INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH
Volume 2, Issue 1, 2025

MACHINE LEARNING FOR FAULT DETECTION IN THREE-PHASE TRANSMISSION LINES AND ELECTRIC MACHINES

Kifayat Ullah

Department of Electrical Engineering, University of Engineering and Technology, Peshawar, KPK, Pakistan

kifayatullah4@gmail.com

Keywords

Article History
Received: 01 January, 2025
Accepted: 21 February, 2025
Published: 31 March, 2025

Corresponding Author: *
Kifayat Ullah

Abstract
This research explores machine learning algorithms for fault detection and classification in electrical machines and three-phase transmission lines. Using MATLAB Simulink, fault scenarios were simulated to generate datasets, which were preprocessed and divided into training, validation, and testing sets. Algorithms including Decision Trees, XGBoost, k-Nearest Neighbors (KNN), and Random Forest were evaluated. The best-performing model was integrated with Simulink for real-time fault detection. Results indicate that KNN and Random Forest outperform other methods in accurately identifying and classifying faults, enhancing system reliability and reducing downtime.

INTRODUCTION

Transmission lines perform the most important task of transferring electric power from generating stations to load centers. Since the development of the transmission system, power system engineers have been the primary focus for classifying and detecting faults. Transmission line relaying involves three major tasks: detection, classification, and location of the fault [1]. Fault classification must be done as fast and accurate as possible in short duration, it provides a good service for protecting the apparatus as well as an open way for disconnecting the part where this incident happened at fault, and with the help of this, it gives safe way to the system from any damages and ensures that the service to customers is reliable [2]. Additionally, an electrical machine is an electrically driven device that can convert electrical energy into mechanical energy or vice versa. These devices have a wide range of applications, both household machines and industrial equipment. Dependence on electrical machines is increasing; therefore, detecting faults in electrical machines is crucial. Faults in electrical machines come in various types, and they must be detected to ensure the optimal performance, reliability, and longevity of the machines. At a broad level, failure in an electrical machine can be classified as either electrical or mechanical, depending on the root cause of the failure [3]. Fault detection in electrical machines requires a strong Understanding of the domain. Electrical faults in machines can mostly occur in the stator winding due to changes or fluctuations in voltage and current [4]. The reliability of a transmission line is one of the important factors, as a fault in the transmission line leads to a loss of service to its customers. Faults in transmission lines are broadly classified into two categories, symmetric and asymmetric faults, based on the fault condition experienced by each line. Different faults come under the umbrella of each branch. Symmetrical faults are those faults in which all three lines of the three-phase transmission experience the same fault condition during the occurrence of the fault and are out of phase by 120 degrees from one another. A three-phase-to-ground and a three-phase fault fall under this branch. Three-phase to ground faults are considered one of the most severe faults in the transmission line. Solving symmetrical faults is straightforward, accomplished using per-unit calculations. On the other hand, asymmetric faults are those in which at least one of the three lines of a three-phase transmission system experiences a different fault condition [5].

Single line to ground, double 2 line to ground, and double line fault fall under this branch. Solving an asymmetrical fault is complex and challenging. Asymmetric faults are solved using the method of symmetrical components, which was proposed by C.L. Fortescue [6].

Electric motor faults can take many different forms, but one typical problem is overvoltage or undervoltage. Electrical malfunctions and machine damage can occur when an electrical machine is operated at voltage levels exceeding its rated capacity. Over-voltage can cause overheating of electrical components, insulation failure, and accelerated aging. On the other hand, under-voltage can increase current draw, reduce efficiency, and potentially damage the machine. The incorrect connection of the stator windings is another flaw. The performance and dependability of an electrical machine can be severely compromised by incorrect stator winding connections, which can lead to overheating, reduced efficiency, and damage to electrical components. For optimal results, stator windings should be placed and connected according to the manufacturer's guidelines and industry standards [6]. Inter-turn short circuits in windings are another common cause of electrical issues in electric machinery. Direct electrical connections between neighbouring turns in the winding result in these issues, which also include insulation failure, overheating, and excessive currents. Manufacturing mistakes, deteriorating insulation, or mechanical stress can all result in inter-turn short circuits [7]. The primary objective of our proposed work is to design a system utilizing a machine learning algorithm that can accurately identify and categorize faults in this critical electrical system. This project aims to reduce downtime and improve overall system reliability and operation.

