Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022)
IEEE Xplore Part Number: CFP22AV8-ART; ISBN: 978-1-6654-8425-1

2022 International Conference on Electronics and Renewable Systems (ICEARS) | 978-1-6654-8425-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEARS53579.2022.9751897

Electrical Faults-Detection and Classification using Machine Learning

[1] Janarthanam K,
Student of Electronics and Communication Engineering,
Sri Ramakrishna Engineering College, Coimbatore.
janarthanam.1802053@srec.ac.in

[2] Kamalesh P,
Student of Electronics and Communication Engineering,
Sri Ramakrishna Engineering College,
Coimbatore.
kamalesh.1802059@srec.ac.in

[3] Basil T V,
Student of Electronics and Communication Engineering,
Sri Ramakrishna Engineering College, Coimbatore.
basil.18802066@srec.ac.in

[4] Judeson Antony Kovilpillai J,
Faculty of Electronics and Communication Engineering,
Sri Ramakrishna Engineering College, Coimbatore.
judeson@srec.ac.in

Abstract—This paper focuses on the detection of faults occurring in electrical power transmission lines using machine learning techniques and algorithms. In the recent era, the demand for electricity is increasing rapidly but on the other hand the transmission capacity is not developing accordingly. This paper discusses the most common transmission line faults and classification by using machine learning. Using the proposed methods, the faults are analyzed with different combinations of inputs and by doing so an accurate result is obtained. The machine learning algorithms are implemented in Spyder IDE (Scientific Python Development Environment). The objective is aimed to be addressed using this method.

Keywords—SVM** (support vector machine)**, KNN (**K-nearest Neighbour**), Machine Learning, LSTM (**Long Short-Term Memory**), Decision tree model, Random forest classifier model.

I. INTRODUCTION

In Recent years the power system plays a major role in our day-to-day life. They operate under a balanced condition. The major drawback arises in the power system is fault analysis. Detection and control of fault in the transmission line can be done to work in normal flow. Fault is created due to different reasons like short circuiting and some natural disturbance of lighting, earthquake etc. In order to rectify this problem, fault classification can be done. By using machine learning technique different types of faults can be classified. All these different types of faults can be classified, only if the system is in an unbalanced condition. They can be classified as symmetrical and unsymmetrical. Mostly unsymmetrical faults can occur in the transmission line like L-G Fault (single to ground fault) and L-L Fault (Line to Line faults). The new method discussed below has considerable advantage over the traditional and existing methods. For obtaining best and the accurate results, we have totally made use of four algorithms and the best of these outcomes are considered as result.

II. RELATED WORKS

Zuraida Muhammad and Shabinar Abd Hamid (2020) implemented a ml algorithm technique to improve the quality of power and overcome the problem in the transmission using ANN (Artificial Neural Network). The quality of power can be improved by analysing fault detection and fault classification. Firstly, the fault can be created by using an impedance technique. After the stimulation, faulty current and voltage are measured and taken as an input. The feed forward network along with back propagation algorithms are used to develop the fault classification and detection. The detection and the classification performance were measured using MSC (mean square error). The deduction is achieved 5.6 148 of MSC tolerance and mainly provides 100% of accuracy. The classification achieved 0.893955 of MSC tolerance and provided the 70 % of perfection from this proposed method the fault can easily be classified.

Zakaria Hussian (2020) implemented a deep learning method to detect the problem in the power system. This proposed method introduces a new tool called a LSTM (Long Short-Term Memory) to detect and analyse faults in the power systems. In the transmission line, a fault current signal is extracted. This signal is fed to the Lstm(long short-term memory) network mainly acting as an input for the fault classification process. The gaussian noise level in the range of 20 DB to 30 DB of SNR (Signal to noise ratio) is added to strengthen the proposed model. After the simulation process, different values are obtained. The obtained values range from 0 to 10. Based on the values, fault can be easily classified. if the obtaining value is zero, the system is considered a non-fault system. if it ranges from 1 to 10, Ten different types of faults can be analysed depending upon the values. In the LSTM method Results are easily obtained compared to other methods. This Proposed method gives 100% of accuracy for detection and classification process in the transmission line.

