

# POWER SYSTEM FAULT DIAGNOSIS USING MACHINE LEARNING

## Abstract:

Power Systems have many challenges which include fault diagnosis, load frequency control, unit commitment, load scheduling, optimization etc. In the above-mentioned, fault diagnosis is one of the major issue. This can be resolved by using traditional and artificial intelligence-based techniques. This paper focus on fault detection, classification and location identification in electrical transmission systems using machine learning techniques. The simulation results concluded that the present method is efficient in detection, classification and location estimation of the faults on the transmission lines with satisfactory performance.

**Keywords:** Fault Diagnosis, Artificial Intelligence, Deep Neural Network, Recurrent Neural Network.

## I. Introduction:

Electrical power systems are important assets of every nation, We mostly depend on the electrical power. The electrical power systems were grown very rapidly in the past few decades that resulted in increase in the number of lines in operation and their length. The transmission lines are exposed to open environment so that the faults are unavoidable. These faults occurs in a result of lightning, short circuits, faulty equipment, maloperation, human errors, overloading and ageing etc.

When a fault occurs in transmission lines it is very important to detect, classify and to find the fault location to restore the power delivery. The time needed for the restoration of the power will reflect in power quality. Therefore a sophisticated detection technique and an accurate location on the line is an important requirement for fault detection.

Most of the faults can cause large currents or voltage changes so that they can be detected using the traditional protective relays. But some faults, such as high impedance faults will cause small current and voltage changes. So that it is difficult to detect by a traditional protective relay. For such problems, we need an efficient fault detection, classification and location identification method [1].

## II. Power System Faults:

A fault is an abnormal condition in electrical systems. The faults in the electric transmission lines are short circuit faults and open circuit faults etc. Open circuit faults are very rare in the transmission lines but the short circuit faults are very common these faults are may be due to natural climatic conditions and mis-operation. The transmission of electric power is being done using 3 phase lines. The short circuit faults in the 3 phase transmission lines are classified as symmetrical faults and unsymmetrical faults [2].

### A. Symmetrical faults:

Symmetrical faults are most severe faults and rarely occurs in the power system. These faults are balanced. These faults are of two types LLL fault and LLL-G faults, When ground involves in the fault then that is called as LLL-G fault else called as LLL fault. These faults remain balanced in the system. The analysis can be done by using per phase analysis.

### *B. Un-symmetrical faults:*

These faults are very common and less severe than the symmetrical faults. These faults are classified as the line to ground (L-G), line to line (L-L), double line to ground (LL-G) faults. These faults are unbalanced and cause unbalanced currents to flow in the phases. The study of un-symmetrical faults can be done by using symmetrical components.

## **III. Fault Detection and Classification Techniques:**

There are plenty of techniques proposed over the past years. Those techniques have their own advantages and disadvantages. The fault classification and location identification must be very fast to improve power quality.

### *A. Discrete Wavelet Transform:*

Discrete wavelet transform (DWT) is a very important technique for the feature extraction from certain frequency bands in signals. Discrete wavelet transforms with Multi-Resolution Analysis (MRA) can be used to analyze the high-frequency signals for a short duration. The main drawbacks of DWT is the choice of appropriate mother function suitable for the application, computational complexity and time etc [1][3][4].

### *B. Artificial Neural Networks:*

Artificial neural networks are (ANN) are a family of non-linear statistical models and learning algorithms that are intended to imitate the behaviour of connected neurons in biological neural systems. Different ANN models have been used for different applications. Feedforward neural network (FNN) is the simplest neural network configuration which can be characterized as a single layer or multi-layer perceptrons. An FNN often has an input layer, output layer and at least one hidden layer. The node or neurons will fully be connected with adjacent layer to process the data. The weights will be assigned and the bias for the nodes decides the output of the network, for a given input.

