

# An Intelligent Decision Tree Framework For Power System Fault Classification And Localization

Ashish Lonare

M.Tech Student, School of  
Electrical & Electronics  
Engineering,  
Scope Global Skills University,  
Bhopal, M.P., India  
[ashishlonare101@gmail.com](mailto:ashishlonare101@gmail.com)

Anjana Tripathi

Assistant Professor  
School of Electrical & Electronics  
Engineering, Scope Global Skills  
University  
Bhopal, M.P., India

Balram Yadav

HOS  
School of Electrical & Electronics  
Engineering,  
Scope Global Skills University,  
Bhopal, M.P., India

**Abstract:** This paper is majorly directed towards exploring the newly emerging idea in the field of machine learning, which is termed as decision tree (DT). In this thesis, we have discussed about how they are being designed using mathematical formulae, and incorporate the basic principle of the classification of the input data. The DT is quite simple to design when compared to other available choices. They are easy to train, and give efficient performance for practical purposes. Since the DT is used in various diverse applications, they are also tweaked to make them usable for a particular operation by using modified DT, or Kernels. The DT are now being in usage for various industrial purposes too and are doing better than the conventional machine learning, AI based systems. As discussed a power system can be encountered with the faults named as AG, BG, CG, AB, BC, CA, ABG, BCG, CAG, ABC and ABCG phase fault. Hence it should be equipped suitably to tackle these faults in the most appropriate manner. Tackling these faults means to classify and finding out its location and graveness. In past on occurrence of fault, current is measured from either ends of the line which were then used in the algorithm to classify them. It could be the raise in magnitude of current magnitude as the data in algorithm. Since each fault react differently i.e. different characteristics of current when it occurs.

**Keywords:** Decision tree, Conventional machine learning, phase faults, AI based systems

## LINTRODUCTION

An electric power system contains of generation, transmission and distribution of electric energy. Overhead transmission lines are the maximum

advantageous approach for transportation of electrical energy from sources of generation to load centers to bring about electricity use and consumption. The quick development of electric power systems in the course of recent decades has brought about an extensive increment of the quantity of lines in operation and their aggregate length. These lines are uncovered to faults that are caused by short circuits, overload, tree branches, faulty equipments, lightning, human errors. In most cases, electrical faults manifest themselves as mechanical damage, which should be repaired earlier than returning the line to service. Any fault, if no longer detected and separated rapidly will in this manner, develop into a system wide unsettling influence causing blackouts and even ensuing power outages [1, 2]. When a fault occurs on a power system, economic losses may be decreased and the line service can be maintained if the location of the fault may be as it should be decided, specifically while generation, transmission and distribution happens over a longer distance, thereby enhancing the safety and quality of the power supply. Electric utility services would thus be able to be kept up if the detection, classification and location of the fault on a line can be precisely decided. These suitability factors and deductions make fault analyzers and their algorithms, a critical tool in keeping the competence of smart grids. Faults cause short to long term power

outages for customers and can result in extensive losses especially for the producing enterprise [3].

### 1.1 DECISION TREE

In knowledge discovery of databases, classification called as supervised learning since the labeled instances are classified into the predefined classes. A good increase in momentum of research is observed in the area of classification in recent decade due to wide applicability in real world applications. Decision tree is one of the most popular supervised learning models for knowledge discovery. Decision trees are used to make decisions for the unseen cases with the help of the model build with the trained classes. A decision tree consists of set of root node, branches and leaf nodes. A decision is made in the decision tree by traversing from root node to the leaf node through branches of the tree. The main issue with knowledge based systems is the inability to explain the reason behind the decision making. Where as decision trees not only provide a decision but also complete reason for the decision making. The decision rules can be generated from the output of the decision trees. These decision rules can form as the core unit for building the expert systems. The decision tree model are also efficient in terms of elaborating the reasons behind the decision making [3]. While implementing decision trees there is no need of human experts [4]. The above reasons motivated to use decision trees in real world applications.

## II. POWER QUALITY PROBLEMS

Power quality issues are of most concern now-a-days. The wide spread of electronic gear, for example, data innovation hardware, power gadgets, for example, ASD, PLC, energy-efficient lighting, prompted a total difference in electric burdens nature. These heaps are all the while the significant causers and significant casualties of force quality issues. Because of their non-linearity, this multitude of burdens cause aggravations in the voltage waveform. Alongside innovation

advance, the association of overall economy has developed towards globalization and the net revenue of numerous exercises will in general diminish. The expanded awareness of by far most of cycles (modern, benefits and surprisingly private) to PQ issues turns the accessibility of electric power with quality an essential element for intensity in each area. The most basic regions are the consistent cycle industry and the data innovation administrations. At the point when an unsettling influence happens, monetary misfortunes might occur, with a subsequent loss of usefulness and seriousness. Albeit numerous endeavors have been taken by the utilities, a few shoppers require a degree of PQ higher than given by present day electric organizations. This suggests that a few estimates should be taken to accomplish more elevated levels of force quality.