Monica (2019) proposed a new method to detect and control the fault in the transmission line using IoT. This proposed method uses a voltage sensor that can be connected to the microcontroller to sense the voltage flow in the transmission line. If the voltage flow limit varies from the normal level, the fault can be easily identified. If the limit exceeds the voltage flow, a relay is used to trip the circuit. The GPS(Global Positioning System) is connected to the controller to detect fault location. the flame sensor is used to detect the fire or flame occurring in the transmission line, which can be kept under 100 cm distance due to high temperature. All the collected data is sent to the cloud by using esp8266. The location and the detection of fault occurred results can be viewed in liquid crystal display (LCD).

III.EXISTING METHOD

The project is divided into two parts, fault detection and classification. The model is simulated in MATLAB(MATrix LABoratory). Then the training and the classification is done in the classification learner app which is provided by MATLAB. Using this, all the models got trained and validated at a faster rate. In addition to all these, the dataset was trained for around 24 Machine Learning models and the models.

In this method different input configurations are considered for training the model. A combination of these models with a higher accuracy is considered for the output calculations. There are mainly four models considered.

Another commonly used method is the end-to-end fault detection technique. This method takes the parameters of current and voltages at the fault occurrence time. These values are then used for the training of the model. Targets are categorized into two types. For detection model targets are Fault, No Fault decisions. For classification models targets are No-fault, Line to Ground fault (LG), Line to Line fault (LL), Line to Line to Ground fault (LLG) and Line to Line-to-Line fault (LLL).

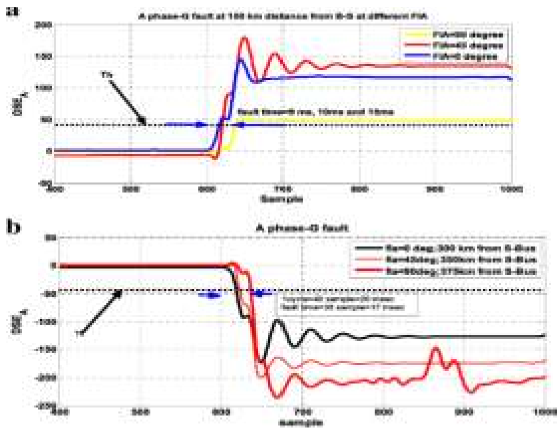


Figure 3.1:fault samples

The terminal method is another way of detecting the faults in the transmission line. The image of fault samples is shown in Figure 3.1. In this method a resistor is used for detecting the fault occurrence and by measuring the values of voltage and current at the time of fault occurrence, it can be segregated. This method is used to detect the fault location from the starting point to the end without tracing.

IV.PROPOSED SYSTEM

The working procedure adopted for the classification and detection of electric faults is followed as a step-by-step process. The first step is to read input from the dataset. Then the pre-processing is carried out in the Spyder IDE. Then the dataset is trained. The trained dataset is then classified into faulty and non-faulty models. The faulty models are then classified into more distinct faults and studied for experimental analysis. Several algorithms are used for this process. The general algorithm used for the classification and detection of electric faults in electric lines is depicted in the Figure 4.1.

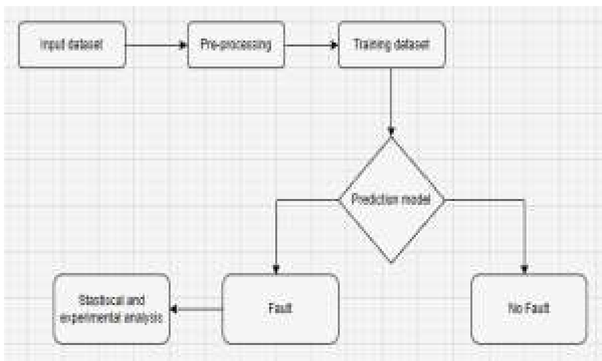


Figure 4.1:Proposed workflow

Normally a power system operates under a balanced condition. If any unbalance occurs in the system represents the fault. These faults may occur due to natural reasons or due to any improper maintenance of the transmission lines, which leads to faults like insulation failures. These faults occurred in the transmission lines are categorised into many as follows,

