IJITCE

ISSN 2347- 3657

International Journal of
Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Smart Detection And Classification of Electrical Faults Using Machine Learning

[1]K. Purusotham, [2]Neeluru Kavya, [3]Ganne Venuteja,[4]Narala Vandana,[5]Mangala Murali Mohan,

[1]Assistant Professor, Department of EEE, Ananthalakshmi Institute of Technology and Sciences, Itikalapalli, Near Sk University, Ananthapur.
[2,3,4,5]Student, Department of EEE, Ananthalakshmi Institute of Technology and Sciences, Itikalapalli, Near Sk University, Ananthapur.

Abstract—

The application of machine learning methods and techniques for the identification of power transmission line defects is the main topic of this research. In contrast to the fast growth in power consumption in recent decades, transmission capacity has not kept pace with this expansion. The most frequent transmission line issues and how to classify them using machine learning are covered in this study. An accurate result is produced by analyzing the flaws with various combinations of inputs using the given methodologies. Spyder IDE, which stands for Scientific Python Development Environment, is where the machine learning techniques are carried out. This strategy is designed to tackle the target. Machine learning, decision tree models, LSTM (long short-term memory), KNN (k-nearest neighbor), and SVM (support vector machine) are some of the terms used in this context.

INTRODUCTION

The electrical grid has become more important to our daily lives in recent years. In a state of equilibrium, they function. Fault analysis is the main problem with the power system. It is possible to restore normal flow by detecting and controlling a defect in the transmission line. Various natural disturbances, such as lightning strikes, earthquakes, and short circuits, may cause faults. The use of fault categorization may help fix this issue. It is possible to categorize several kinds of errors using machine learning. Only when the system is imbalanced can all these fault kinds be categorized. They are either symmetrical or they are not. Unsymmetrical faults, such as the L-G Fault (which goes from ground to line) or the L-L Fault (which goes from line to line), are common on transmission lines. Below, we'll go over the new approach and all the ways it excels above the old and conventional ways. We have used four algorithms in their entirety to get the greatest and most accurate results, and we only utilize the best of these results as our result.

II. RELATED WORKS

In their 2020 study, Zuraida Muhammad and Shabinar Abd Hamid used a ml algorithm approach to enhance power quality and address transmission issues with the use of an Artificial Neural Network (ANN). By studying fault detection and fault classification, power quality may be enhanced. The use of an impedance approach is the first step in producing the malfunction. The input is the measurement of the defective current and voltage after stimulation. The fault categorization and detection system is built using a feed forward network and back propagation techniques. Mean squared error (MSC) was used to quantify the performance of both the detection and classification tasks. The deduction primarily gives 100% accuracy and achieves a tolerance of 5.6 148 of MSC. The suggested technique was able to obtain a fault classification accuracy of 70% and an MSC tolerance of 0.893955. In 2020, Zakaria Hussian used a deep learning technique to identify the power system issue. A novel tool called a Long Short-Term Memory (LSTM) is introduced in this technique to analyze power system issues and identify them. It is possible to detect a fault current signal in the transmission line. The Lstm (long short-term memory) network, which is mostly used for fault categorization, receives this signal as input. To bolster the suggested model, a gaussian noise level between 20 and 30 dB SNR (signal to noise ratio) is included. Variegated results emerge from the simulation. Values between zero and ten are what we get. It is easy to categorize faults using the values. The system is deemed non-faulty if the obtained value is zero. Depending on the values, ten distinct sorts of defects may be analyzed if it varies from 1 to 10. The LSTM approach outperforms other methods in terms of how easily it produces results. When it comes to transmission line identification and categorization, our proposed technique is foolproof. Using the Internet of Things (IoT), Monica (2019) suggested a novel way to identify and manage

ISSN 2347–3657

Volume 13, Issue 2, 2025

transmission line faults. The suggested approach detects the transmission line voltage using a voltage sensor that may be linked to the microcontroller. It is easy to detect a problem if the voltage flow limit deviates from the typical level. A relay is used to trip the circuit if the limit surpasses the voltage flow. To pinpoint the exact position of a malfunction, the controller may communicate with a Global Positioning System (GPS). Due to the high temperatures, the transmission line, which may be retained at a distance of less than 100 cm, can be detected by the flame sensor if a fire or flame were to break out. The data is sent to the cloud using the esp8266. A liquid crystal display (LCD) shows where the defect was and how it was detected.

III.EXISTING METHOD

Fault detection and categorization are the two main components of the project. In MATLAB (MATrix LABoratory), the model is run via a simulation. Then, we use the MATLAB-provided classification learner program to do the training and classification. This allowed for the rapid training and validation of all the models. Approximately twenty-four Machine Learning models were also trained on the dataset. This approach takes into account a variety of input combinations while training the model. The output computations take into account a mix of these more accurate models. On the whole, four models are taken into account. Efficient end-to-end fault detection is another popular approach. This approach uses the current and voltage values measured at the moment the problem occurred. After that, the model is trained using these variables. Two categories of targets are distinguished. The goals of the detection model are judgments involving fault or no fault. Classification models aim for No-fault, Line-to-Ground (LG), Line-to-Line (LL), Line-to-Line-to-Ground (LLG), and Line-to-Line-to-Line (LLL) faults.