LITERATURE

To protect single as well as double circuit transmission lines from different types of faults, such as LG, LL, LLG, and LLL, Yadav and Swetapadma have proposed a wavelet and LDA-based fault detector and classifier, considering the saturation of the current transformer [8]. Ilesh Nandkumar Phalle et al. proposed a method for three-phase transmission line fault detection using IoT. The method they had suggested was a fault-identification system utilizing the internet through a modern combination of components, including the motor, rectifier, regulator, ATMEGA328P microcontroller, honey detector, relay driver IC ULN2003, relay, fault switch, and an IoT module [9]. Xiaoyang Tong and Hao Wen developed a pilot impedance-based protection scheme to detect faults on transmission lines. The faulted line can be accurately detected by using the defined new pilot impedances with synchronized currents and voltages from both line ends [10]. Soufiane Belagoune and colleagues presented one such deep learning architecture based on LSTM for fault detection in transmission lines. In the results, they found that all 31 fault types, including AB faults, ABC faults, and AC faults, were classified with 100% accuracy. Additionally, the other class of faults included ABG, ACG, and AG faults, which were classified with 96.77% accuracy; BCG and BG faults were classified with 93.54% accuracy. However, CG faults were classified with 87.09% accuracy. [11]. Fezan Rafique et al. proposed an end-to-end machine learning model for fault detection and classification in power transmission lines; however, the paper stated that the features needed to be predefined. This paper applies such concepts using an LSTM model to get better results. [12]. In addition, in the research work of Prerana P. Wasnik et al. presented a semi-supervised machine learning approach using k-nearest neighbors for fault detection and classification in transmission lines, achieving an accuracy of 97%.[13]. A hybrid methodology was proposed by Nguyen Nhan Bon and Le Van Dai for the identification, classification, and localization of faults on transmission lines, where various machine learning techniques were combined into one. In its testing, the proposed methodology has shown good accuracy and fast processing time [14]. Shukla and Koley proposed a KNN-based protection scheme for fault detection and classification in a six-phase transmission system [15]. Khalfan Al Kharusi, Abdelsalam El Haffar, and Mostefa Mesbah took one step and presented the application of machine learning to fault detection and classification in their works on transmission lines that operate under inverter-based generator connections. Their results showed high accuracy across different classifiers, with the Bag ensemble classifier achieving 100% accuracy, while the Adaboost ensemble classifier attained an accuracy of 99.4% [16]. Fault Detection in Electrical Machines Javier de las Morenas and colleagues demonstrated the application of machine learning on fault diagnosis of edge electrical machines. Machine learning techniques are applied to classify faults in edge fault identification [17]. Muhammad Faraz Tariq et al. proposed a data-driven approach for robust fault detection and isolation of three-phase induction motors using a system identification model. [18]. P. Tian et al. proposed one of the ways of detecting field winding faults for synchronous motors using the analysis of transient stray fluxes and currents. Further, the mutual application of stray flux signals and currents between the time of startup and the time of trouble to detect faults. [19]. Emanuele Principi et al. introduced the unsupervised data-driven approach for detecting faults in electrical machines. Vibration signals were detected in electric motors in order to identify faults, first compressing attributes through the deep autoencoder network, and then the classifier distinguishing the faults [20]. In this light, Konstantinos N. Gyftakis proposed a detection technique that is more advanced for rotor electrical faults in induction motors at startup based on the analysis of zero-sequence current while the motor was operating in a transient mode [21]. The technique suggested by Ana L. Martinez-Herrera and his coworkers was in terms of analyzing the startup transient current signal, through current signal homogeneity and kurtosis analysis. These features were fed to a feed-forward backpropagation artificial neural network for training it in the task of fault classification [22]. A solution to condition monitoring for small induction motors has been developed by Sayedabbas Sobhi et al. using machine-learning algorithms. Work on the development of architecture for anomaly detection showed that time series forecasting with the use of a random forest and decision tree machine learning model resulted in an accuracy of 0.96% and 0.956%, respectively [23]. Fault Detection of 3 Phase Induction Motor Using Vibration Analysis, a technique proposed by Vivek Dahifale et al, they aimed to perform vibration analysis for a three-phase induction motor to identify various faults such as under voltage, over voltage, overloading, and bearing faults. The existence of a fault can be predicted by studying the Fast Fourier Transform (FFT) of the vibration signal. By comparing the peak frequency in the FFT of the vibration signal with the FFT of the motor under normal operating conditions at the rated frequency, the faulty condition can be identified accordingly [24].

The objective of this research is to utilize machine learning algorithms that can accurately identify and classify faults in these electrical systems, thereby decreasing downtime and increasing overall system reliability by employing Decision trees, XGBoost, k-NN, and random forest Algorithms. The results show that K-Nearest Neighbours (KNN) and Random Forest algorithms outperformed other tested techniques in terms of fault detection and classification in transmission lines and electrical machines, resulting in increased efficiency, reduced operational costs, and improved system resilience in the energy manufacturing sector.

METHADOLOGY

The Study started at the design stage of modeling a circuit using MATLAB Simulink. This circuit was designed to simulate all possible fault scenarios in transmission lines and electrical machines, as shown in Figures 1 and 2. The aim was to simulate real operational conditions when various types of faults occurred. We generated our dataset by simulating the fault conditions correctly. Various types of data have been collected and then stored in an Excel sheet. The approach employed made sure that the data was in an ordered state for ready access in the further stages of our Study.

DATA PREPROCESSING

Before entering the model training section, the data collected was pre-processed. The steps followed in this pre-processing are very crucial. We first classified the data regarding faults into different categories based on the actual scenario, which in turn told us how many times each type of fault occurred. Hence, it provided valuable information regarding the distribution of fault occurrences. Then, find the missing values using suitable imputation methods. To enhance data quality, we removed irrelevant entries that may occur in the dataset. We used pie charts and bar graphs to visualize the data, as they provide valuable insights into the distribution of the faults and help clarify the characteristics of the dataset. The data visualization for Transmission lines is demonstrated in Figure 3, and for the induction motor in Figure 4.