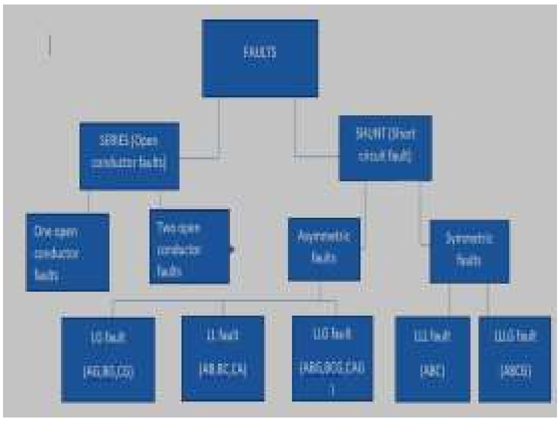
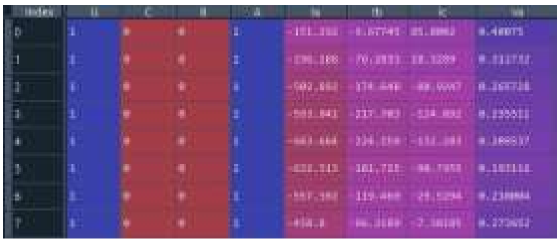


Figure 4.2:Classification of faults

Figure 4.2 shows the classification of electric faults. In this proposed system, we use python along with spyder IDE for the detection and classification of faults. In the first stage, the short-circuit fault is detected and based on the values of current and voltages obtained in the fault, it is further classified into sub-categories. The parameters of line voltage and line current are made to compare with the preloaded datasets and hence the faults are classified.

The dataset represents the dataset to train the model to detect the type of fault. Here in this dataset, I_a, I_b, I_c are the three currents of the 3-phase line system and V_a, V_b, V_c are the respective lines.

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



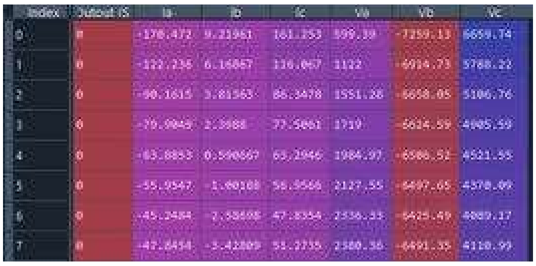
Index	G	C	B	A	Ia	Ib	Ic	Va
0	1	0	0	1	-151.332	-9.67745	85.8062	0.40075
1	1	0	0	1	-236.186	70.2833	18.3289	0.312732
2	1	0	0	1	-502.892	174.646	-80.9247	0.265728
3	1	0	0	1	-593.942	217.703	-124.892	0.235511
4	1	0	0	1	-643.664	224.159	-132.283	0.209537
5	1	0	0	1	-632.315	161.725	-98.7058	0.183114
6	1	0	0	1	-557.392	115.460	-29.5294	0.238884
7	1	0	0	1	-458.8	86.3189	-7.38185	0.273652

Figure 4.3:Classification:dataset

The dataset shown in Figure 4.3 is for the classification of shunt faults

Classification Dataset

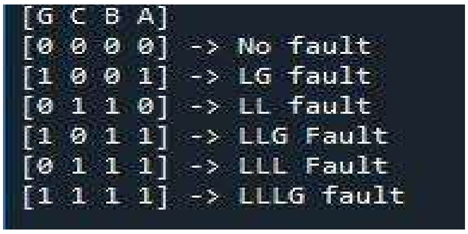
Under normal conditions the current varies from -100A to 100A and the voltage per unit is -0.6 to 0.6. While during fault, we notice some absurd and random behaviour and the value of Line current even touches -/+800A mark. The classification dataset is shown in Figure 4.3. The dataset is downloaded from the weblink kaggle/input/electrical-fault-detection-and-classification/classData.csv



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	65.2946	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

Figure 4.4 shows the image of the detection dataset.The dataset is downloaded from the weblink kaggle/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

The types of faults are listed in Figure 4.5.

