FAULT DETECTION AND CLASSIFICATION IN TRANSMISSION LINE BY USING DECISION TREE

[1]P. P. Wasnik, [2]Dr. N. J. Phadkule, [3]Dr. K. D. Thakur
[1]PG Scholar, [2]Associate Professor, [3]Assistant Professor
[1]Department of Electrical Engineering,
[1]Government College of Engineering, Amravati, India.

Abstract : The paper presents a novel approach of Machine learning for the protection of power system. The decision tree algorithm based on semi-supervised-machine learning is designed for the automated detection as well as classification of the transmission line faults. The decision tree makes the predictions based on the decisions for constructing optimal-classification tree. To extract the features in the current and voltage signal, discrete wavelet transform is applied, which decomposes the signals into smaller components. This feature vector is used by the classifier for class prediction. The performance of the algorithm is tested on the Three phase series compensated network in MATLAB environment. The results shows that the proposed algorithm can detect and classify the different faults reliably while keeping the computation burden suitably low which make the implementation more feasible.

***IndexTerms* - Fault classification, fault detection, decision Tree, dwt, machine learning.**

I. INTRODUCTION

The transmission lines forms one of the most significant part of the power system as they play a most important role in transmission of power from the generating station to the customer. The normal operation of the transmission line can be affected by many factors and the most sever damages are caused by the faults. In the transmission line with three phase power source, there are eleven types of faults that can occur. The faults in the order of decreasing frequency of occurrence are: single line –to-ground fault (LG), line-to –line fault (LL), double line- to –ground fault (LLG), three line fault (LLL) and three phase –to –ground fault (LLLG). It is then very important to detect these faults; otherwise it will cause disturbances to the system which can further led to substantial outages in the firmly interconnected system working well within the limits.

The current trend of increasing the power transfer over long distance and operating the systems more closer to their limits, makes security more complex and at the same time more complicated to handle. The Machine Learning (ML) is the advanced emerging method in the area of Artificial Intelligence (AI), [1] and [2], now can be used as problem solver in the big data analysis. This encourages the ML to be used for various applications in the various fields. Many researchers are working on the concept of ML for different problems in power system, in [3] and [4] such as economic dispatch, security assessment, Load monitoring and scheduling, voltage stability prediction, fault analysis etc. The fault analysis using ML concept is the most novel approach to obtain the precise analysis and quick recovery of the power system components after the fault.

The semi-supervised learning is the types ML, which combines the power of both the supervised and unsupervised learning. It requires less initial labeled data to train the algorithm and hence more advantageous in the application where large unlabeled data is present and getting labeled data is tedious. The application of such powerful ML technique in the fault prediction could results in improving the protection procedures for the power transmission system. The study [5] and [6] introduces the application of ML for detecting and classifying fault on the transmission line. In addition, it will also reduce the time required to clear the faults, thus increasing the overall power system reliability and efficiency.

The decision tree algorithm provides an interesting alternative to the conventional methods, as presented in study [7] and [8]. The decision tree method, tool of ML, falls into the category of the learning from examples. The underline principle consists of extracting information about a given problem and organizing it into decision trees which are both interpretable and appropriate for automatic use, this enables one subsequently to conclude knowledge about new, unseen situations, whenever they arise. One of the most important steps for designing the fault classifier is the preprocessing step for extracting features. The Discrete Wavelet Transform (DWT) is the signal processing tool, mostly used for unveiling the hidden information from the signal, explained in [9].

The present work introduces the semi-supervised ML approach for detection and classification of transmission line faults. The fault current and fault voltages are decomposed into smaller components by using the DWT, used as feature vector for class prediction. The decision tree algorithm is designed as the fault classifier. Finally, these extracted key features were fed to decision tree algorithm for fast and accurate fault recognition.

II. DISCRETE WAVELET TRANSFORM

The pre-processing of the data is one of the useful techniques to improve the performance and the speed of protection scheme. This can be achieved by unveiling the useful information from the signals, which also reduces the dimensions of the input data. The DWT is one of the preferable signal decomposition tools in the area of power system. The digital representation of the Continuous Wavelet Transform (CWT) is called as DWT. DWT can be realized by multi-stage filter bank having wavelets as low pass (LP) and high pass (HP).