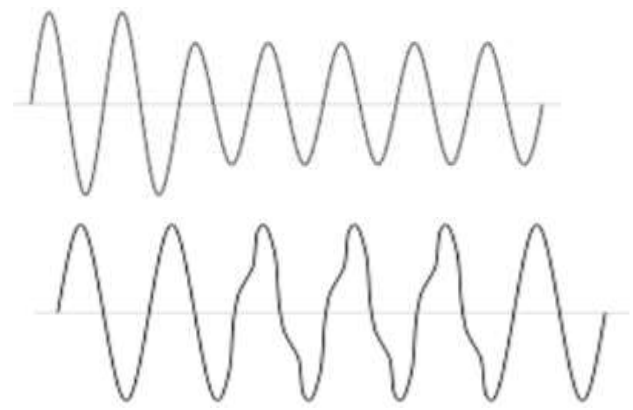


Figure 2.1 : Under voltage

Figure 2.2: Harmonic

## III. ELECTRICAL FAULTS

In electrical power systems consisting of generators, transformers, transmission lines and distribution circuits, most of the faults, about two-thirds, are liable to occur in the transmission lines. A fault in a circuit is any failure which interfaces with the normal flow of current. The deficiencies happen in power framework because of protection disappointment of gear's, flashover of lines started by a lightning stroke, because of extremely durable harm to guides and towers or because of unintentional defective tasks. The faults can

be broadly classified into shunt faults(short circuits) and series faults(open conductors). The shunt issue includes hamper guide and ground or short out between at least two guides. The shunt shortcomings are portrayed by expansion in current and fall in voltage and recurrence. Shunt shortcomings are delegated follows:

1. LG fault
2. LL fault
3. LLG fault
4. Three phase fault

Table 3.1: Types of faults

Types	Symbol	% of Appearance	Acuteness
Line to Ground	L-G	75-80%	Fall under the category of not so grave
Line to Line	L-L	10-15%	Sever compared to L-G but less than L-L-G
Double Line to Ground	L-L-G	3-10%	Graver
Three phase	3- $\phi$	1-5%	Most grave

