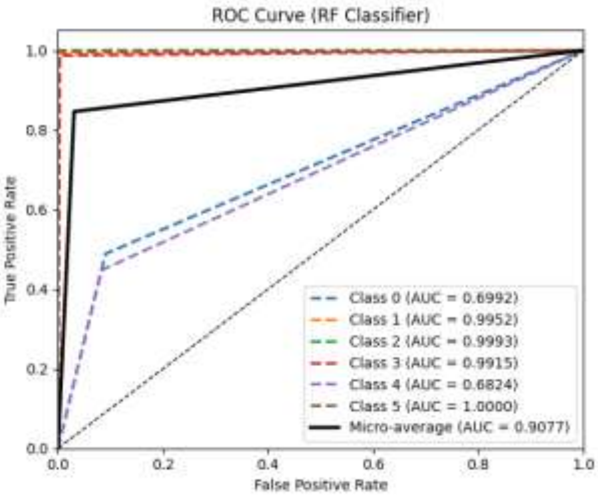


(a)RF

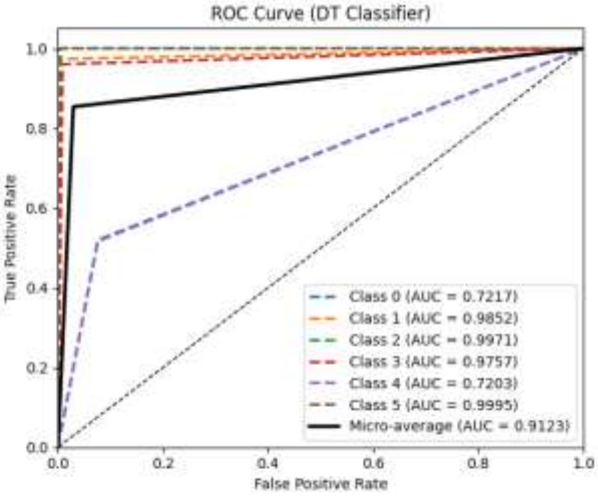
(b)DT

Figure 4.1. Confusion matrix of the models with the highest accuracy

However, certain confusions were observed between the classes "no fault" (Class 0) and "three-phase and ground fault" (Class 5, LLLG). In the RF model, 112 fault-free samples were classified as LLLG, and 121 LLLG samples were classified as fault-free. Similarly, in the DT model, mutual confusion was observed between 99 and 101 samples. This is due to the similarity of the signal characteristics of these two classes. Overall, both the RF and DT models can detect complex fault scenarios with high accuracy and produce reliable results, particularly in the diagnosis of multi-phase short circuits. Furthermore, the use of feature selection or data balancing techniques in the future to eliminate the limited confusion arising from signal similarities between classes could further increase the overall accuracy of the system. The ROC curves for these models are presented in Figure 4.2.



(a)RF



(b)DT

Figure 4.2. The ROC curves of the models that achieved the highest accuracy

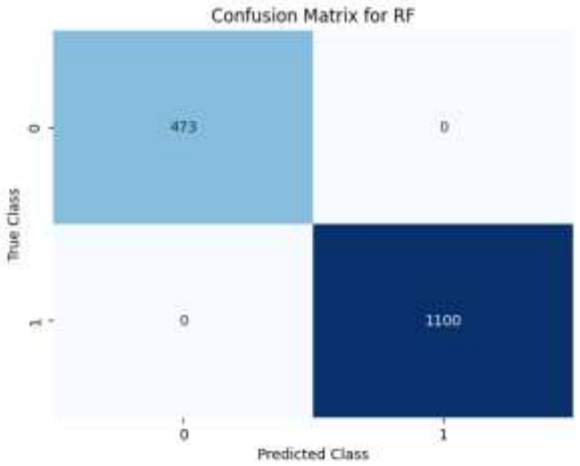
When Figure 4.2 is analyzed, it is seen that both RF (a) and DT (b) models can distinguish fault classes with high accuracy. The RF model demonstrated superior discrimination power, with AUC values exceeding 0.99, particularly for Classes 1, 2, 3, and 5, with an overall performance of micro-average AUC = 0.9077. The DT model also demonstrated similarly high performance, producing AUC values in the range of 0.97–0.99 for Classes 1, 2, 3, and 5, with an overall AUC of 0.9123. Both models experienced some discrimination difficulties for the "no fault" (Class 0) and "three-phase fault" (Class 4) classes. Despite this, the overall performance of both models is quite high, demonstrating their reliability and applicability in terms of fault detection and classification. In addition to classifying fault types, additional experimental analyses were also conducted to evaluate the system's fault detection performance. The findings regarding fault detection are presented in Table 4.2.