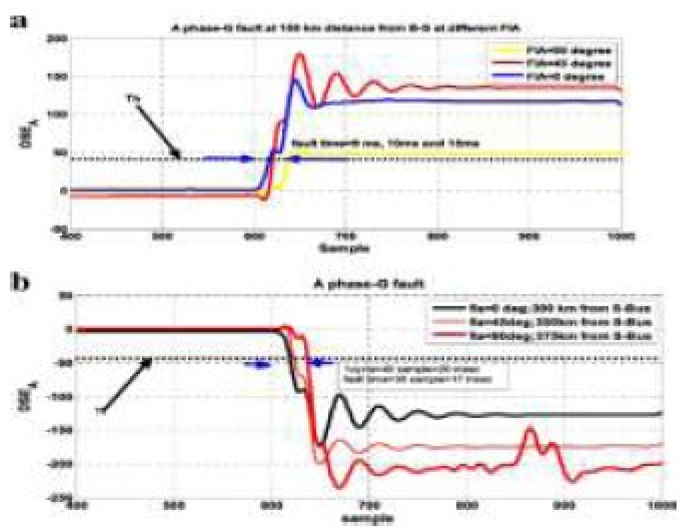


Figure 3.1:fault samples

As an additional tool, the terminal approach may identify transmission line problems. Figure 3.1 displays the picture of the defective samples. One way to separate a defect from its source is to use a resistor to measure the voltage and current levels at the moment the fault occurs. Without tracing, this approach may locate the defect site from the beginning to the conclusion.

IV.PROPOSED SYSTEM

A methodical approach is taken to the technique that has been implemented for the purpose of electrical fault categorization and detection. Dataset input reading is the first stage. After that, the Spyder IDE is used to do the preprocessing. Training the dataset follows. The next step is to sort the training dataset into models that are defective and those that are not. After then, the flawed models are broken down into more specific issues and put through experimental testing. For this procedure, many algorithms are used. Figure 4.1 shows the overall method that is used to classify and identify electric faults in electric lines.

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

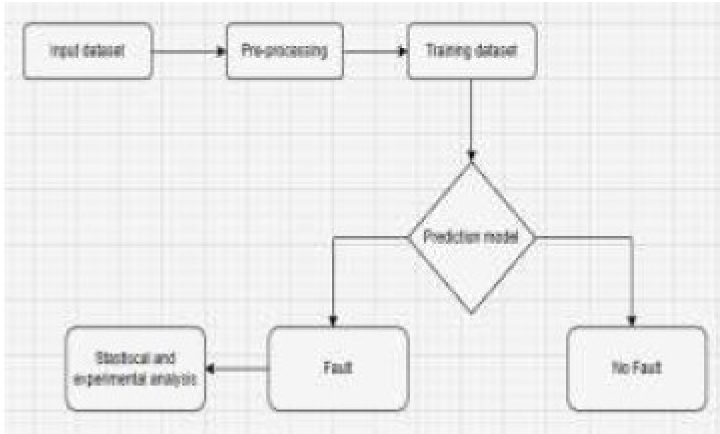


Figure 4.1:Proposed workflow

In an ideal world, a power system would be perfectly balanced. When there is an imbalance in the system, it signifies the problem. Problems like insulation failures may develop in transmission lines as a result of either natural disasters or careless maintenance. The following are some of the various types of transmission line faults:

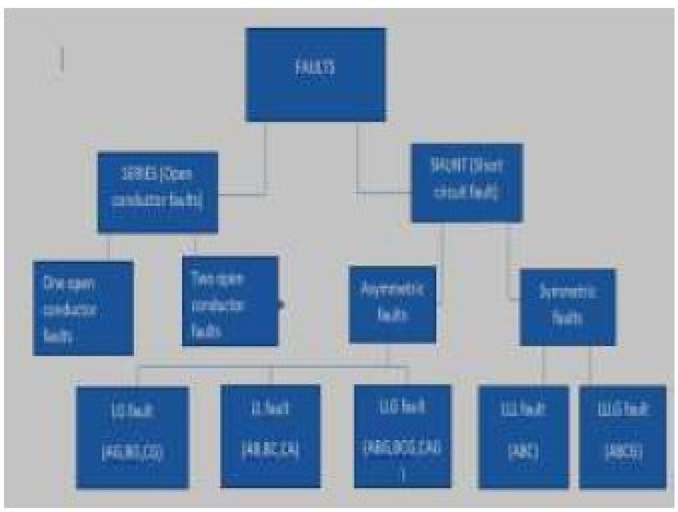


Figure 4.2:Classification of faults

Figure 4.2 displays the categorization of electrical problems. Python and the Spyder IDE are used in this system's suggested fault detection and categorization process. Initially, the short-circuit defect is identified and then categorized according to the current and voltage values measured inside the fault. The faults are classified by comparing the line voltage and line current characteristics with the preloaded datasets. You may think of the dataset as the data used to train the model to identify the fault type. Va, Vb, and Vc are the lines in this dataset, while Ia, Ib, and Ic are the three currents in a three-phase line system.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Inputs - [Ia, Ib, Ic, Va, Vb, Vc]
Outputs - [G C B A]