In consideration to the unique feature of providing multiple-resolutions in both time and frequency by wavelets, the sub-band information can be extracted from the original signal. When applied to faults, this sub-band information of faulted power system is seen to provide useful signatures of faults, so that fault classification can be done elegantly. In studies [9], [10] and [11], the

DWT technique is used for detecting the faults on the power system under different operating conditions. The DWT in combination with other algorithms like k nearest neighbor, artificial neural network and decision tree is studies in [12], [13], and [14]

The DWT of any signal say, *a(k)* can be calculated by passing it through the series of filters, which results in convolution of the two. The multi-stage filter bank decomposes the given signal or down sampled by two in each stage and is given as input to subsequent stages, in [15] and [16]. The filter outputs are subsampled by 2, the decomposed signals can be written in mathematical form as shown below. Equation1 and equation 3 shows the output of high pass filter and equation2 and equation4 shows the output of low pass filter.

$$a_1(n) = \sum_k h(k-2n)\, a_0(k) \tag{1}$$

$$d_1(n) = \sum_k g(k-2n)\, a_0(k) \tag{2}$$

At level 2, the next scale decomposition is carried out on a1(n) and the decomposed signals are,

$$a_2(n) = \sum_k h(k-2n)\, a_1(k) \tag{3}$$

$$d_2(n) = \sum_k g(k-2n)\, a_1(k) \tag{4}$$

The proposed scheme of protection employs the DWT for feature extraction accomplished in the following stages. Firstly, the simulation of the test system is performed in order to obtain the instantaneous current and voltage samples at the relay locations at Bus 2. Then in the next stage, the current and voltages obtained in first stage are passed through the series of low pass and high pass filters in order to obtain the approximate and detail coefficients. Third stage involves the amplitude of the four level approximations of current and voltage signals used as input to train the fault classifier.

III. DECISION TREE

The Decision Tree (DT) also called as classification tree is a decision based tool, which uses binary tree like structure to figure out the relationship between the input and the output. The building process of the decision tree depends on initial learning dataset, i.e. a set of samples with predefined label, begins at the top node of the tree with complete learning dataset progresses by repetitively creating descendant nodes by splitting the learning dataset into subsets of increasing classification purity [7], [17] and [19].

The DT is the graphical representation, as shown in figure 1, of the decisions to be made, the case situations and the possible outcomes in relation with the cases and decisions. The decision is made by the Gini index which is the measure of the node impurity [15]. The index defines the splitting criteria for the tree at each node in order to obtain the most accurate guess of the decision.

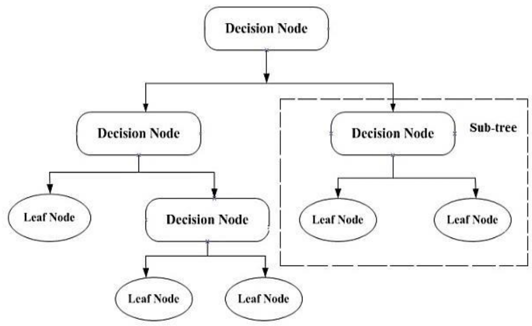

Fig.1. Decision Tree Structure

DT is the data mining classification algorithm used for high-dimension pattern classification [18]. The mathematical representation of the DT algorithm is built on the following definitions,

$$\bar{\boldsymbol{X}}=\{\boldsymbol{X_1}, \boldsymbol{X_2}, \dots\dots\dots.., \boldsymbol{X_m}\}^{\boldsymbol{T}} \tag{5}$$

$$\boldsymbol{X_i}=\{\boldsymbol{x_1}, \boldsymbol{x_2}, \dots\dots, \boldsymbol{x_{ij}}, \dots\dots\dots, \boldsymbol{x_m}\} \tag{6}$$

$$\mathbf{S}=\{\boldsymbol{S_1}, \boldsymbol{S_2}, \dots\dots\dots, \boldsymbol{S_i}, \dots\dots\dots, \boldsymbol{S_m}\}^{\boldsymbol{T}} \tag{7}$$

Where *m* is the number of available observations (cases), the number of independent variables (features), S the *m*-dimensional vector of the categorical (dependent) variable to be predicted from $\bar{X}$, X_i, the *ith* component vector of *n* dimensional independent variables, x_{i1}, x_{i2},….., x_{ij},…..…x_{in} the independent variables (predictors) of the pattern vector X_i and *T* the vector transpose notation. The goal of DT data mining is to predict S based on observing $\bar{X}$. As many DTs with various accuracy levels can be constructed from a given $\bar{X}$, finding the optimal tree is difficult in practice because of the large size of the search space [19].

The DT is applied for the analyzing various problems such as power-quality disturbance [20], solar PV system faults [21], faults on single line [22], faults on double –circuit- transmission-line [23], fault-zone identification [24] and [25].