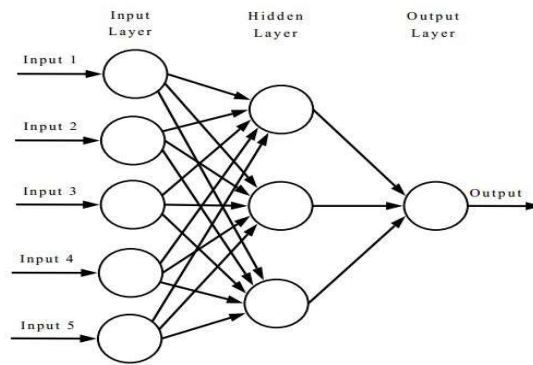


Fig. 1 Artificial neural network

Fig. 1 shows the artificial neural network with multi-input and a single output. From the late 1980 s, researchers are using the Back-propagation algorithm with FNN. There are several types of FNN networks such as Radial basis function networks (RBFN), Probabilistic neural networks (PNN) etc [4][5][6][7].

### *C. Support Vector Machines:*

Support vector machine (SVM) was invented by Cortes and Vapnik in 1995. The main idea of SVM classifiers is to find the optimal hyperplane that maximizes the margin between

two groups of examples. SVM uses non-linear kernel functions to map the examples into higher dimensions. SVM prevents overfitting due to its structural risk-minimizing nature. SVM is a very powerful tool for classification problems. SVM with other techniques were also implemented such as DWT, ST etc. Even though SVM gives better results, it has a problem with parameter optimization [8][9].

#### *D. Decision Trees:*

Decision Trees (DT's) refers to the class of tree-like graphs capable of decision making. DTs will look like trees models with nodes. Concretely, three types of nodes found in a DT, namely root node, internal nodes and leaf nodes. Decision-making starts from the root node and the flow goes along the path that satisfies the test conditions. Decision trees can be trained with many algorithms such as greedy algorithm, random forest etc. Decision trees are easy to understand by humans as per the conditions, but the main drawback is their stability [10][11].

#### *E. Fault Location Identification Techniques:*

The accurate location of faults in the transmission lines and distribution system greatly reduces the time to restore the power. The conventional fault location methods can be classified into two groups, travelling wave based schemes and impedance measurement based schemes [12].

##### *a) Travelling Wave Based Fault Locators:*

Travelling waves will get generated due to switching operations and faults such as short circuit faults and open circuit faults. Travelling waves phenomenon for fault location is classified into four different types. Two of them are due to generated wave analysis and the remaining two are external wave injection to the transmission line at a single end and both ends. The time of reflection of the wave is proportional to the fault location [13][14].

##### *b) Impedance Measurement Based Fault Locators:*

These schemes provide another alternative for the fault location estimation problem. Let us consider that a single line to ground fault occurred in a transmission line with a fault resistance at a distance  $x$  from the sending end. The fault will draw the fault current based on the fault resistance. The measurements units which are placed at sending end and receiving end will measure currents and voltages for double end algorithms, for single-end algorithms measurement will be done at sending end only [13][15].

##### *c) Other Fault Locators:*

The above-mentioned techniques are normal mathematical derivations. Instead of conventional fault location techniques, soft computing techniques such as wavelet transform, artificial neural networks or genetic algorithms were also introduced. But these methods have their problems that result from the line modelling accuracy, data availability [16].

#### *F. Performance Metrics:*

There are several performance measures we can consider to estimate the performance of the network such as Mean Square Error (MSE), Mean Absolute Error (MAE), Correlation Coefficient etc.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Let us consider that we have n number of predictions generated,  $Y_i$  is the predicted output and  $\hat{Y}_i$  is the actual output. Mean square error is the average squared difference between the estimated values and the actual value. Mean absolute error is the average absolute difference between estimated output and actual output.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

Correlation Coefficient denoted by 'r' represents, how closely the predicted output is matched with actual output. The closer the absolute value of r to one, the better that the data are described by a linear equation.

#### IV. Deep Neural Networks:

In recent days, the computational power of computers increased very much and the cost of computing is reduced. Deep learning is a subset of machine learning. The usage of machine learning and deep learning algorithms increased due to computational power and data availability. Deep neural networks are the improved version of the artificial neural networks. These networks are similar to the feedforward neural networks with multiple hidden layers between the input and output layers. The requirement of data for deep neural networks is more [17]. Fig.2 shows a simple deep neural network with one input layer, one output layer and three hidden layers.