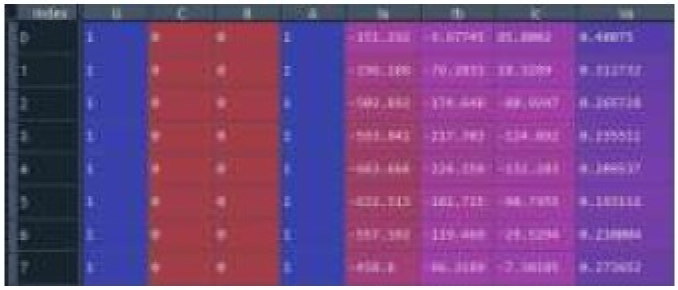
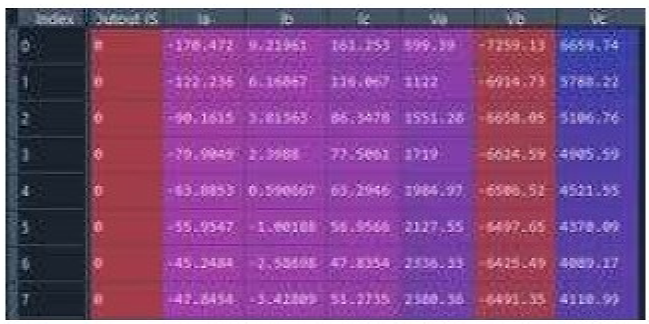


Figure 4.3:Classification:dataset

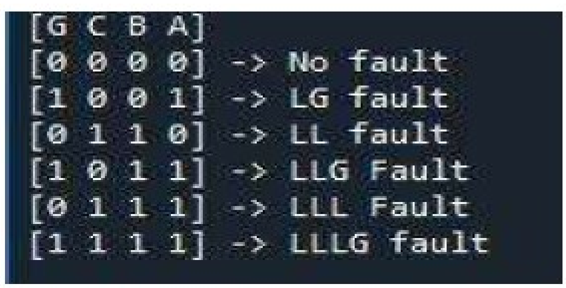
Figure 4.3 displays a dataset that is used for shunt fault classification. A Database for Classification Both the current and the voltage per unit typically range from -100A to 100A under typical circumstances. As the fault progresses, we see irrational and unpredictable conduct, and the line current value may go as low as -800A. Figure 4.3 displays the categorization dataset. You may get the dataset by clicking on the provided link.



Index	Output (S	Ia	Ib	Ic	Va	Vb	Vc
0	0	-170.472	9.21961	161.253	599.39	-7259.13	6659.74
1	0	-122.236	6.16867	116.067	1122	-6914.73	5788.22
2	0	-90.1615	3.81563	86.3478	1551.28	-6658.05	5106.76
3	0	-79.9049	2.3988	77.5061	1719	-6624.59	4905.59
4	0	-63.8853	0.590667	63.2946	1984.97	-6506.52	4521.55
5	0	-55.9547	-1.00188	56.9566	2127.55	-6497.65	4370.09
6	0	-45.2484	-2.58698	47.8354	2336.33	-6425.49	4089.17
7	0	-47.8454	-3.42809	51.2735	2380.36	-6491.35	4110.99

Figure 4.4:Detect Dataset

The picture of the detection dataset is shown in Figure 4.4. You can get the dataset at this URL:kaggle.com/input/electrical-fault-detection and classification/detectdataset.csv.



```
[G C B A]
[0 0 0 0] -> No fault
[1 0 0 1] -> LG fault
[0 1 1 0] -> LL fault
[1 0 1 1] -> LLG Fault
[0 1 1 1] -> LLL Fault
[1 1 1 1] -> LLLG fault
```

Figure 4.5:Output sample

Figure 4.5 shows the different kinds of defects. The classification algorithm. One popular and sophisticated classifier is SVM, which stands for support vector machine. Both linear and non-linear data types are acceptable in support vector machine classification. Thanks to its superior signal detection capabilities compared to other methods, the SVM is still going strong today. Following this, we take the SVM model:

V.EXPERIMENTAL RESULTS:

International Journal of
Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

Identifying and categorizing transmission line defects is the primary goal of the project. What follows is an analysis of their findings. Some machine learning techniques made use of the datasets of two detection and classification models. Additionally, the method that produced the most precise outcome is shown below. Presented below is a bar graph illustrating the quantity of values found in both datasets. There is exactly one value in each dataset.

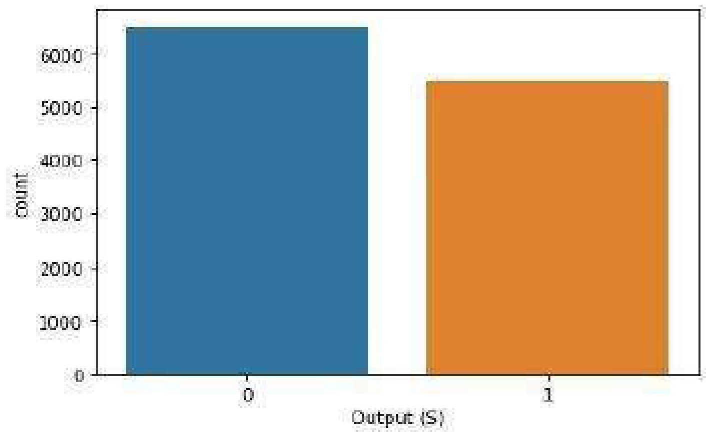


Figure 4.6:Classification and detection dataset range

The values of both datasets are shown in Figure 4.6. The following is a visualization of the present in the detection dataset.