IV. SYSTEM STUDIED

The Three phase series compensated network simulated in MATLAB environment is used to evaluate the performance of the proposed algorithm. Figure 2 shows the single line diagram of three phase test system used for fault detection and classification of transmission line. The test system is operating at base voltage of 735 kV, with the base power of 100 MW at 60 Hz. The faults of different nature are simulated on the transmission line between Bus B1 and Bus B2, of length 300 km. The fault data required for the classification is obtained from the relay location at Bus 2. The magnitude of current and voltage varies with the location and type of the fault and during each fault type changes in the magnitude are different, hence fault voltage and fault current can be sufficient to be used for detection and classification purpose.

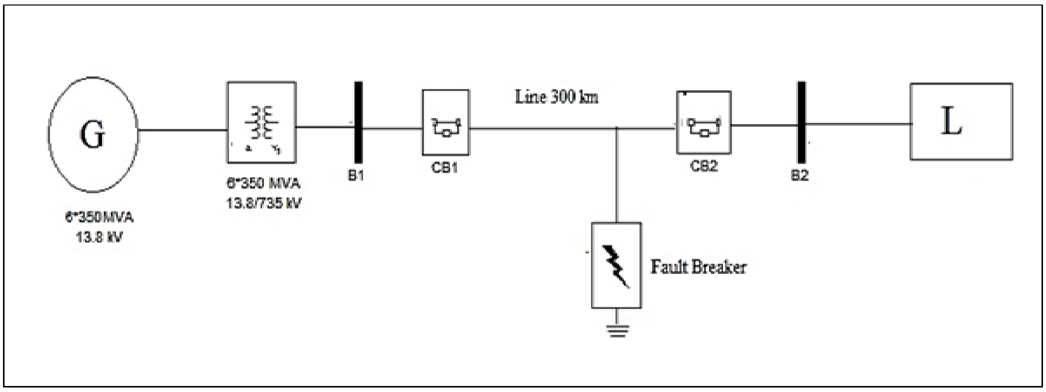

Fig. 2. Three Phase Transmission Line Network for fault detection and classification

The decision tree algorithm works as classifier to detect the abnormal situation and then classify it according to its nature of occurrence. The classifier is initially trained to set the principle of operation. Assume the three phases be A, B, C and G is the ground. The table 1 shows the types of fault and their class number, which includes all the 11 types of fault on three phase system along with the normal operation with the class number 12.

Table 1. Fault types and class number

Fault Type	Class Number
Phase A-to-G	1
Phase B-to-G	2
Phase C-to-G	3
Phase A-B-to-G	4
Phase B-C-to-G	5
Phase A-C-to-G	6
Phase A-B	7
Phase B-C	8
Phase A-C	9
Phase A-B-C	10
Phase A-B-C-to-G	11
Normal Operation	12

The decision tree algorithm generates the generalized model or the tree, during the training and the test sample is classified based on the node conditions [7].

V. RESULT AND DISCUSSION

In order to evaluate the performance of the classifier, 22 the fault cases are considered i.e. all 11 type of faults at two different locations are simulated in MATLAB using fault breaker block along with two cases of normal operation. The various fault conditions are created using different parameters listed in table 2. The result consists of accurately detected and labeled fault cases. The confusion matrix is plotted as shown in figure 3, to evaluate the accuracy of classifier operation. The green block shows the correctly classified fault case. The red blocks shows the incorrectly classified fault cases, the coloums shows the total test cases belonging to that particular class and the row shows the predicted class output. The blue block shows the total classification accuracy of the classifiers.

Table-2: Experimental conditions

Fault Type	AB, BC, AC, AG, BG, CG, ABG, BCG, ACG, ABG and ABCG
Fault Locations	280 and 300 km
Fault Resistance	0.001 and 0.1 Ω

The classifier is tested for 24 different situations including fault situation and normal operataion. The confusion matrix shows the accuracy of the classification for each fault and overall classfication accuracy offered by the classifier. The classifiers offers 100 % accuacy of decision and for all 24 cases 8.3% of overall classification accuracy.

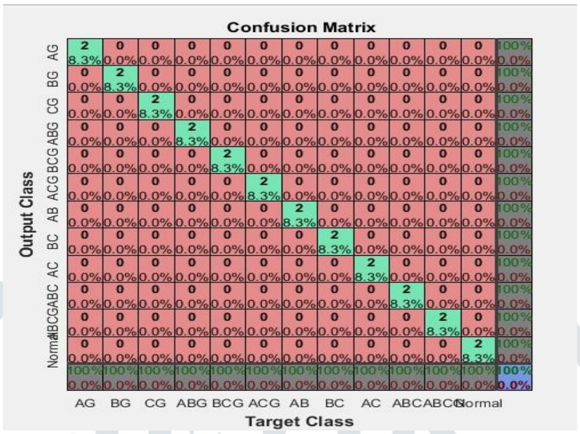

Fig.3. Confusion Matrix

It is evident in the figure 3, that the proposed algorithm is capable of detecting and classifying the fault cases successfully for each of the 24 cases correctly. The result confirms the 100% accuracy of the DT algorithm as fault classifier and also its suitability for real time implementation.