Classification Algorithm

SVM stands for support vector machine which is a widely used and advanced classifier. In SVM classification the data can be either linear or non- linear. The SVM is working perfectly till today due to its ability towards detection of signal in a better manner than any other techniques. The SVM model is taken as follows,

V.EXPERIMENTAL RESULTS:

The project is mainly focused on detection and classification of faults in transmission lines. The results of the same are discussed below. The datasets of two models for classification and detection were used and implemented for certain machine learning algorithms. The algorithm with the most accurate result is also given below. The number of values in both the datasets are depicted in a bar graph below. Both the datasets have no duplicate values.

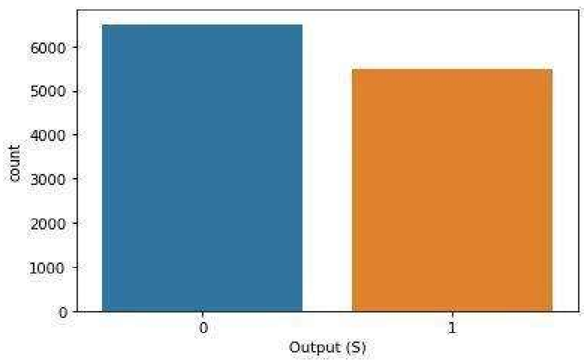


Figure 4.6:Classification and detection dataset range

Both the datasets are plotted with their values in Figure 4.6. The current in the detection dataset is plotted and shown below.

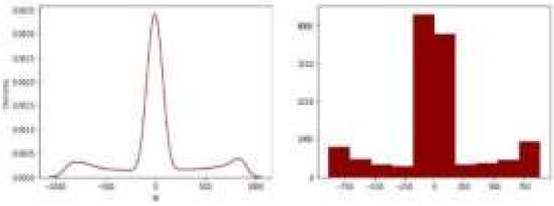


Figure 4.7:Current Ia

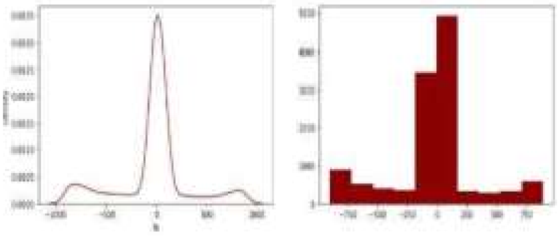


Figure 4.8:Current Ib

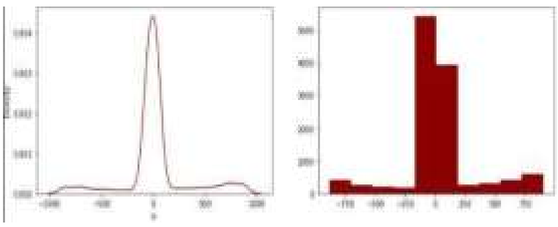


Figure 4.9:Current Ic

Figures 4.7,4.8,4.9 represent the current graph of the three lines with values from the classification dataset. Current value is taken in the x axis while density is taken along the y axis. The voltage graph is also plotted for the values in the dataset and the graph resulting shows the value where the voltages lie in the dataset.

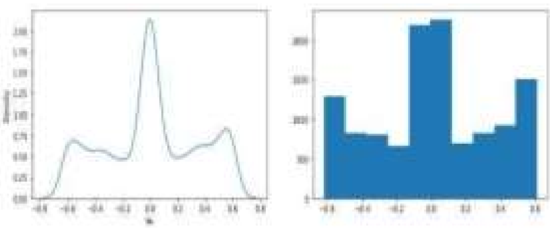


Figure 4.10:Voltage Va

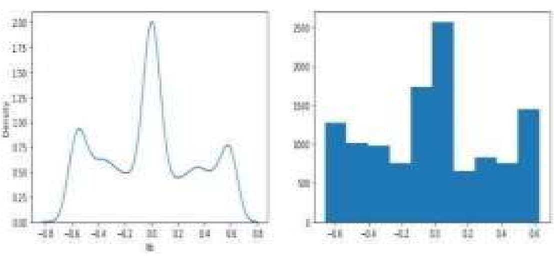


Figure 4.11:Voltage Vb

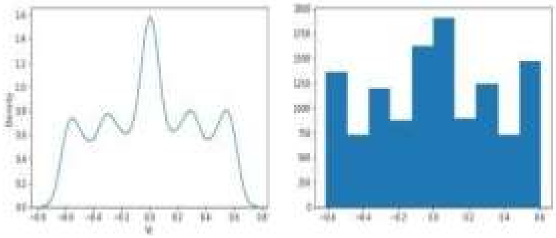


Figure 4.12:Voltage Vc

Figures 4.10,4.11,4.12 represent the voltage graph of the three lines with values from the classification dataset. Voltage value is taken in the x axis while density is taken along the y axis.The three lines A, B,C and their graphs are plotted based on the values of the current and voltage and the line mediums are shown below. All the data is normally distributed. The signal flow graph with respect to lines A,B and C are plotted.