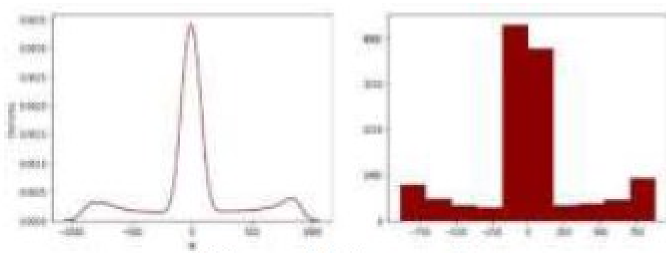


Figure 4.7:Current Ia

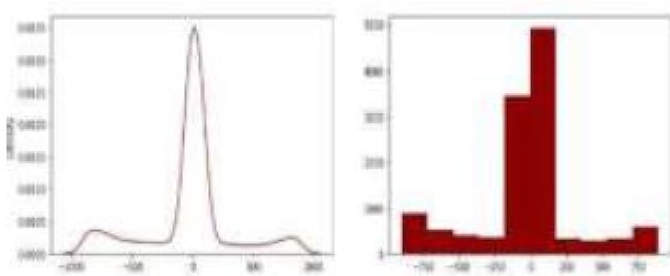


Figure 4.8:Current Ib

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

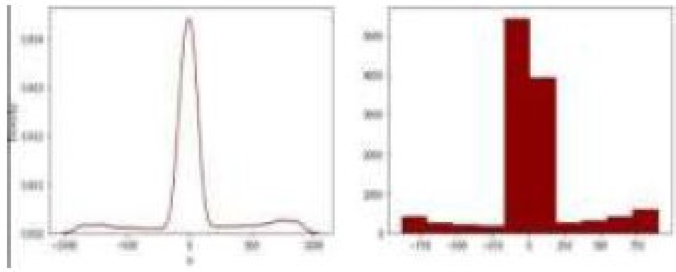


Figure 4.9:Current Ic

As shown in Figures 4.7, 4.8, and 4.9, the three lines containing values from the classification dataset are now graphed. The x-axis represents the current value, while the y-axis represents density. Additionally, the dataset's values are used to construct a voltage graph, which displays the values of the voltages inside the dataset.

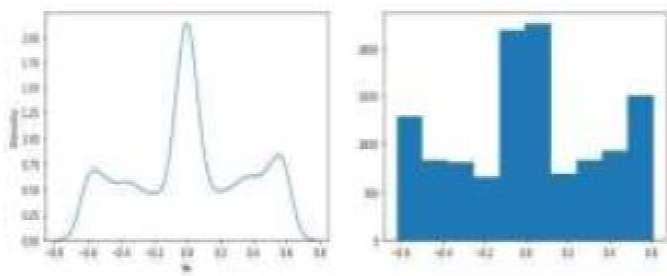


Figure 4.10:Voltage Va

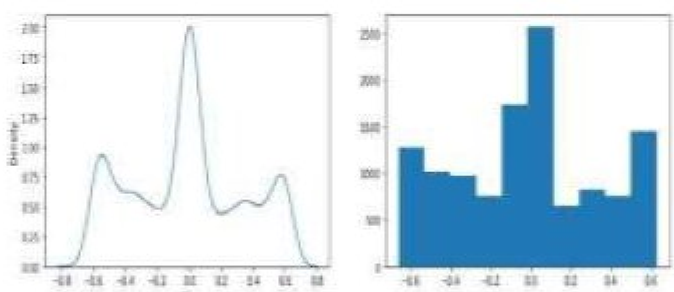


Figure 4.11:Voltage Vb

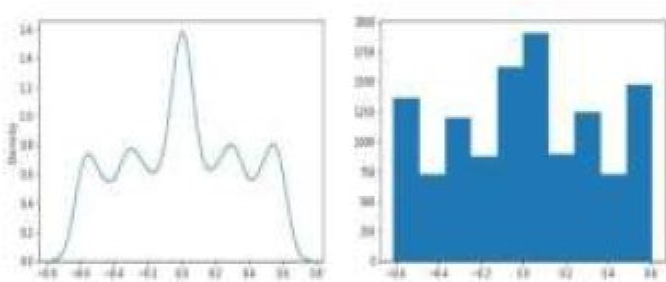


Figure 4.12:Voltage Vc

Using values from the classification dataset, Figures 4.10, 4.11, and 4.12 show the voltage graph of the three lines. The x-axis represents voltage value and the y-axis density. Following are the line mediums and the three lines (A, B, and C) that represent them on a graph, with the current and voltage values serving as the basis. Normally dispersed data is all there is. We can see the signal flow graph in relation to lines A, B, and C.