DATA SPLITTING

After preprocessing, the data were split into three subsets: training, validation, and testing datasets. The data was split as follows: 70% for training, and a substantial fraction of the data was used to train the machine learning model; 20% of the data was reserved for validation. 10% was reserved for testing: the subset used to evaluate the performance of the trained models and subsequently provide an unbiased estimate of the model's accuracy and reliability.

MODEL DEVELOPMENT

A supervised machine learning approach to determine the most effective model for fault detection. At the model development stage, several algorithms were considered to determine the most effective model for fault detection. The considered algorithms are Random Forest, gradient boosting, ridge regression, decision tree, and voting classifier. These algorithms are trained on the preprocessed dataset and then tested for performance. Therefore, we select one particular algorithm that yields the best results in terms of accuracy and reliability in fault detection.

MODEL EVALUATION

The performance of the trained models was determined using confusion matrices. They provided a very detailed view of the accuracy, precision, recall, and F1 score for either model. The confusion matrices helped in showing the strengths and limitations of each model, and these models were then compared with one another, as shown in Figure 5. As a result of this comparison, we selected the best-performing model, which would be effective in detecting faults. The best model was then subjected to further testing and verification to assess its viability and robustness.

INTEGRATION WITH SIMULINK

After evaluating and selecting the best model, the Simulink model was integrated with the trained machine-learning algorithm for the transmission line and electrical machine circuit. The integrated network was interfaced in such a way that real-time fault detection could now be achieved

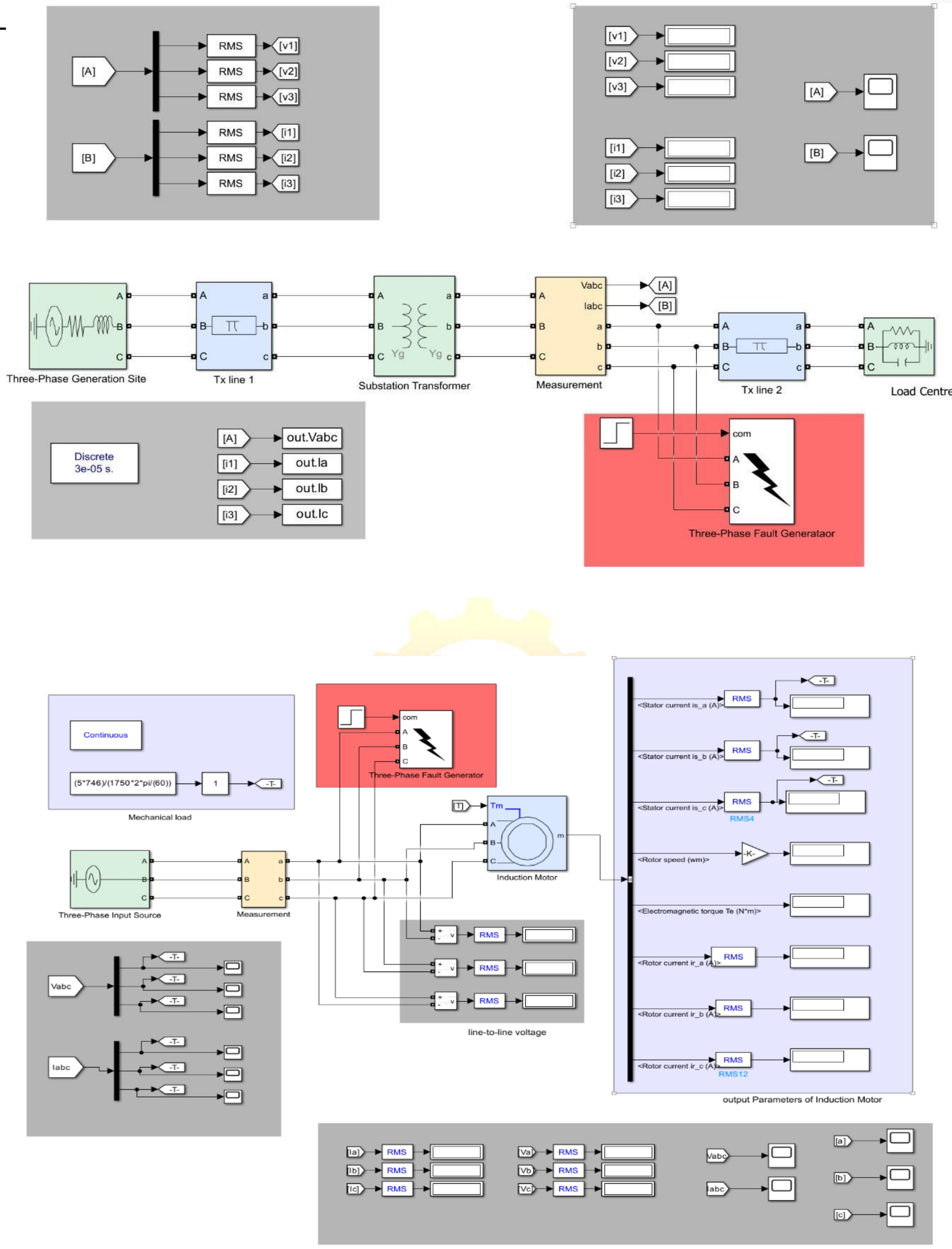
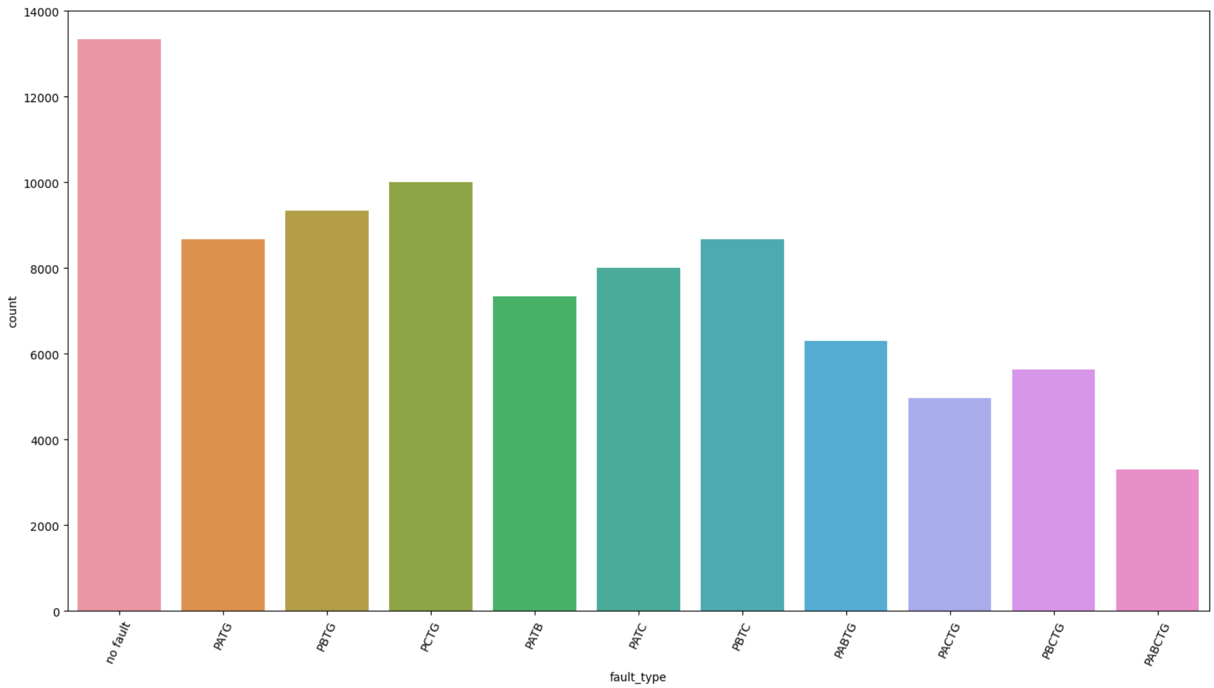