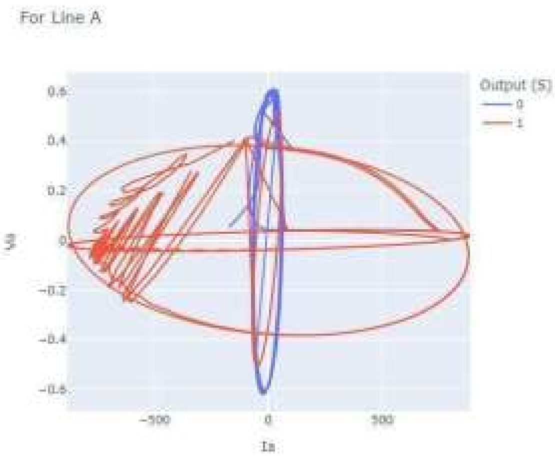


Figure 4.13: Signal flow graph of line A

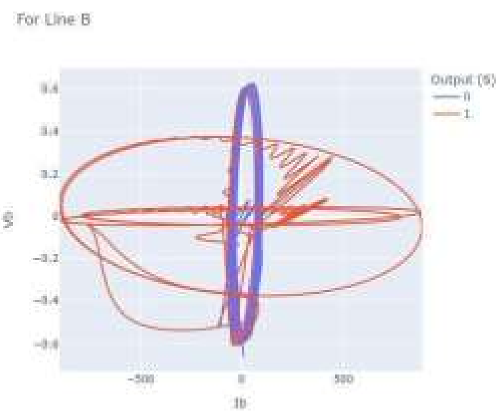


Figure 4.14: Signal flow graph of Line B

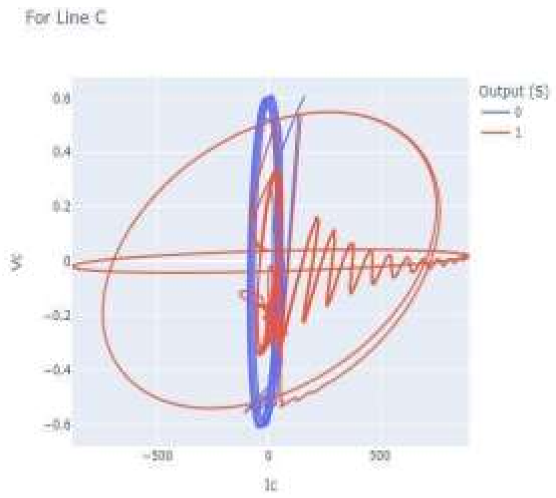


Figure 4.15: Signal flow graph of Line C

The graphs are plotted with the current as the x component and voltage as the y component. Figures 4.13,4.14,4.15 represent the signal flow graphs of Lines A,B and C. It is observed that normally the line current varies from -100 to 100 amperes and voltage between -0.6V and 0.6V. During fault, a random pattern and absurd behavior is observed. The value of the line current even touches the value of -/+800 ampere mark. The classification dataset is used to classify the faults based on the values of [G C B A] where the values of ground and the lines are taken into account.

The fault types are given as

[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

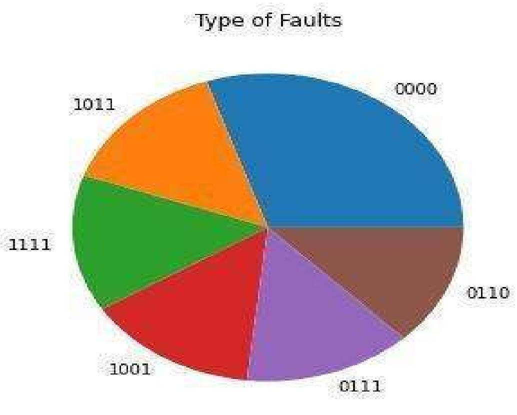


Figure 4.16: No of values at different faults

A pictorial representation of the number of faults in the dataset is shown in Figure 4.16. The types of faults have been classified based on the values and their corresponding graphs are plotted according to the current and voltage values. The graphs are plotted for the three lines A, B and C for different faults.