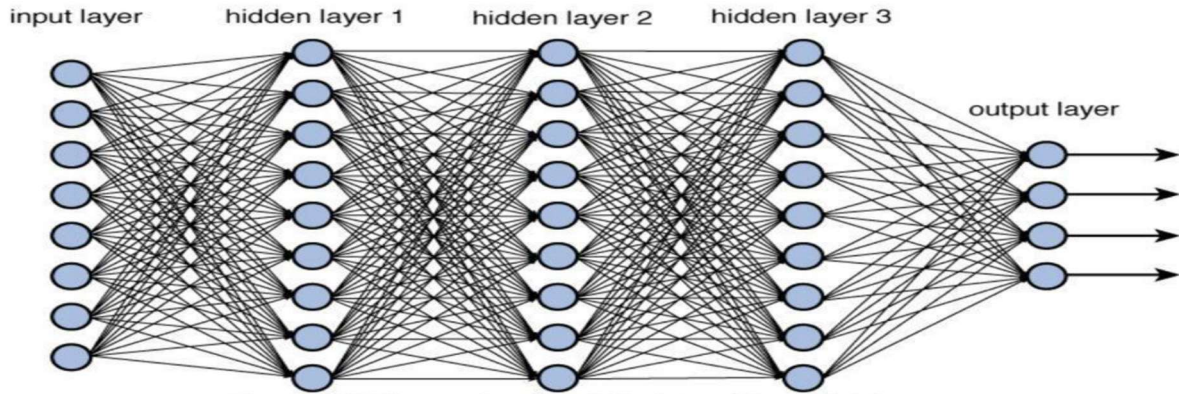


Fig.2 Deep Neural Network

There are several deep neural networks which can do a better job in some specific applications. Deep neural networks can extract features automatically without the help of other feature extraction techniques. In this paper, we mainly focus on recurrent neural networks. Recurrent neural networks (RNN) have many architectures such as fully recurrent, partially recurrent, long short term memory (LSTM) recurrent neural networks etc.

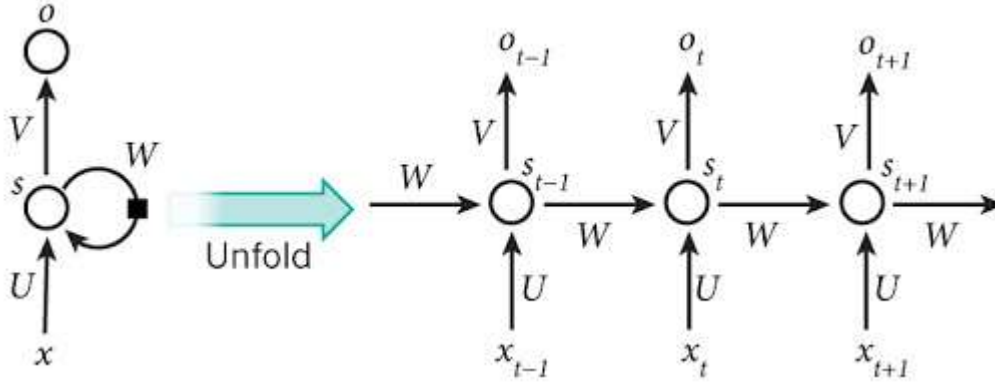


Fig.3 Recurrent Neural Network and the unfolding in time

The recurrent neural networks have feed-back connections to the network itself. RNN's are extremely powerful in modelling sequential data and non-sequential data. RNN can also be used in image processing, video processing, speech recognition, natural language processing etc. The partially recurrent structure adds a feedforward connection, through a synapse, from the input axon to the layer after the 1st hidden layer. In this case, the recurrent structure acts as a state for the feedforward structure [18].

## V. Simulation and Results:

A simple three-phase system is studied in this paper as shown in fig.4. The length of the transmission line is 200 km, the system frequency is 50 Hz and the line voltage is 220 kV. The transmission line connects two sources and has positive sequence impedance of  $Z_1 = 4.76 + j59.75 \Omega$  and zero sequence impedance  $Z_0 = 77.70 + j204.26 \Omega$ .