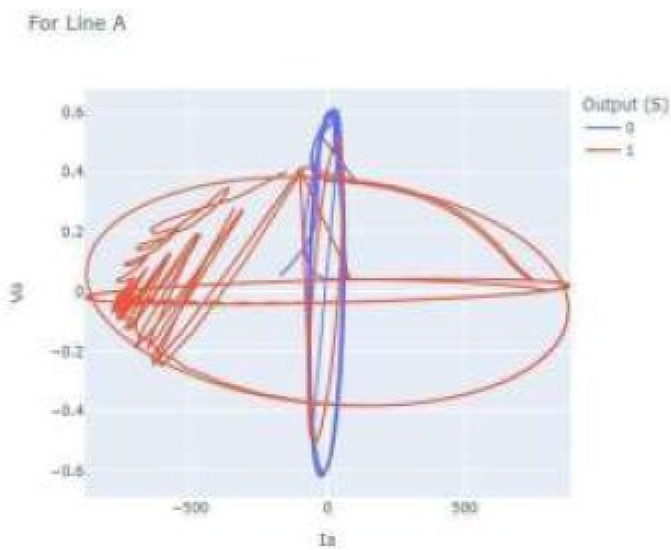


Figure 4.13: Signal flow graph of line A

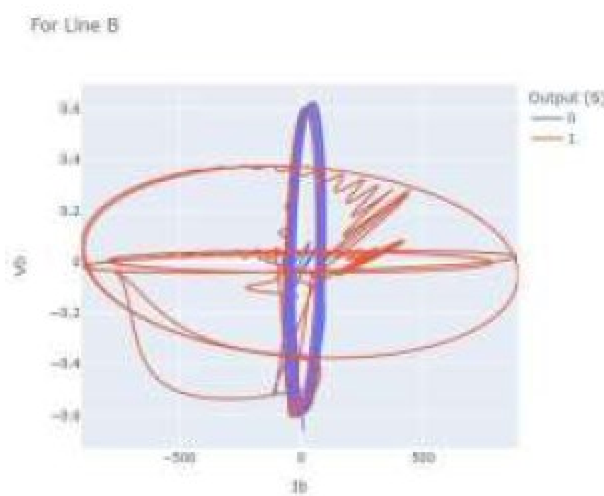


Figure 4.14: Signal flow graph of Line B

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

For Line C

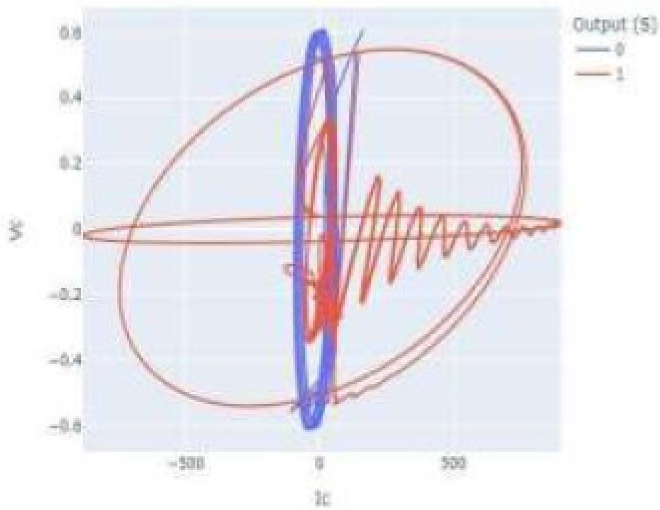


Figure 4.15: Signal flow graph of Line C

A current graph would have the current as the x-axis and a voltage graph as the y-axis. Line A, Line B, and Line C's signal flow graphs are shown in Figures 4.13, 4.14, and 4.15. Line currents typically range from -100 to 100 amps and voltages from -0.6 V to 0.6 V. A pattern of irrational and seemingly random behavior is seen during fault. There is a point where the line current falls on the -/+800 ampere range. The faults are classified using the classification dataset according to the values of [G C B A], where the ground and line values are considered. The following are the defect types:

[G C B A]

[0 0 0 0] ->No fault

[1 0 0 1] ->LG fault

[0 1 1 0] ->LL fault

[1 0 1 1] ->LLG fault

[0 1 1 1] ->LLL fault

[1 1 1 1] ->LLLG fault

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

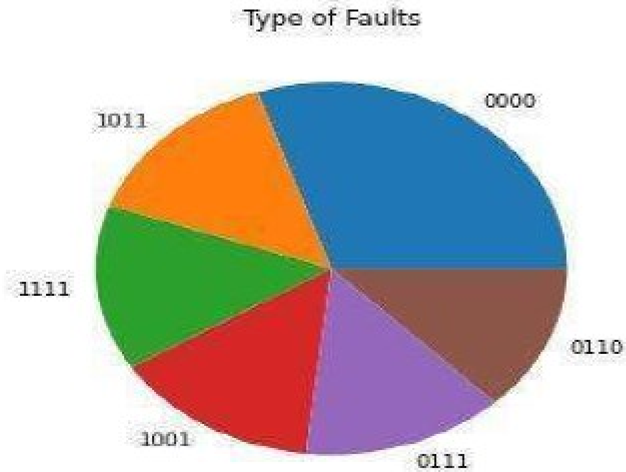


Figure 4.16: No of values at different faults

Figure 4.16 is a graphical depiction of the total number of errors found in the dataset. Current and voltage readings are used to classify the fault kinds, which are then used to produce matching graphs. For various faults, the graphs are drawn for lines A, B, and C.