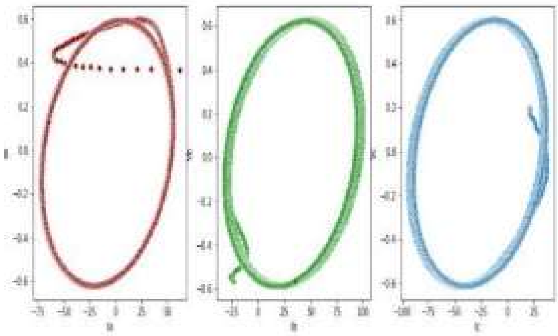


Figure 4.17:No fault power line

For line ground fault, the fault occurs between phase A and ground where the amount of current flowing in line A is 10 times more than the normal current.

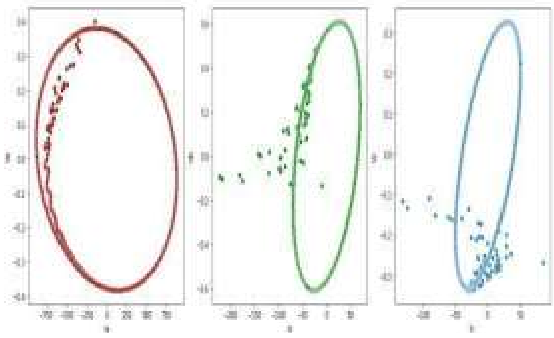


Figure 4.18:Line ground fault

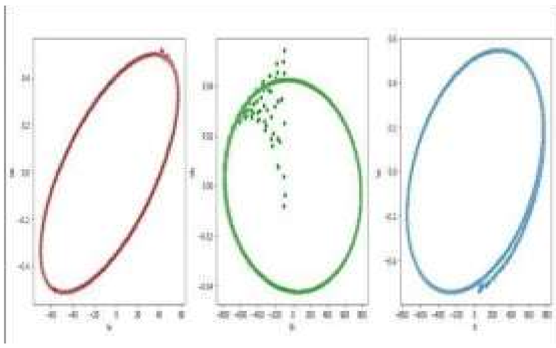


Figure 4.19: Line to Line Fault

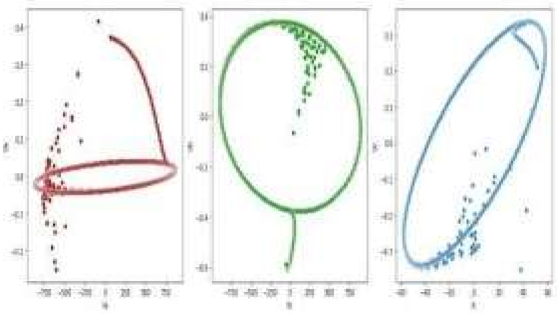


Figure 4.20:Line Ground Fault

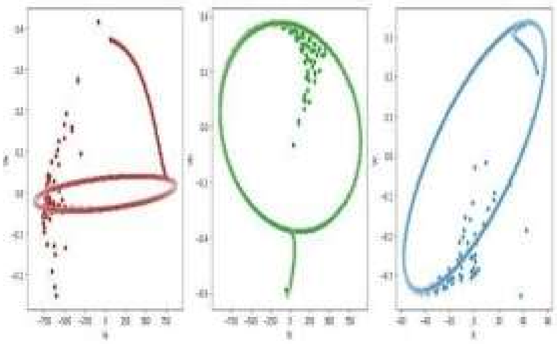


Figure 4.21: Line - Line fault

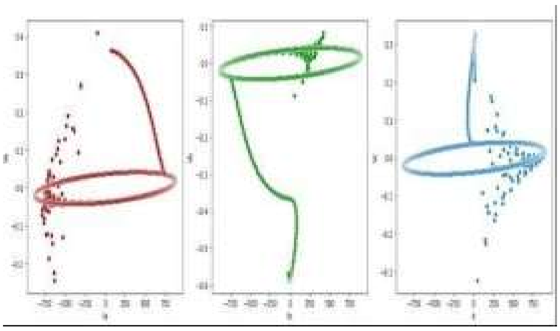
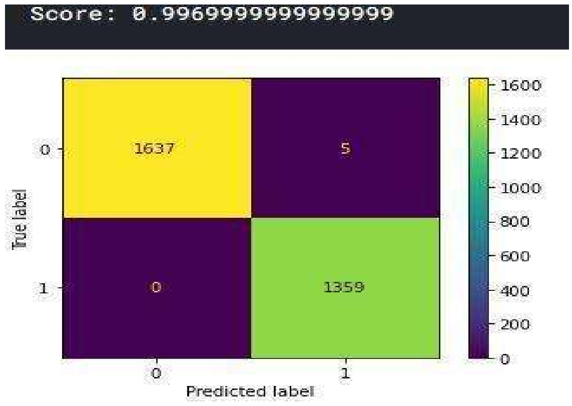


Figure 4.22: Line - Line Ground Fault

The different types of faults occurring are shown in Figures 4.17,4.18,4.19,4.20,4.21,4.22.