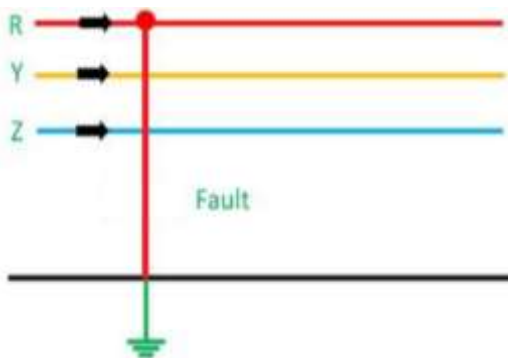


Figure 3.1: Single Line to Ground Fault

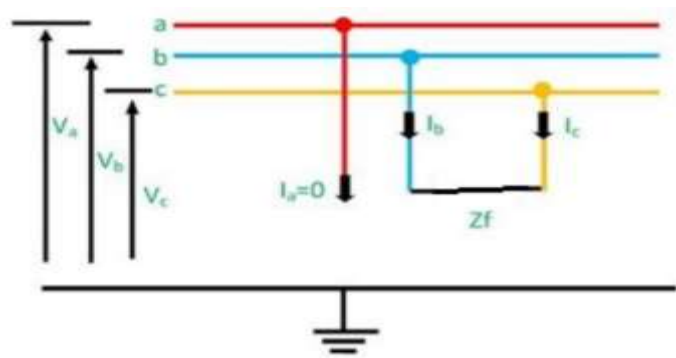


Figure 3.2: Line to Line Fault

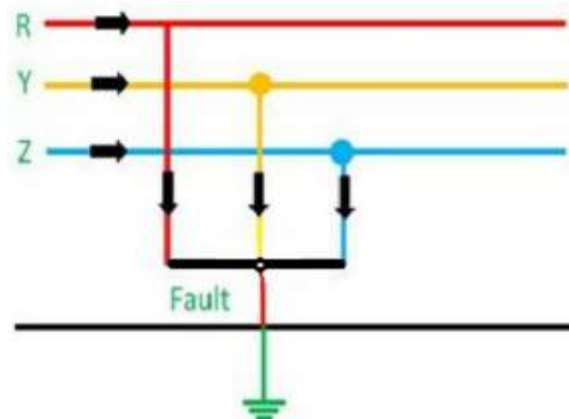
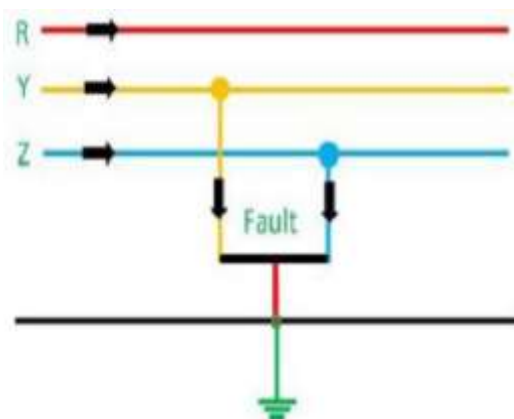


Figure 3.3: Double Line to Ground Fault

Figure 3.4: Three Phase Fault

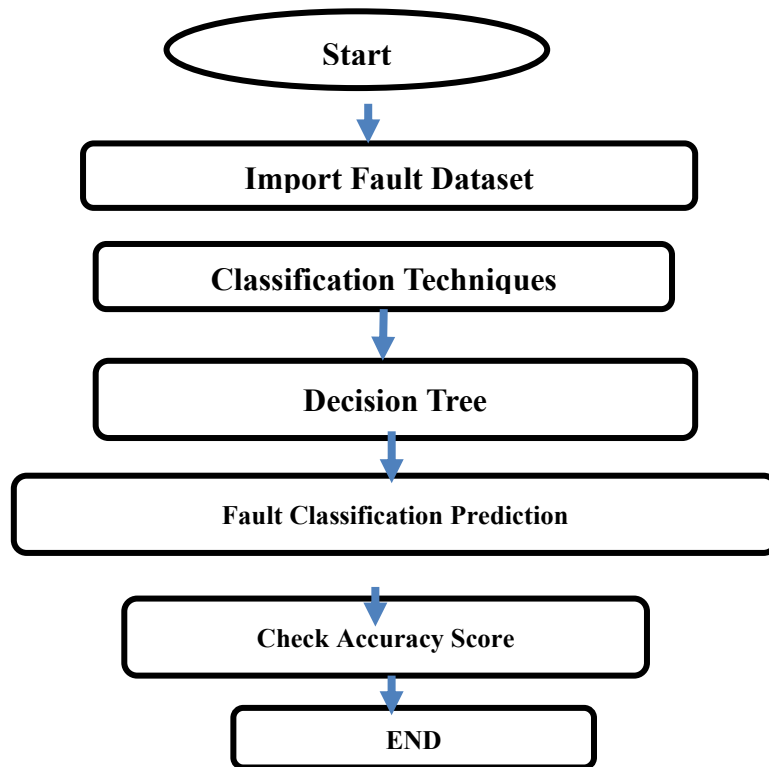
**IV. DATA PRE-PROCESSING IN DECISION TREES**

Decision trees are supervised classifiers that learn from the training dataset and classify target instances on a trained decision tree. Small variations in data change the structure of decision trees. Data preprocessing is an important step that helps to enhance the quality of data and help the drawing out of meaningful information

from the data [1]. Data preprocessing is the procedure of planning (cleaning and sorting out) the information to make it reasonable for preparing and building Machine Learning models. Information preprocessing is an information mining process that changes crude information into a readable and understandable format and suitable to create data mining models [18].

*4.1 Flowchart & Algorithm*

4.2



**V. RESULTS**

```

[4] df = pd.DataFrame()
for i in ["/content/drive/MyDrive/AB Fault/ABFAULT25.xlsx", "/content/drive/MyDrive/BC Fault/BCFAULT25.xlsx", "/content/drive/MyDrive/CA Fault/CAFAULT25.xlsx"]:
    a = pd.read_excel(i)
    a.rename(columns = {'Fault Type': 'Fault_Type'}, inplace = True)
    df = pd.concat([df, a])
  
```

```

from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

accuracy_score(Y_Test, Y_RF_pred)

0.4950166112956811
    
```