VI. CONCLUSION

The paper presents a decision tree algorithm based on semi-supervised machine learning for fault identification and classification on the transmission line. The goal is to develop the most accurate method of fault analysis. The performance of the algorithm is tested on Three-phase series compensated network for each 11 types of fault. The feature extraction process on the fault current and voltage signal is carried out by using four levels of decomposition in DWT. The feature vectors are then used for class prediction by the fault classifier. The results have shown that, the DT algorithm is able to detect and classify the fault cases accurately and successfully. The need of initial label data required for classifier training is considerably reduced due to the involvement of the semi-supervised learning approach. It is observed that, successfully trained classifier is able to classify any number of fault cases well within the limiting criteria.

VII. ACKNOWLEDGMENT

This paper is submitted under the project and thesis work for the master of technology (Electrical Power System) from Government college of Engineering, Amravati, under the guidance of Dr. N. J. Phadkule, Associate Professor, and Dr. K. D. Thakur, Assistant Professor, Department of Electrical Engineering, Government college of Engineering Amravati.

REFERENCES

[1] Debirupa Hore**.**2018.Application Of Artificial Intelligence And Machine Learning In The Field Of Electrical Engineering: An Overview**.**International Conference on Smart Systems (ICSS).ISBN: 978-93-88350-20-4..

[2] Annina Simon. Mahima Singh Deo. S. Venkatesan and D.R. Ramesh Babu**.** 2015**.** An Overview of Machine Learning and its Applications**.** International Journal of Electrical Sciences & Engineering (IJESE).Volume 1, DOI- 289980169.

[3] Seyed Mahdi Miraftabzadeh. Federica Foiadelli. Michela Longo and Marco Pasetti. 2019. A Survey of Machine Learning Applications for Power System Analytics. IEEE International Conference on Environment and Electrical Engineering and IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe). 10.1109/EEEIC.2019.8783340.

[4] Radu-Adrian Tîrnovan. and Maria Cristea. 2019. Advanced techniques for fault detection and classification in electrical power transmission systems: An overview. 8th International Conference on Modern Power Systems (MPS). 10.1109/ MPS.2019. 8759695.

[5] Halil Alper Tokel. Rana Al Halaseh. Gholamreza Alirezaei. and Rudolf Mathar. 2018. A New Approach for Machine Learning-Based Fault Detection and Classification in Power Systems. IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). 10.1109/ ISGT.2018. 8403343.

[6] Prerana P. Wasnik. Dr. N. J. Phadkule. and Dr. K. D. Thakur. 2019. Fault Detection And Classification Based On Semisupervised Machine Learning Using KNN. in Proc, International Conference on Innovative Trends and Advances in Engineering and Technology (ICITAET).

[7] Chengxi Liua. Zakir Hussain Rathera,b. Zhe Chena. and Claus Leth Bak. 2013. An overview of decision tree applied to power systems. International Journal of Smart Grid and Clean Energy. vol. 2.10.12720/sgce.2.3.413-419.

[8] S. M. Shahrtash. and A. Jamehbozorg. 2008. A Decision Tree Based Method for Fault Classification in Transmission Lines. IEEE/PES Transmission and Distribution Conference and Exposition. 10.1109/ TDC.2008.4517258.

[9] Suman Devi. Nagendra K. Swarnkar. Sheesh Ram Ola. and Om Prakash Mahela. 2016. Detection of Transmission Line Faults Using Discrete Wavelet Transform. Conference on Advances in Signal Processing (CASP). 10.1109/ CASP. 2016. 7746152.

[10] Toshiba Suman. Om Prakash Mahela. and Sheesh Ram Ola. 2016. Detection of Transmission Line Faults in the Presence of Solar PV Generation Using Discrete Wavelet. IEEE 7th Power India International Conference (PIICON). 10.1109/ POWERI. 2016.8077203.

[11] Tapas Kumar Patel. Prafulla Chandra Panda. Sarat Chandra Swain. and Subodh Kumar Mohanty.2017. A Fault Detection Technique in Transmission Line By using Discrete Wavelet Transform. Second International Conference on Electrical, Computer and Communication Technologies (ICECCT). 10.1109/ICECCT.2017.8118052.