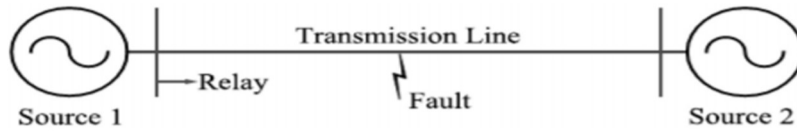


Fig.4. The studied system with sources at both sides

The system is modelled in MATLAB/SIMULINK, with which the data used in this paper is simulated [19]. Three-phase voltage and current signals are collected by the relay employed at source 1. The fault locations are 0 km, 50 km, 100 km, 150 km, 200 km. The fault resistance is 0.01  $\Omega$ , 5  $\Omega$ , 10  $\Omega$ , 15  $\Omega$ , 20  $\Omega$ . The simulated faults are a-g, b-g, c-g, ab, bc, ac, ab-g, bc-g, ac-g, abc, no-fault situations. The collected samples had been used for the training, cross-validation and testing. The 80% of data had been used for the training, 5% of data for cross-validation and remaining 15% data for testing the network.

We had used a partially recurrent neural network for the detection and classification another network for the fault location estimation. We had used NeuroSolutions software for the training and testing of the network. The network consists of one input layer, one output layer and 5 hidden layers for the classification problem. For fault location estimation also we had used the same network parameters. The NeuroSolutions software has a default function approximation network. We can also use that network for the fault location estimation.

While training, the network processing elements in each hidden layer are affecting training time and the minimum number of epochs needed for better results.

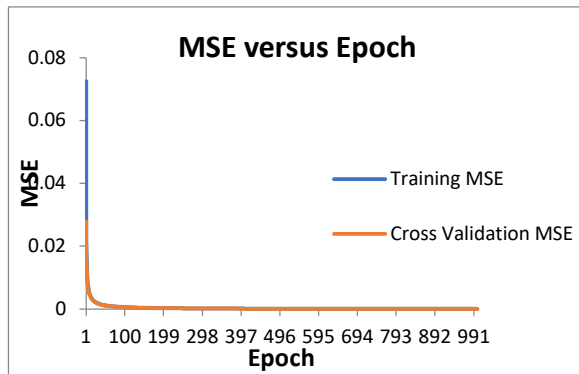


Fig.3 Training curve for the fault classification

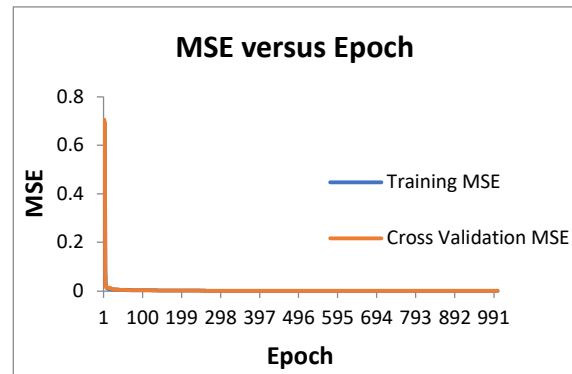


Fig.4 Training curve for the fault location estimation

Table .1 Performance table for fault classification

<i>Performance</i>	<i>No-Fault</i>	<i>AG</i>	<i>BG</i>	<i>CG</i>	<i>AC</i>	<i>BC</i>	<i>CA</i>	<i>ABG</i>	<i>BCG</i>	<i>CAG</i>	<i>ABC</i>
MSE	3.7658 3E-06	0.0002 58642	8.8098 4E-06	5.3432 8E-06	3.7774 4E-06	6.2246 9E-06	3.4831 6E-06	4.5409 E-06	1.2794 4E-05	4.4779 E-06	2.7368 9E-06
MAE	0.0013 21413	0.0064 29595	0.0019 57132	0.0013 64944	0.0013 24566	0.0017 31786	0.0013 58941	0.0014 98223	0.0021 51007	0.0013 34154	0.0011 24385
r	0.9999 514	0.9986 52115	0.9999 53224	0.9999 70318	0.9999 79301	0.9999 64841	0.9999 80127	0.9999 73549	0.9999 85429	0.9999 71576	0.9999 84784
Percent Correct	100	100	100	100	100	100	100	100	100	100	100

Table .1 shows the performance of the classification network with testing data. Classification accuracy of classification network is 100%. Table. 2 shows the performance of the fault location estimation network, it shows mean absolute error as 1.2764 km.