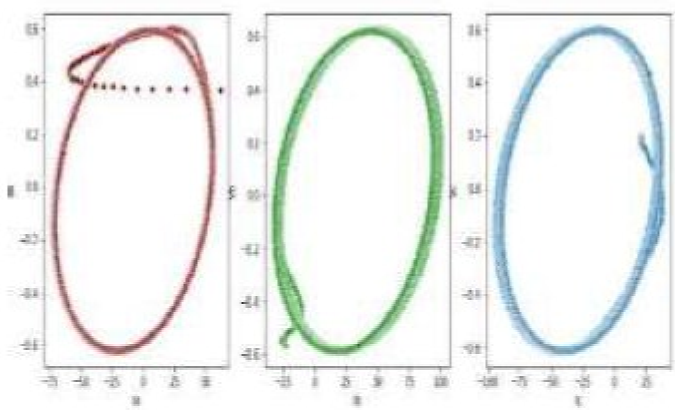


Figure 4.17:No fault power line

When line A's current flows ten times the usual amount—between phase A and ground—that's a line ground fault.

ISSN 2347–3657

Volume 13, Issue 2, 2025

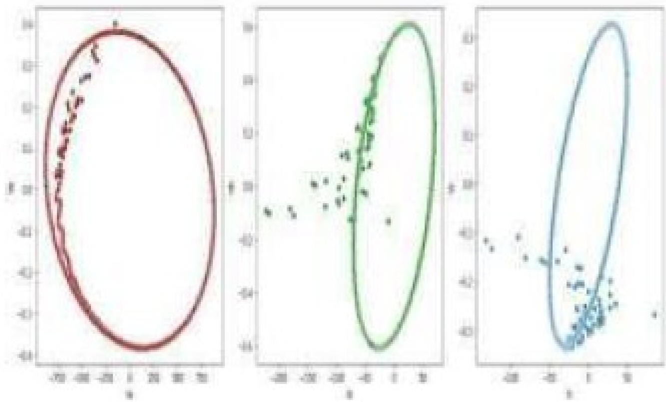


Figure 4.18:Line ground fault

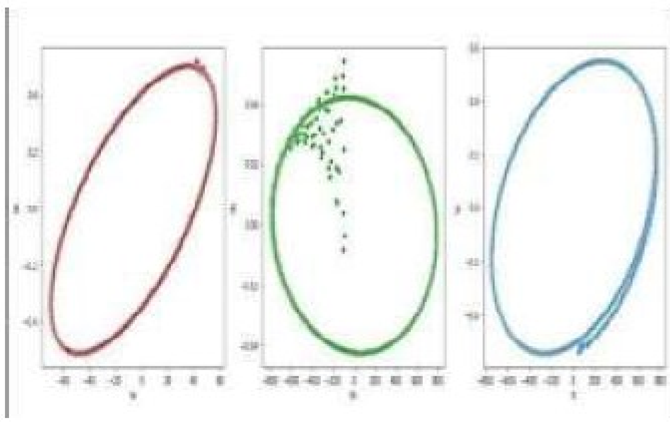


Figure 4.19: Line to Line Fault

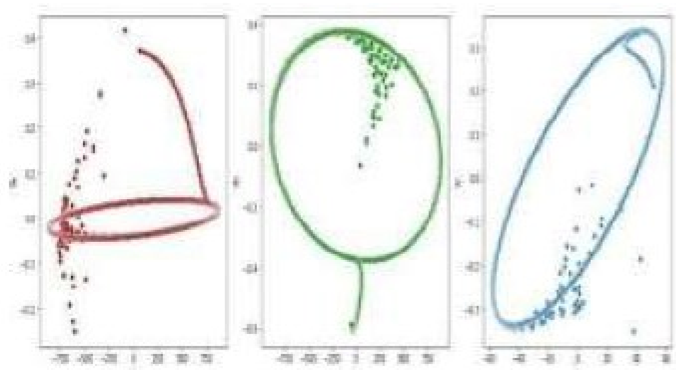


Figure 4.20:Line Ground Fault

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

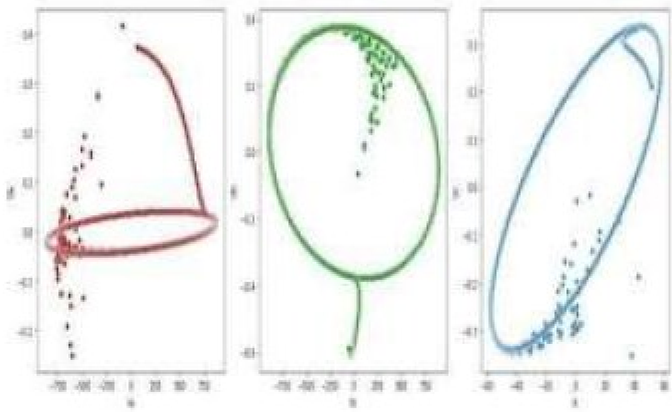


Figure 4.21: Line - Line fault

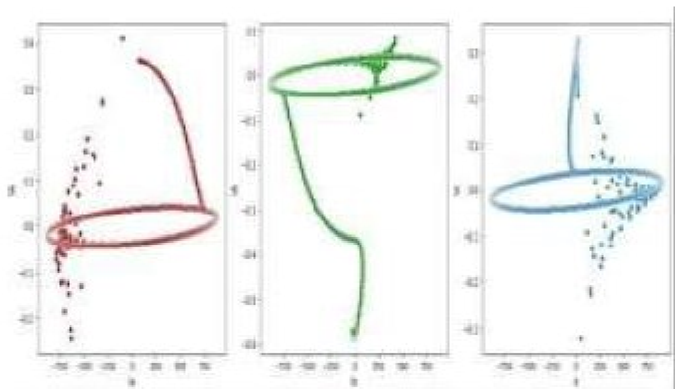


Figure 4.22: Line - Line Ground Fault

Figures 4.17, 4.18, 4.19, 4.20, 4.21, and 4.22 depict the many fault types that may occur. Various methods are evaluated using the detection dataset; the most efficient one is selected according to the score. Detection methods include Support Vector Machines (SVMs), Decision Tree Models, Random Forest Classifiers, and K-Nearest Neighbors.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

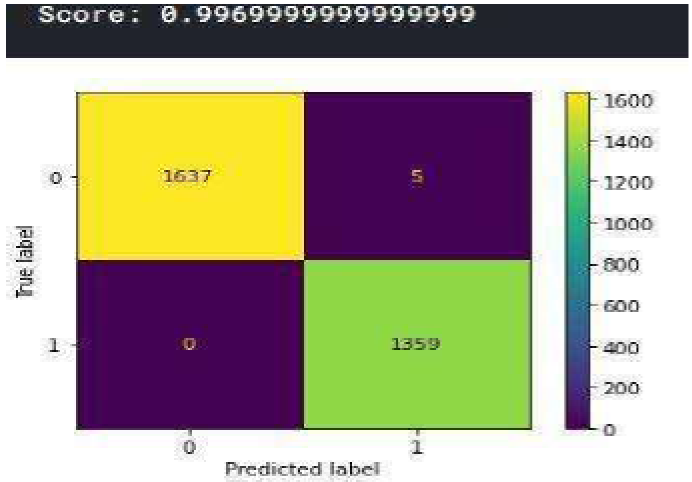