[12] M. Gowrishankar. P. Nagaveni. and P. Balakrishnan. 2016. Transmission Line Fault Detection and Classification Using Discrete Wavelet Transform and Artificial Neural Network. Middle-East Journal of Scientific Research. 10.5829/idosi.mejsr.2016.24.04.23063.

[13] M.Pavani. S.Raghunath Sagar. and G.Ganesh. 2018.Fault Classification in Transmission Lines Using Wavelet Transform and K-Nearest Neighbors.International Journal of Education and Applied Research (IJEAR)**.** Vol. 9.

[14] P. Hariramakrishnan. and S. Sendilkumar**.**2018. Data Mining Based Power Quality Disturbance Detection Using Wavelet Transform**.** International Journal of Engineering Technology**.**

[15] Tamer S. Abdelgayed. Walid G. Morsi. and Tarlochan S. Sidhu. 2017. Fault detection and classification based on co-training of semi-supervised machine learning. IEEE Trans. on Industrial, DOI10.1109/THE 2017.2726961.

[16] Subhra Jana. and Abhinandan De.2017. Transmission Line Fault Pattern Recognition using Decision Tree based Smart Fault Classifier in a Large Power Network. IEEE Calcutta Conference (CALCON), 978-1-5386-3745-6/17.

[17] L.Wehenke. and M. Pavella. 1993.Decision tree approach to power systems security assessment. International Journal of Electrical Power & Energy Systems.Volume 15.Issue 1. Pages 13-36.10.1016/0142-0615(93)90014-E.

[18] Prakash K. Ray. Soumya R. Mohanty. Nand Kishor. and João P. S. Catalão. 2013. Optimal Feature and Decision Tree-Based Classification of Power Quality Disturbances in Distributed Generation Systems. IEEE Transactions On Sustainable Energy. 10.1109/TSTE. 2013.2278865.

[19] Urmil B. Parikh. Bhavesh R. Bhalja Jr. Rudra Prakash Maheshwari. and Biswarup Das. 2007. Decision Tree Based Fault Classification Scheme for Protection of Series Compensated Transmission Lines. International Journal of Emerging Electric Power Systems.Volume 8.Issue 6.10.2202/1553-779X.1630.

[20] J. Upendar. C.P. Gupta. and G.K. Singh. 2012. Statistical decision-tree based fault classification scheme for protection of power transmission lines. International Journal of Electrical Power & Energy Systems.Volume 36.Issue 1. 10.1016/j.ijepes.2011.08.005.

[21] Ye Zhao. Ling Yang. Brad Lehman . Jean-François de Palma.Jerry Mosesian. and Robert Lyons. 2012.Decision Tree-Based Fault Detection and Classification in Solar Photovoltaic Arrays. Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC).10. 1109/APEC.2012. 6165803.

[22] Arash Jamehbozorg. and S. Mohammad Shahrtash.2010.A Decision-Tree-Based Method for Fault Classification in Single Circuit Transmission Lines. IEEE Transactions On Power Delivery. Vol. 25. No. 4. 10.1109/Tpwrd.2010.2053222.

[23] A. Jamehbozorg. and S. M. Shahrtash. 2010. A Decision Tree-Based Method for Fault Classification in Double-Circuit Transmission Lines. IEEE Transactions on Power Delivery. Volume 25. 10.1109/ TPWRD.2010. 2050911.

[24] Prasad Ranjan Ghosh. Srikant Mohapatra. Arjyadhara Pradhan. and Soubhagya Prusty. 2020. Fault Classification and Fault Zone Identification in a Thyristor Controlled Series Compensator Based Transmission Line by using Decision Tree. International Journal of Innovative Technology and Exploring Engineering (IJITEE).Volume-9 Issue-3.10.35940/ijitee..

[25] S.R. Samantaray. 2009. Decision tree-based fault zone identification and fault classification in flexible AC transmissions-based transmission line. IET Generation, Transmission & Distribution. 10.1049/iet-gtd.2008.0316. Vol. 3. Iss. 5.pp.

[26] Rajesh Kumar. Om Prakash Mahela. and Mahendra Kumar.2019.A Current Based Algorithm Using Harmonic Wavelet Transform and Rule Based Decision Tree for Transmission Line Protection. 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU). 978-1-7281-1253-4.

[27] S.R. Samantaray. I. Kamwa.and Geza Joos.2011. Decision tree based fault detection and classification in distance relaying. Engineering Intelligent Systems for Electrical Engineering and Communications. vol 18 no 2.

[28] Subodh Kumar Mohanty. Anshudip Karn. and Shobhan Banerjee. 2020. Decision Tree Supported Distance relay for Fault Detection and Classification in a series compensated line. IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020).10.1109/PESGRE45664.2020.9070724.