Table .2 Fault location estimation results

<i>Performance</i>	<i>Function Approximation Network</i>	<i>Partially Recurrent Neural Network</i>
MSE	13.83166315	3.158307138
MAE	2.763306481	1.276490649
r	0.99857872	0.99969047

## VI. Conclusions and Future Scope:

The power system fault diagnosis is a very large area to study. The drawbacks of traditional fault classification and fault location estimation techniques can be overcome by using deep learning techniques. Partially recurrent neural networks are classifying with satisfactory accuracy. The fault location estimation is also good with partially recurrent neural network when compared to the function approximation network.

The future scope of this paper is to analyze the system performance by varying the fault inception angle and the value of fault resistance.

## REFERENCES

- [1] K. Chen, C. Huang, and J. L. He, "Fault detection, classification and location for transmission lines and distribution systems: A review on the methods," *High Voltage*, vol. 1, no. 1, pp. 25–33, 2016.
- [2] Ed Richard C. Dorf Boca Roton, *The electrical engineering handbook*, CRC Press LLC, 2000.
- [3] Valsan SP, Swarup KS. Fault detection and classification logic for transmission line using multi resolution wavelet analysis. *Electric Power Comp Syst* 2009; 36(4):321–44.
- [4] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines based on wavelet transform and ANN," *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 2058–2063, Oct. 2006.
- [5] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, "Artificial neural network and support vector machine approach for locating faults in radial distribution systems," *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 710–721, Apr. 2005.
- [6] R. N. Mahanty and P. B. D. Gupta, "Application of RBF neural network to fault classification and location in transmission lines," *IEEE Proc. Gener. Transm. Distrib.*, vol. 151, no. 2, pp. 201–212, Mar. 2004.
- [7] J. Upendar, C. P. Gupta, and G. K. Singh, "Discrete wavelet transform and probabilistic neural network based algorithm for classification of fault on transmission systems," in *Proc. Annu. IEEE India Conf. (INDICON)*, vol. 1. Kanpur, India, 2008, pp. 206–211.
- [8] P. K. Dash, S. R. Samantaray, and G. Panda, "Fault classification and section identification of an advanced series-compensated transmission line using support vector machine," *IEEE Trans. Power Del.*, vol. 22, no. 1, pp. 67–73, Jan. 2007.
- [9] U. B. Parikh, B. Das, and R. Maheshwari, "Fault classification technique for series compensated transmission line using support vector machine," *Int. J. Elect. Power Energy Syst.*, vol. 32, no. 6, pp. 629–636, 2010.
- [10] A. Jamehbozorg and S. M. Shahrtash, "A decision tree-based method for fault classification in double-circuit transmission lines," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2184–2189, Oct. 2010.
- [11] M. K. Jena, L. N. Tripathy, and S. R. Samantray, "Intelligent relaying of UPFC based transmission lines using decision tree," in *Proc. 1st Int. Conf. Emerg. Trends Appl. Comput. Sci. (ICETACS)*, Shillong, India, 2013, pp. 224–229.
- [12] A.A. Girgis, C.M. Fallon, D.L. Lubkeman, A fault location technique for rural distribution feeders. *IEEE Trans. Ind. Appl.* 29(6), 1170–1175, 1993.
- [13] IEEE guide for determining fault location on ac transmission and distribution lines. *IEEE Std C37.114-2004*, 2005, pp: 1–36.
- [14] H.W. Dommel, J.M. Michels, High speed relaying using traveling wave transient analysis, *IEEE PES Winter Power Meeting* (1978)
- [15] R.H. Salim, K.C.O. Salim, A.S. Bretas, Further improvements on impedance-based fault location for power distribution systems. *IET Gener. Trans. Distrib.* 5(4), 467–478 (2011)
- [16] M.M Saha, J. Izykowski, E. Rosolowski, *Fault location on power networks*, Springer, 2010.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [18] NeuroDimension, Inc, "NeuroSolutions", vol. Version 5.07, 2005.

- [19] K. Chen, J. Hu, J. He, "Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder", *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1, Aug. 2016.