Figure 1 Simulink circuit of Tx line

fault_type



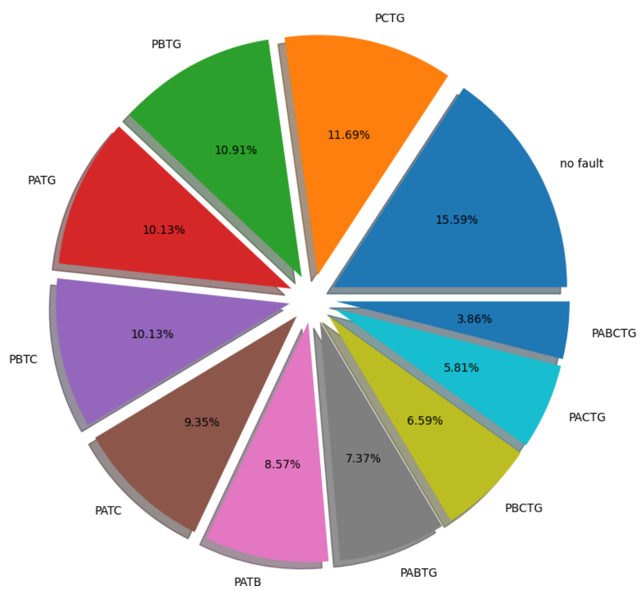
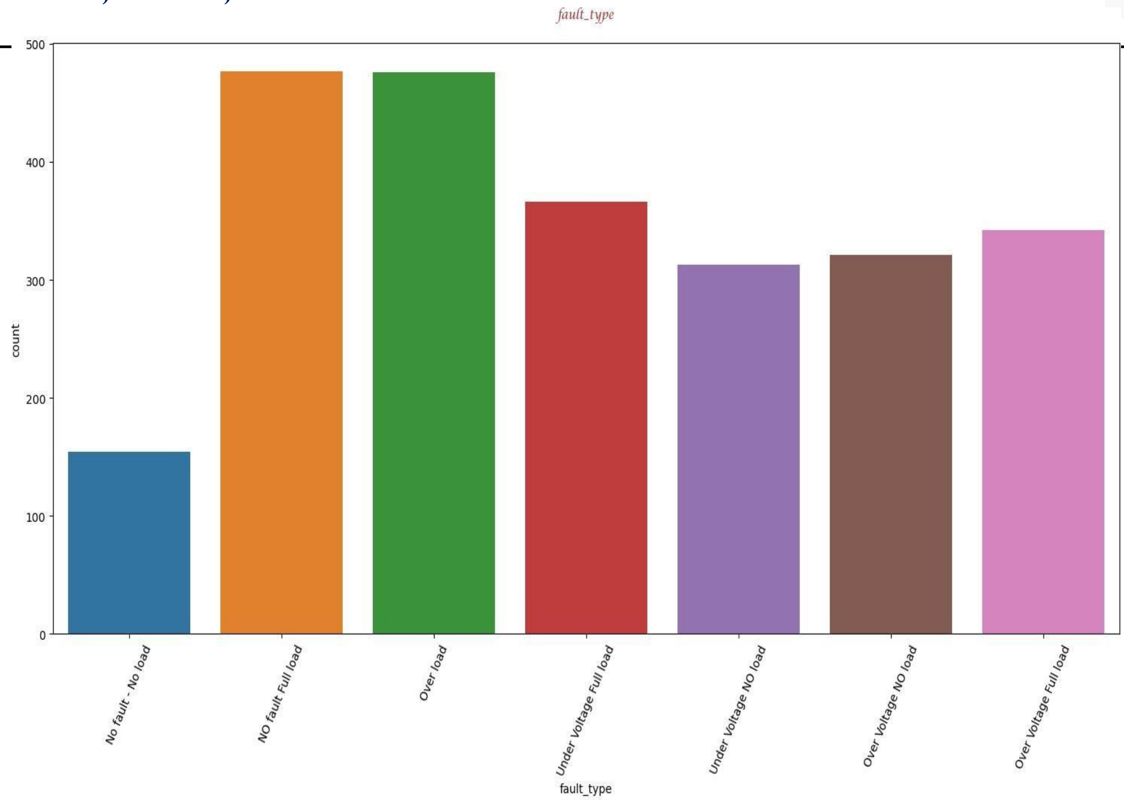