25 Ω				Accuracy
Fault	AB	BC	CA	49.50

```

[65] df = pd.DataFrame()
for i in ['/content/drive/MyDrive/AB Fault/ABFAULT25.xlsx', '/content/drive/MyDrive/BC Fault/BCFAULT25.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT25.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT25.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT25.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT25.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT25.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT25.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT25.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT25.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT25.xlsx', '/content/drive/MyDrive/AB Fault/ABFAULT50.xlsx',
'/content/drive/MyDrive/BC Fault/BCFAULT50.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT50.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT50.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT50.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT50.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT50.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT50.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT50.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT50.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT50.xlsx', '/content/drive/MyDrive/AB Fault/ABFAULT75.xlsx',
'/content/drive/MyDrive/BC Fault/BCFAULT75.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT75.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT75.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT75.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT75.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT75.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT75.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT75.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT75.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT75.xlsx']:
a = pd.read_excel(i)
a.rename(columns = {'Fault Type': 'Fault_Type'}, inplace = True)
df = pd.concat([df, a])

a = pd.read_excel(i)
a.rename(columns = {'Fault Type': 'Fault_Type'}, inplace = True)
df = pd.concat([df, a])
    
```

```

from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

accuracy_score(Y_Test, Y_RF_pred)

0.7642792384406165
    
```

25 & 50 & 75 Ω												Accuracy
Fault	AG	BG	CG	AB	BC	CA	ABG	BCG	CAG	ABC	ABCG	76.42

```

[81] df = pd.DataFrame()
for i in ['/content/drive/MyDrive/AB Fault/ABFAULT25.xlsx', '/content/drive/MyDrive/BC Fault/BCFAULT25.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT25.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT25.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT25.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT25.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT25.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT25.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT25.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT25.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT25.xlsx', '/content/drive/MyDrive/AB Fault/ABFAULT50.xlsx',
'/content/drive/MyDrive/BC Fault/BCFAULT50.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT50.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT50.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT50.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT50.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT50.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT50.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT50.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT50.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT50.xlsx', '/content/drive/MyDrive/AB Fault/ABFAULT75.xlsx',
'/content/drive/MyDrive/BC Fault/BCFAULT75.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT75.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT75.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT75.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT75.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT75.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT75.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT75.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT75.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT75.xlsx', '/content/drive/MyDrive/AB Fault/ABFAULT100.xlsx',
'/content/drive/MyDrive/BC Fault/BCFAULT100.xlsx', '/content/drive/MyDrive/CA Fault/CAFAULT100.xlsx',
'/content/drive/MyDrive/ABC Fault/ABCFAULT100.xlsx', '/content/drive/MyDrive/ABC Fault/ABCFAULT100.xlsx', '/content/drive/MyDrive/ABG Fault/ABGFAULT100.xlsx',
'/content/drive/MyDrive/BCG Fault/BCGFAULT100.xlsx', '/content/drive/MyDrive/BG Fault/BGFAULT100.xlsx', '/content/drive/MyDrive/CAG Fault/CAGFAULT100.xlsx',
'/content/drive/MyDrive/CG Fault/CGFAULT100.xlsx', '/content/drive/MyDrive/AG Fault/AGFAULT100.xlsx']:
a = pd.read_excel(i)
a.rename(columns = {'Fault Type': 'Fault_Type'}, inplace = True)
df = pd.concat([df, a])
    
```

```

from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

accuracy_score(Y_Test, Y_RF_pred)

0.7864429834584647
    
```

25 & 50 & 75 & 100 Ω												Accuracy
Fault	AG	BG	CG	AB	BC	CA	ABG	BCG	CAG	ABC	ABCG	78.64

**VI.CONCLUSION**

Protective relaying system is a versatile tool for protection of electric power systems. Because of the drastic changes occurring in power systems, the necessity for providing better protective relaying systems for transmission lines is essential. This work gives brief overview on different existing techniques for fault analysis and apart from the existing methodologies a novel fault detection scheme is presented in this work. It uses post fault current samples of all the phases. As discussed already a power system can be encountered with the faults named as AG, BG, CG, AB, BC, CA, ABG, BCG, CAG, ABC and ABCG phase fault. Hence it should be equipped suitably to tackle these faults in the most appropriate manner. Tackling these faults means to classify and finding out its location and graveness. In past on occurrence of fault, current is measured from either ends of the line which were then used in the algorithm to classify them. It could be the raise in magnitude of current magnitude as the data in algorithm. Since each fault react differently i.e. different characteristics of current when it occurs.

**Conflict of Interest:** The corresponding author, on behalf of second author, confirms that there are no conflicts of interest to disclose.

**Copyright:** © 2025 by Ashish Lonare, Anjana Tripathi, Balram Yadav Author(s) retain the copyright of their original work while granting publication rights to the journal.