Figure 4.23:SVM model

Figure 4.23 depicts the label model and the score of SVM

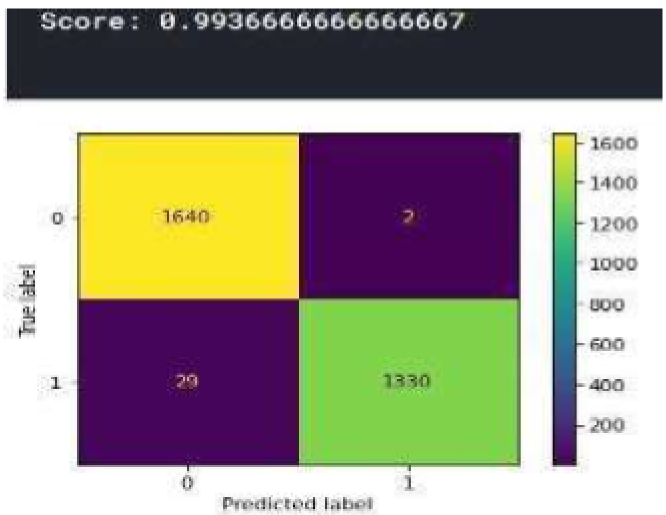


Figure 4.24:Decision Tree Model

Figure 4.24 depicts the label model and the score of Decision Tree Model

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Score: 0.9956666666666667

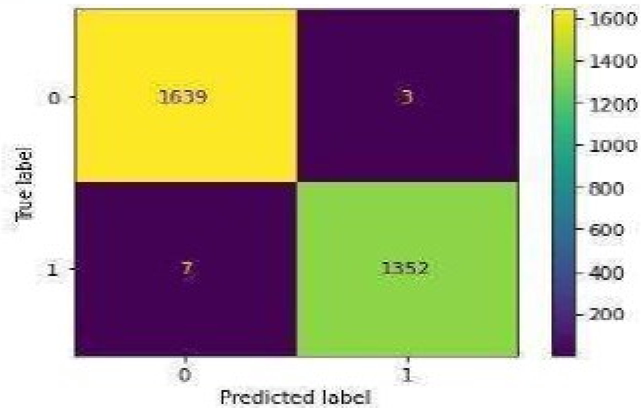


Figure 4.25:KNN Model

Figure 4.25 depicts the label model and the score of KNN Model

Score: 0.9903333333333334

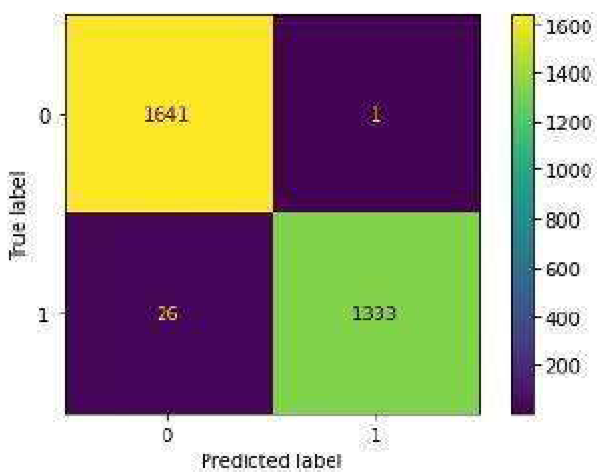


Figure 4.26:Random Forest Classifier Model

The label model and the score of the Random Forest Classifier are shown in Figure 4.26. Model

Since SVM can efficiently forecast all signals, it has outperformed the other models in fault detection. This is because, in certain circumstances, other models fail to detect faults even when there is a defect. It is the Labelling procedure that establishes the algorithms' correctness. After calculating the scores of each algorithm, SVM emerged victorious with a score of 0.996999999999, proving its efficiency.