Figure 2 Simulink circuit of Induction Motor



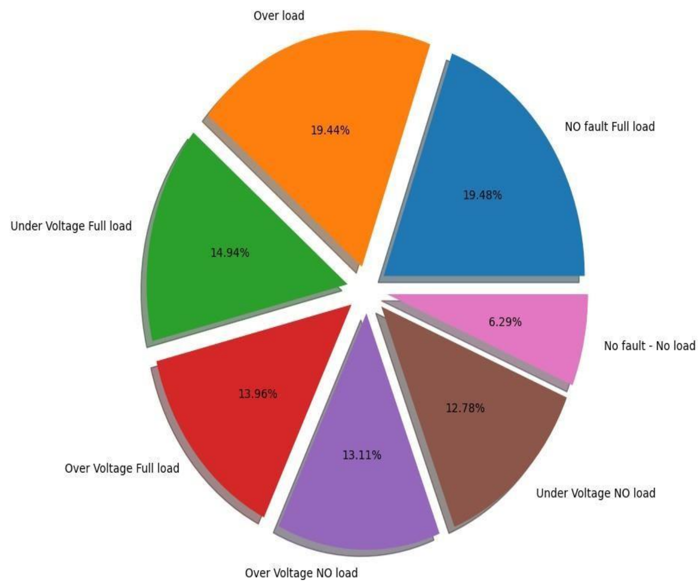


Figure 3 Bar graph and PI chart of Tx line fault data

Figure 4 Bar graph and PI chart of Induction Motor data Faults

within the MATLAB environment. The linkage of the Simulink model with the machine-learning algorithm enables the system for real-time detection and classification of faults. The ability to detect in real-time is what makes this research applicable and realistic for the developed system.

FAULT DETECTION TESTING

Overall, testing was conducted on the integrated model to check for functionality and reliability. The developed test cases have been used to describe a variety of fault cases that occur within the Simulink circuits, and the system's detection and classification of these faults were tested here. Fault detection testing validated that the machine-learning model, trained on the dataset extracted to identify and categorize different fault types, was effective. Results from those tests demonstrate the system's robustness and accuracy in fault detection.

OPTIMIZATION AND FINE-TUNING

The trained machine learning model is used to improve its accuracy based on the feedback from the Simulink. Optimization and fine-tuning are important, as they bring out the best from the model, ensuring that it runs at its best possible level during application in the real world.

RESULT AND DISCUSSION

By employing machine learning algorithms, the following results are obtained:

Fault Detection in Transmission Lines

In our research, we tested several machine learning algorithms to determine the most effective method for fault detection in transmission lines. We employed machine learning algorithms, including Logistic Regression, Ridge Regression, Decision Tree, Random Forest, KNN, Naive Bayes, and XG-Boost. The results obtained for each algorithm are provided in Table 1.

ACCURACY COMPARISON

From the above tabular demonstration, it can be summarized that Decision Tree, Random Forest, KNN, and XG-Boost algorithms performed exceptionally well. However, the Logistic Regression and Ridge Regression algorithms yielded abysmal performance, with accuracies of 64% and 55.5%, respectively. Whereas Naive Bayes also gave very promising accuracy. The Random Forest algorithm and KNN algorithm demonstrate excellent performance in terms of accuracy, indicating that they are effective in detecting and classifying faults in transmission lines.