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License, allowing others to distribute, remix, adapt, and build upon it, even for commercial purposes, with proper attribution. Author(s) are also permitted to post their work in institutional repositories, social media, or other platforms.

**REFERENCES:**

- [1] Yu Liu , Sakis Meliopoulos, “Protection and fault locating method of series compensated lines by wavelet base energy traveling wave” , Power & Energy Society General Meeting, pp 1 – 5, July 2017.
- [2] Yanhui Xi , Zewen, Li “ Fault location based on travelling wave identification using an adaptive extended Kalman filter”, IET Generation, Transmission & Distribution, vol. 12,pp. 1314 -1322, 2018.
- [3] Jana, Soumyadip, Nath, Sudipta, Dasgupta, Aritra, 2012. Transmission line fault classification based on wavelet entropy and neural network. Int. J.Electr. Eng. Technol. 3 (July–September (2)), 94–102.
- [4] Hasabe, R.P., Vaidya, A.P., 2014a. Detection and classification of faults on 220 kv transmission line using wavelet transform and neural network.Int. J. Smart Grid Clean Energy 3 (July (3)), 283–290.
- [5] M. Choudhury, A. Ganguly, "Transmission line fault classification using discrete wavelet

- transform", *Energy Power and Environment: Towards Sustainable Growth (ICEPE) 2015 International Conference on*, pp. 1-5, 2015.
- [6] Jamil, M., Sharma, S.K., Singh, R., "Fault detection and classification in electrical power transmission system using artificial neural network", *SpringerPlus*, 2015, 4, (334), pp. 1–13.
- [7] P. Ray, D. P. Mishra, S. Mohaptra, "Fault classification of a transmission line using wavelet transform & fuzzy logic", *Proc. IEEE 1st Int. Conf. Power Electron. Intell. Control Energy Syst. (ICPEICES)*, pp. 1-6, Jul. 2016.
- [8] P. Ray, D. Prasad, "Application of Wavelet Technique for Fault Classification in Transmission Systems", *Procedia Computer Science Elsevier*, vol. 92, pp. 78-83, 2016.
- [9] S.A. Gafoor, P.V. Ramana Rao, "Wavelet based fault detection classification and location in transmission line", *Power and Energy Conference PECon '06 IEEE International*, vol. 28, pp. 114-118, 2006.
- [10] T. Lobos, P. Kostyla, J. Pospieszna, MJaroszewski, Location of Faults On Transmission Lines Using Wavelet Transforms, *International Conference on High Voltage Engineering and Application*, November 9-13, 2008, pp: 633-636.
- [11] P K Murthy, J Amarnath, S Kamakshiah et al., "Wavelet transform approach for detection and location of fault HVDC system[C]", *Proceedings of 2008 IEEE Region 10 and the Third International Conference on Industrial and Information Systems*, pp. 1-6, December 8-10, 2008.
- [12] B. R. Reddy, M. V. Kumar, M. Suryakalavathi et al., "Fault detection classification and location on transmission lines using wavelet transform", the proceedings of the 2009 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp. 409-411, 2009.
- [13] Zhengyou He, Ling Fu, Sheng Lin, and Zhiqian Bo, "Fault detection and classification in EHV transmission line based on wavelet singular entropy", *IEEE Trans. Power Del.*, vol. 25, no. 4, 2010, pp. 2156-2163.
- [14] A. Yadav, A. Swetapadma, "A novel transmission line relaying scheme for fault detection and classification using wavelet transform and linear discriminant analysis", *Ain Shams Engineering Journal*, 2014.
- [15] R. C. Mishra, P. M. Deoghare, C. Bhale, S. Lanjewar, "Wavelet Based Transmission Line Fault Classification And Location", *IEEE International Conference on Smart Electric Grid (ISEG)*, pp. 1-5.
- [16] Majid Jamil, Rajveer Singh, Sanjeev Kumar Sharma, "Fault identification in electrical power distribution system using combined discrete wavelet transform and fuzzy logic", *Journal of Electrical Systems and Information Technology*, vol. 2, pp. 257-267, 2015.
- [17] N. A. Sundaravaradan , Rounak Meyur , P. Rajaraman , "A wavelet based novel technique for detection and classification of parallel transmission line faults" , *International Conference on Power and Embedded System (SCOPES) – 2016*, Vol. 2, pp. 1951 - 1955 , 2016.
- [18] S. Kirubadevi, S. Sutha, "Wavelet based transmission line fault identification and classification", *International Conference on Computation of Power Energy Information and Communication (ICCPEIC) – 2017*, pp. 737 – 741, 2017.
- [19] J.R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, no. 1, pp. 81-106, 1986.