V.CONCLUSION

ISSN 2347–3657

Volume 13, Issue 2, 2025

This article shows the results of testing the accuracy for numerous input configurations and analyzing them. It uses several machine learning models on different power system models to accurately detect, classify, and identify faults. We tested several algorithms and picked the one with the highest efficiency rating based on its accuracy.

REFERENCES

[1] A. Mukherjee, P.K. Kundu, A. Das, Applicat ion of Principal Component Analysis for Fault Classificat ion in Transmission Line with Ratio-Based Method and Probabilist ic Neural Net -work: a Comparat ive Analysis. J. Inst . Eng. India Ser. B 101(4), 321–333 (2020).
[2] B. Patel, P. Bera, Fast fault detect ion during power swing on a hybrid t ransmission line using WPT. IET Gener. Transm. Dis-trib. 13(10), 1811–1820 (2019)
R. Godse, S. Bhat , Mathematical morphology-based feature- ext ract ion technique for detect ion and classificat ion of faults on power t ransmission lines. IEEE Access 8, 38459–38471 (2020)
[3] M. Kazim, A.H. Khawaja, U. Zabit, Q. Huang, Fault Detect ion and Localizat ion for Overhead 11-kV Dist ribut ion Lines with Magnet ic Measurements. IEEE Trans. Inst rum. Meas. 69(5), 2028–2038 (2019)
[4] H. Shu, N. An, B. Yang, Y. Dai, Y. Guo, Single Pole-to-Ground Fault Analysis of MMC-HVDC Transmission Lines Based on Capacit ive Fuzzy Ident ification Algorithm. Energies 13(2), 319 (2020).
[5] B. Chat terjee, S. Debnath, Cross correlat ion aided fuzzy based relaying scheme for fault classificat ion in t ransmission lines. Engineering Science and Technology, an International Journal 23(3), 534–543 (2020).
[6] A. Mukherjee, P.K. Kundu, A. Das, Power system fault ident ificat ion and localizat ion using mult iple linear regression of principal component distance indices. Internat ional Journal of Applied Power Engineering 9(2), 113–126 (2020)
[7] R. Fahim, S. Niloy, A.H. Shat il, M.R. Hazari, S.K. Sarker, S.K. Das, An Unsupervised Protect ion Scheme for Overhead Transmission Line with Emphasis on Situat ions during Line and Source Parameter Variat ion, in: ICREST 2021 - 2nd Int . Conf. Robot . Elect r. Signal Process, 2021.
[8] Amanuel, Thomas, Amanuel Ghirmay, Huruy Ghebremeskel, Robel Ghebrehiwet,and Weldekidan Bahlibi. " Comparative Analysis of Signal P rocessing Techniques for Fault Detect ion in Three Phase Induct ion Mot or." Journal of Elect ronics 3, no. 01(2021): 61-76.
[9] H. Liang, Y. Liu, G. Sheng, and X. Jiang, ―Fault - cause ident ificat ion method based on adapt ive deep belief network and t ime – frequency characterist ics of t ravelling wave, ‖ IET Generation, Transmission & Dist ribut ion, vol 13, no. 5, pp. 724-732, 2019.
[10] Azzouz, B. Boussaid, A. Zouinkhi and M. N. Abdelkrim, "Mult i-faults classificat ion in WSN: A deep learning approach," 2020 20th Internat ional Conference on Sciences and Techniques of Automat ic Cont rol and Computer Engineering (STA), Monast ir, Tunisia, 2020, pp.343-348,doi: 10.1109/STA50679.2020.9329325.
[11] M. U. Usman, J. Ospina, and M. O. Faruque, " Fault classificat ion andlocat ion ident ificat ion in a smart dist ribut ion network using ann," in 2018 IEEE Power & Energy Society General Meeting (PESGM). IEEE, 2018,pp. 1–6.
[12] M.-F. Guo, N.-C. Yang, and W.-F. Chen, " Deeplearning- based fault classificat ion using hilbert - huang t ransform and convolutional neural network in power dist ribut ion systems," IEEE Sensors Journal, 2019.
[13] Mirshekali, H.; Dashti, R.; Shaker, H.R.; Samsami, R.; Torabi, A.J. Linear and Nonlinear Fault Locat ion in Smart Dist ribut ion Network under Line Parameter Uncertainty. IEEE Trans. Ind. Informatics 2021.
[14] Yin, X.; Li, X.; Zhang, Y.; Zhang, T.; Lu, C.; Ai, Q.; Li, Z.; Sun, Z. A Survey of Deep Learning and Its Applicat ion in Dist ribut ion Network. In Proceedings of the 2020 International Conference on Art ificial Intelligence in Informat ion and Communicat ion, ICAIIC 2020; Institute of Electrical and Elect ronics Engineers Inc., Fukuoka, Japan, IEEE, 2020; pp. 643–646