DISCUSSION ON ALGORITHM PERFORMANCE

High accuracy has been attained from the Random Forest and K-NN algorithms. Hence, these two can be used for real-time operation for fault detection in transmission lines. The Random Forest algorithm, with its ensemble approach, provides excellent generalization capabilities by combining the predictions of multiple decision trees. In contrast to KNN, which is particularly suitable for most classification tasks due to its reliance on the closeness of data points in the feature space. Points to the fact that Logistic Regression and Ridge Regression perform poorly because they are unable to handle complex datasets with non-linear relationships and intricate fault patterns. In most cases, these algorithms fail to accurately represent the underlying dynamics of the condition, leading to inaccuracies. The following is the confusion matrix shown in Figure 5 of the KNN algorithm.

Table 1 Comparison of Different ML Models on Tx Dataset

Algorithms used on Tx dataset	Obtained Accuracy in %
Logistic Regression	64
Ridge Regression	55.5
Decision Tree	99.8
Random Forest	99.97
KNN	99.99
Naive Bayes	98.4
XG-Boost	99.03

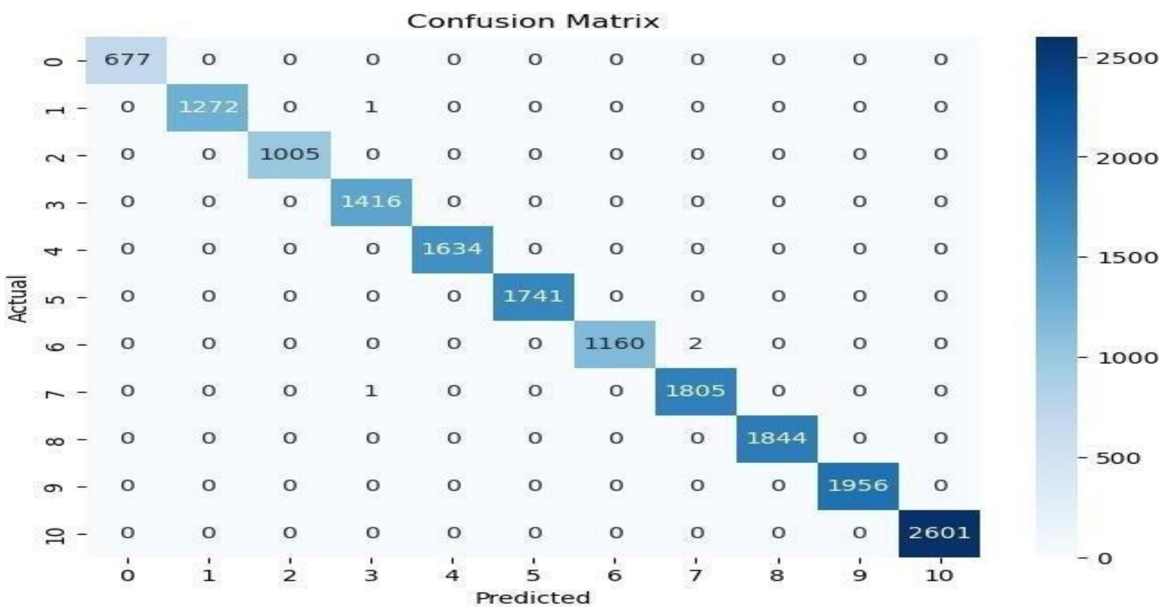


Figure 5 TXL confusion matrices

FAULT DETECTION IN ELECTRICAL MACHINES

Similarly, tests are conducted on machine learning algorithms for fault detection in electrical machines. The accuracy of the machine learning algorithms in fault detection is shown in Table 2.

Table 2 Comparison of Different ML Models on Induction Motor Dataset

Algorithms used on IM dataset	Obtained Accuracy in %
Logistic Regression	52

Ridge Regression	48
Decision Tree	97
Random Forest	99
KNN	98
Naive Bayes	89
XG-Boost	96

DISCUSSION

The Decision Tree, Random Forest, KNN, and XG-Boost algorithms showed an excellent rate of accuracy. For Logistic Regression and Ridge Regression, the results were low. The accuracy of Naive Bayes was moderate. The most suitable model after comparison was the Random Forest algorithm, followed by the KNN. The results in Fault Detection for the machine confirm the findings for the algorithms, which reconfirm the reliability. The reason behind achieving high accuracy with the Random Forest and KNN algorithms in fault detection of electrical machines is that these salgorithms are highly effective in all types of fault detection. Moreover, the Random Forest algorithm's ability to handle high-dimensional data and avoid overfitting makes it suitable for real-time applications. For this reason, the Logistic Regression and Ridge Regression algorithms did not perform well in fault detection of electrical machines, indicating that these algorithms cannot handle the complexity of the fault patterns. Additionally, they may not perform accurately in responding to the non-linearity and noise present in the dataset. Figure 6 represents a confusion matrix of the decision tree classifier.