Table 4.2. Experimental Findings for Fault Detection

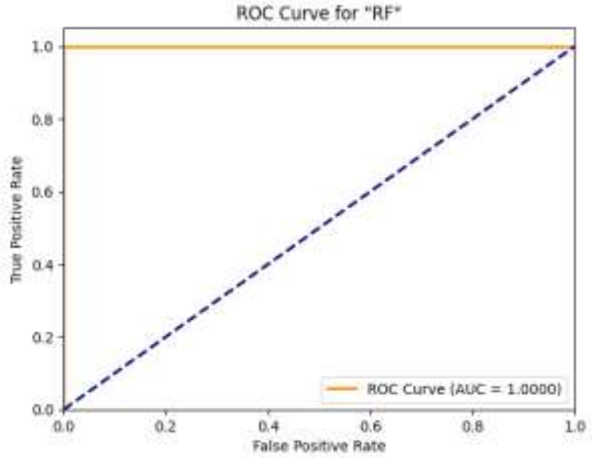
Algorithms	Performance evaluation metrics				
	Accuracy	Precision	Recall	F-Score	AUC (micro)
DT	0.9987	0.9979	0.9991	0.9985	0.9991
kNN	0.9612	0.9546	0.9530	0.9538	0.9530
NB	0.7177	0.6585	0.6445	0.6498	0.6445
SVM	0.8436	0.8121	0.8255	0.8181	0.8255
RF	1.0000	1.0000	1.0000	1.0000	1.0000
LR	0.6033	0.6150	0.6362	0.5914	0.6362
XGBoost	1.0000	1.0000	1.0000	1.0000	1.0000
LGBM	0.9994	0.9995	0.9989	0.9992	0.9989
AdaBoost	0.8086	0.8926	0.6818	0.7065	0.6818
MLP	0.7355	0.6809	0.6579	0.6657	0.6579

The results presented in Table 4.2 demonstrate that the proposed system has a very high level of fault detection performance. Specifically, the Random Forest (RF) and XGBoost algorithms achieved error-free classification by reaching a value of 1.0000 in all performance metrics (Accuracy, Precision, Recall, F-Score, and AUC). The LightGBM (LGBM) and Decision Tree (DT) models also achieved similarly high performance with accuracies of 99.94% and 99.87%, respectively. In contrast, the accuracy values of the Naive Bayes (NB), Logistic Regression (LR), and Multi-Layer Perceptron (MLP) algorithms were relatively low, indicating that these models were unable to adequately capture the complex patterns in the dataset. Overall, the findings reveal that ensemble-based methods (RF, XGBoost, LGBM) are the most reliable, stable, and highly accurate approaches to fault detection. The confusion matrix and ROC curve of the high-performing RF model are given in Figure 4.3.

The results presented in Figure 4.3 demonstrate that the RF model demonstrates exceptional classification accuracy in the fault detection task. An examination of the confusion matrix (a) reveals that the model predicts both fault-free cases (Class 0) and faulty cases (Class 1) with full accuracy. This finding demonstrates that the RF algorithm is able to robustly learn patterns in the dataset and perfectly distinguish between the two classes. The ROC curve also supports this conclusion, with an AUC of 1.0000 demonstrating that the model accurately distinguishes between positive and negative classes. These results demonstrate that the RF model has extremely high discriminatory power in the fault detection problem.



(a)Confusion Matrix



(b)ROC Curve

Figure 4.3. Confusion matrix (a) and ROC curve (b) of the RF model that achieved the highest accuracy

Despite the high accuracy rates achieved in this study, the research has several limitations. First, due to the limited size and diversity of the dataset used, the generalizability of the model to different conditions may be limited. Furthermore, limited confusion occurs between situations with similar characteristics, such as "no fault" and "three-phase + ground fault." These confusions can be mitigated in the future through data preprocessing or feature selection techniques to improve system performance.