The detection dataset is used and different algorithms are tested and the one with the most efficient result is found based upon the score. SVM (Support Vector Machine), Decision tree model, Random Forest Classifier and KNN(K-Nearest Neighbour) are the algorithms used for detection



.

Figure 4.23:SVM model

Figure 4.23 depicts the label model and the score of SVM

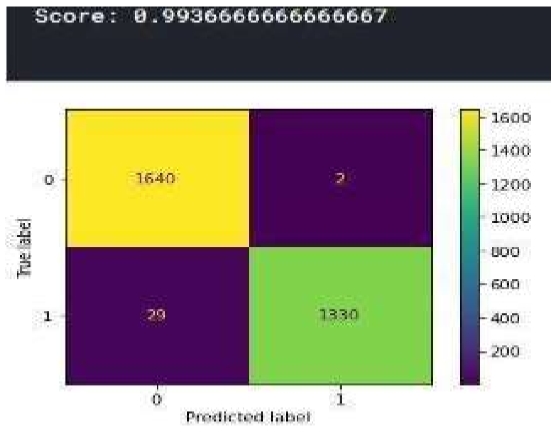


Figure 4.24:Decision Tree Model

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

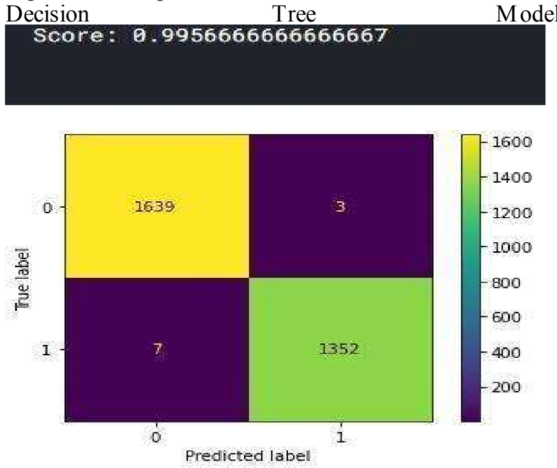


Figure 4.25:KNN Model

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

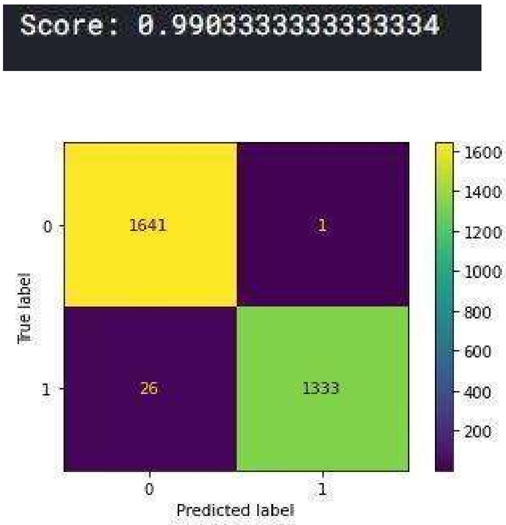


Figure 4.26:Random Forest Classifier Model

Figure 4.26 depicts the label model and the score of Random Forest Classifier Model