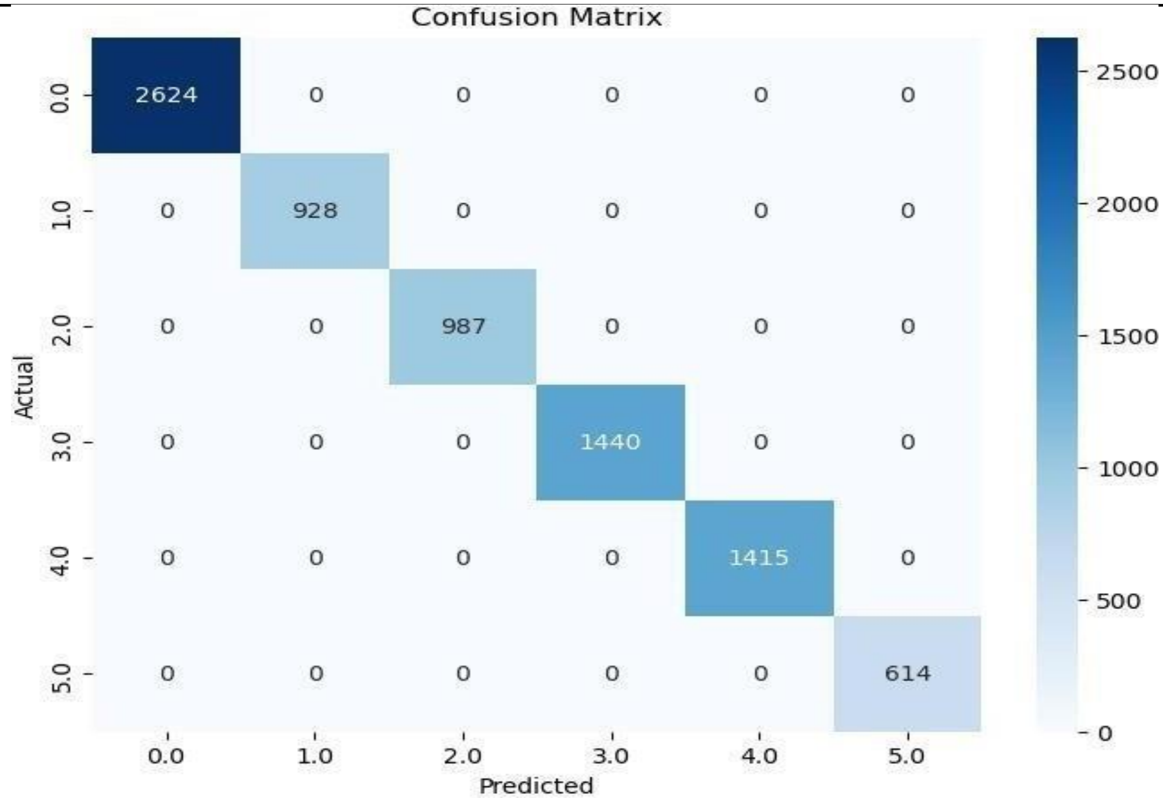


Figure 6 Electrical machine confusion matrices

CONCLUSION

This study indicated the significant capabilities of machine learning in enhancing fault detection and categorization in transmission lines and electrical machines through the simulation of specific fault conditions in the MATLAB Simulink environment. The datasets were generated on which the machine learning model was trained in MATLAB Simulink. The integration between the machine learning algorithm and Simulink was performed to ensure the accuracy and efficiency of the designed network, as well as to enable real-time fault detection applications. Furthermore, it demonstrated that machine learning techniques can effectively detect and classify faults, resulting in reduced downtime and operational costs. This advancement in fault detection technology is not only applicable to the energy manufacturing sector but also holds promise for broader applications within modern infrastructure.

REFERENCES

H. H. Goh et al., “Transmission Line Fault Detection: A Review,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 8, no. 1, p. 199, Oct. 2017, doi: https://doi.org/10.11591/ijeecs.v8.i1.pp199-205.

K. M. Silva, B. A. Souza, and N. S. D. Brito, “Fault Detection and Classification in Transmission Lines Based on Wavelet Transform and ANN,” IEEE Transactions on Power Delivery, vol. 21, no. 4, pp. 2058–2063, Oct. 2006, doi: https://doi.org/10.1109/tpwrd.2006.876659.

D. P. Kothari and I. J. Nagrath, Electric machines. New Delhi Tata Mcgraw-Hill, 2010.

D. Basak, A. Tiwari and S. P. Das, "Fault diagnosis and condition monitoring of electrical machines - A Review," 2006 IEEE International Conference on Industrial Technology, Mumbai, India, 2006, pp. 3061- 3066, doi: 10.1109/ICIT.2006.372719.

J. Grainger and W. Stevenson, Power System Analysis. McGraw-Hill Series in Electric, 1994.

S. Akbar, T. Vaimann, B. Asad, A. Kallaste, M. U. Sardar, and K. Kudelina, "State-of-the-Art Techniques for Fault Diagnosis in Electrical Machines: Advancements and Future Directions," Energies, vol. 16, no. 17, p. 6345, Jan. 2023, doi: https://doi.org/10.3390/en16176345.

Boldea, I., & Nasar, S. A. (2002). "The Induction Machine Handbook." CRC Press. Joksimović, G. M., & Penman, J. (2000). "The Detection of Inter-Turn Short Circuits in the Stator Windings of Operating Motors." IEEE Transactions on Industrial Electronics, 47(5), 1078-1084.