Results

This paper investigates the applicability of machine learning-based approaches for the detection and classification of faults in a three-phase transmission line system. While traditional diagnostic methods remain laborious and slow, the investigated machine learning approaches have proven superior in meeting high speed and accuracy requirements. In the developed method, three-phase voltages and three-phase currents are used as inputs to neural networks. Tree-based and ensemble methods such as DT, RF, XGBoost, and LGBM have been found to outperform other algorithms in fault classification. In particular, the RF and XGBoost models exhibited the highest performance in terms of accuracy and AUC values. These results demonstrate that the proposed system offers high accuracy, stability, and reliability in fault detection and classification problems. Furthermore, testing the model with data sets obtained from different systems constitutes an important research area for future studies. Moreover, future research needs to focus on real-world data to overcome the problem of inadequacy of synthetic datasets and develop more comprehensive solutions to the problem of fault localization, which is as vital as fault detection and classification.

References

[1] Goni, M. O. F., Nahiduzzaman, M., Anower, M. S., Rahman, M. M., Islam, M. R., Ahsan, M., ... & Shahjalal, M. (2023). Fast and accurate fault detection and classification in transmission lines using extreme learning machine. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 3, 100107.

[2] Jamil, M., Sharma, S. K., & Singh, R. (2015). Fault detection and classification in electrical power transmission system using artificial neural network. SpringerPlus, 4(1), 334.

[3] Shakiba, F. M., Azizi, S. M., Zhou, M., & Abusorrah, A. (2023). Application of machine learning methods in fault detection and classification of power transmission lines: a survey. Artificial Intelligence Review, 56(7), 5799-5836.

[4] S.A. Aleem, N. Shahid, I.H. Naqvi, Methodologies in power systems fault detection and diagnosis, Energy Syst 6 (1) (2015) 85–108.

[5] Chen, K., Hu, J., Zhang, Y., Yu, Z., & He, J. (2019). Fault location in power distribution systems via deep graph convolutional networks. IEEE Journal on Selected Areas in Communications, 38(1), 119-131.

[6] Saber, A., Emam, A., & Elghazaly, H. (2017). A backup protection technique for three-terminal multisection compound transmission lines. IEEE Transactions on Smart Grid, 9(6), 5653-5663.

[7] Tong, H., Qiu, R. C., Zhang, D., Yang, H., Ding, Q., & Shi, X. (2021). Detection and classification of transmission line transient faults based on graph convolutional neural network. CSEE Journal of Power and Energy Systems, 7(3), 456-471.

[8] Vaish, R., Dwivedi, U. D., Tewari, S., & Tripathi, S. M. (2021). Machine learning applications in power system fault diagnosis: Research advancements and perspectives. Engineering Applications of Artificial

Intelligence, 106, 104504.

[9] Alimi, O. A., Ouahada, K., & Abu-Mahfouz, A. M. (2020). A review of machine learning approaches to power system security and stability. IEEE access, 8, 113512-113531. http://dx.doi.org/10.1109/ACCESS.2020.3003568.

[10] Gururajapathy, S. S., Mokhlis, H., & Illias, H. A. (2017). Fault location and detection techniques in power distribution systems with distributed generation: A review. Renewable and sustainable energy reviews, 74, 949-958.http://dx.doi.org/10.1016/j.rser.2017. 03.021.

[11] Chen, Y. Q., Fink, O., & Sansavini, G. (2017). Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction. IEEE Transactions on Industrial Electronics, 65(1), 561-569.

[12] Guo, M. F., Yang, N. C., & Chen, W. F. (2019). Deep-learning-based fault classification using Hilbert–Huang transform and convolutional neural network in power distribution systems. IEEE Sensors Journal, 19(16), 6905-6913.

[13] Roy, B., Adhikari, S., Datta, S., Devi, K. J., Devi, A. D., Alsaif, F., ... & Ustun, T. S. (2023). Deep learning based relay for online fault detection, classification, and fault location in a grid-connected microgrid. IEEE Access, 11, 62674-62696.

[14] Dasgupta, A., Debnath, S., & Das, A. (2015). Transmission line fault detection and classification using cross-correlation and k-nearest neighbor. International Journal of Knowledge-based and Intelligent Engineering Systems, 19(3), 183-189.

[15] E Sathya Prakash. Electrical fault detection and classification.[Dataset].Kaggle. https://www.kaggle.com/datasets/esathyaprakash/electrical-fault-detection andclassification/data?select=classData.csv

[16] Susanto, E. R., & Cahyana, A. (2025). Penerapan Algoritma XGBoost untuk Prediksi Diabetes: Analisis Confusion Matrix dan ROC Curve. Fountain of Informatics Journal, 10(1), 40-50.