SVM has done better in Fault Detection than the rest of the models because it's able to predict all the signals in the most efficient manner while in other models there are cases where there is actually fault but the model is not able to identify it. The accuracy of the algorithms is determined by the Labelling method. The scores of all the algorithms are determined and SVM is found to be the algorithm with the highest score 0.996999999999 and hence it is proved to be efficient.

V.CONCLUSION

Accurate detection, classification and identification of fault are performed using various machine learning models on various power system models and the accuracy for different configuration of input were tested and the results were analyzed in this paper as shown above. The accuracy was evaluated for different algorithms and the most efficient one is determined and used.

REFERENCES

[1] A. Mukherjee, P.K. Kundu, A. Das, Application of Principal Component Analysis for Fault Classification in Transmission Line with Ratio-Based Method and Probabilistic Neural Net-work: a Comparative Analysis. J. Inst. Eng. India Ser. B 101(4), 321–333 (2020).

[2] B. Patel, P. Bera, Fast fault detection during power swing on a hybrid transmission line using WPT. IET Gener. Transm. Dis-trib. 13(10), 1811–1820 (2019) R. Godse, S. Bhat, Mathematical morphology-based feature- extraction technique for detection and classification of faults on power transmission lines. IEEE Access 8, 38459–38471 (2020)

[3] M. Kazim, A.H. Khawaja, U. Zabit, Q. Huang, Fault Detection and Localization for Overhead 11-kV Distribution Lines with Magnetic Measurements. IEEE Trans. Instrum. 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 Capacitive Fuzzy Identification Algorithm. Energies 13(2), 319 (2020).

[5] B. Chatterjee, S. Debnath, Cross correlation aided fuzzy based relaying scheme for fault classification in transmission lines. Engineering Science and Technology, an International Journal 23(3), 534–543 (2020).

[6] A. Mukherjee, P.K. Kundu, A. Das, Power system fault identification and localization using multiple linear regression of principal component distance indices. International Journal of Applied Power Engineering 9(2), 113–126 (2020)

[7] R. Fahim, S. Niloy, A.H. Shatil, M.R. Hazari, S.K. Sarker, S.K. Das, An Unsupervised Protection Scheme for Overhead Transmission Line with Emphasis on Situations during Line and Source Parameter Variation, in: ICREST 2021 - 2nd Int. Conf. Robot. Electr. Signal Process, 2021.

[8] Amanuel, Thomas, Amanuel Ghirmay, Huruy Ghebremeskel, Robel Ghebrehiwet,and Weldekidan Bahlibi. “Comparative Analysis of Signal Processing Techniques for Fault Detection in Three Phase Induction Motor.” Journal of Electronics 3, no. 01(2021): 61-76.

[9] H. Liang, Y. Liu, G. Sheng, and X. Jiang, ―Fault-cause identification method based on adaptive deep belief network and time – frequency characteristics of travelling wave, ‖ IET Generation, Transmission & Distribution, vol 13, no. 5, pp. 724-732, 2019.

[10] Azzouz, B. Boussaid, A. Zouinkhi and M. N. Abdelkrim, "Multi-faults classification in WSN: A deep learning approach," 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), Monastir, Tunisia, 2020, pp.343-348,doi: 10.1109/STA50679.2020.9329325.

[11] M. U. Usman, J. Ospina, and M. O. Faruque, “Fault classification and location identification in a smart distribution 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, “Deep-learning-based fault classification using hilbert-huang transform and convolutional neural network in power distribution systems,” IEEE Sensors Journal, 2019.

[13] Mirshekali, H.; Dashti, R.; Shaker, H.R.; Samsami, R.; Torabi, A.J. Linear and Nonlinear Fault Location in Smart Distribution 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 Application in Distribution Network. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2020; Institute of Electrical and Electronics Engineers Inc., Fukuoka, Japan, IEEE, 2020; pp. 643–646