A. Yadav and A. Swetapadma, "A novel transmission line relaying scheme for fault detection and classification using wavelet transform and linear discriminant analysis," Ain Shams Engineering Journal, vol. 6, no. 1, pp. 199–209, Mar. 2015, doi: https://doi.org/10.1016/j.asej.2014.10.005.

"THREE PHASE TRANSMISSION LINE FAULT DETECTION OVER IOT," International Research Journal of Modernization in Engineering Technology and Science, Dec. 2023, doi: https://doi.org/10.56726/irjmets47019.X.

Tong and H. Wen, "A novel transmission line fault detection algorithm based on pilot impedance," Electric Power Systems Research, vol. 179, p. 106062, Feb. 2020. doi: https://doi.org/10.1016/j.epsr.2019.106062.

S. Belagoune, N. Bali, A. Bakdi, B. Baadji, and K. Atif, "Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems," Measurement, vol. 177, p. 109330, Jun. 2021, doi: https://doi.org/10.1016/j.measurement.2021.109330.

F. Rafique, L. Fu, and R. Mai, "End to end machine learning for fault detection and classification in power transmission lines," Electric Power Systems Research, vol. 199, p. 107430, Oct. 2021, doi: https://doi.org/10.1016/j.epsr.2021.107430.

P. P. Wasnik, N. J. Phadkule, and K. D. Thakur, "Fault detection and classification based on semi- supervised machine learning using KNN," Dec. 2019, doi: https://doi.org/10.1109/icitaet47105.2019.9170220.

N. N. Bon and L. V. Dai, "Fault Identification, Classification, and Location on Transmission Lines Using Combined Machine Learning Methods," International Journal of Engineering and Technology Innovation, vol. 12, no. 2, pp. 91–109, Feb. 2022, doi: https://doi.org/10.46604/ijeti.2022.7571.

Sunil Kumar Shukla and E. Koley, "Detection and classification of open conductor faults in six-phase transmission system using k-nearest neighbour algorithm," 2017 7th International Conference on Power Systems (ICPS), Dec. 2017, doi: https://doi.org/10.1109/icpes.2017.8387285.

K. Al Kharusi, A. El Haffar, and M. Mesbah, "Fault Detection and Classification in Transmission Lines Connected to Inverter-Based Generators Using Machine Learning," Energies, vol. 15, no. 15, p. 5475, Jul. 2022, doi: https://doi.org/10.3390/en15155475.

Javier, F. Moya-Fernández, and Julio Alberto López-Gómez, “The Edge Application of Machine Learning Techniques for Fault Diagnosis in Electrical Machines,” Sensors, vol. 23, no. 5, pp. 2649–2649, Feb. 2023, doi: https://doi.org/10.3390/s23052649.

M. F. Tariq, A. Q. Khan, M. Abid, and G. Mustafa, “Data-Driven Robust Fault Detection and Isolation of Three-Phase Induction Motor,” IEEE Transactions on Industrial Electronics, vol. 66, no. 6, pp. 4707– 4715, Jun. 2019, doi: https://doi.org/10.1109/tie.2018.2866104.

P. Tian, J. A. Antonino-Daviu, C. A. Platero, and L. Dunai, “Detection of Field Winding Faults in Synchronous Motors via Analysis of Transient Stray Fluxes and Currents,” IEEE Transactions on Energy Conversion, vol. 36, no. 3, pp. 2330–2338, Sep. 2021, doi: https://doi.org/10.1109/tec.2020.3041643.

E. Principi, D. Rossetti, S. Squartini, and F. Piazza, “Unsupervised electric motor fault detection by using deep autoencoders,” IEEE/CAA Journal of Automatica Sinica, vol. 6, no. 2, pp. 441–451, Mar. 2019, doi: https://doi.org/10.1109/jas.2019.1911393.

K. N. Gyftakis, D. V. Spyropoulos, and E. Mitronikas, “Advanced Detection of Rotor Electrical Faults in Induction Motors at Start-up,” IEEE Transactions on Energy Conversion, pp. 1–1, 2020, doi: https://doi.org/10.1109/tec.2020.3025786.

A. L. Martinez-Herrera, E. R. Ferrucho-Alvarez, L. M. Ledesma-Carrillo, R. I. Mata-Chavez, M. Lopez- Ramirez, and E. Cabal-Yepez, “Multiple Fault Detection in Induction Motors through Homogeneity and Kurtosis Computation,” Energies, vol. 15, no. 4, p. 1541, Jan. 2022, doi: https://doi.org/10.3390/en15041541.

S. Sobhi, M. Reshadi, N. Zarft, A. Terheide, and S. Dick, “Condition Monitoring and Fault Detection in Small Induction Motors Using Machine Learning Algorithms,” Information, vol. 14, no. 6, p. 329, Jun. 2023, doi: https://doi.org/10.3390/info14060329.

V. Dahifale, M. Keskar, V. Kakyalia, S. Rodge, and G. M. Karve, “Fault Detection Of 3 Phase Induction Motor Using Vibration Analysis,” IJFMR - International Journal For Multidisciplinary Research, vol. 5, no. 5, doi: https://doi.org/10.36948/ijfmr.2023.v05i05.8257.