Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, No.3, February : 2025

A DATA DRIVEN APPROACH FOR PREDICTING TRANSFORMER FAULTSFOR POWER SYSTEMS

Raksha Dehariya, Research Scholar, Department of Electrical and Electronics Engineering, TIT& Science, Bhopal, India.
Saurabh Gupta, Associate Professor, Department of Electrical and Electronics Engineering, TIT& Science, Bhopal, India.
Devendra Sharma, Assistant ProfessorDepartment of Electrical and Electronics Engineering, TIT& Science, Bhopal, India.

ABSTRACT:
The demand for a reliable supply of electrical energy for the exigency of modern world in each and every field has increased considerably requiring nearly a no-fault operation of power systems. The crucial objective is to mitigate the frequency and duration of unwanted outages related to power transformer puts a high pointed demand on power transformer protective relays to operate immaculately and capriciously. The high pointed demand includes the requirements of dependability associated with no false tripping, and operating speed with short fault detection and clearing time. Timely maintenance of the distribution components, specifically transformers is insufficient for absolute reliability. Also, faults cause temporary interruption in power supply during the time of repair and replacement. The best solution to avoid this involves predicting the temporal probability of faults, heuristically. This paper discusses the prediction of faults on transformers using artificial neural networks.The raining algorithm used is the Scaled Conjugate Gradient and KNN Classifer.
Keywords: Transformer Fault Location,Artificial Neural Networks, Scaled ConjugateGradient, KNN, and Accuracy.

INTRODUCTION:
Optimal power system maintenance schemes are reliability centric. But in typical Indian power system scenarios, these schemes are idealistic and potentially more expensive than the net profit that may be obtained using those schemes. As a result, sub-optimal maintenance schemes are forced to be carried out. The schemes used for this purpose are predominantlymaintenance-upon-breakdown, where a faulty component is taken offline after a fault; and pre emptive maintenance where loose schedules are followed for maintenance whether or not the component has broken down. The latter is aimed at reducing system downtime.
The crucial objective to mitigate the frequency and duration of unwanted outages related to power transformer puts a high pointed demand on power transformer protective relays to operate immaculately and capriciously. The high pointed demand includes the requirements of dependability associated with no false tripping, and operating speed with short fault detection and clearing time. Protection of large power transformers is a very challenging problem in power system relaying. [2]The protective systemincludes devices that recognize the existence of a fault, indicates its location and class, detect some other abnormal fault like operating conditions and starts the inceptive steps of opening of circuit breakers to disconnect the faulty equipment of the power system. There are problems which are peculiar to transformer, which are not encountered in other items of power system. One of the major problem is the large magnetizing inrush current, whose magnitude can be as high as internal fault current and may cause false tripping of the breaker. [5] A common differential relay operating on the basis of measurement and evaluation of currents at both sides of the transformer can't avoid the trip signal during inrush condition. Since the transformer inrush current is rich in second harmonic

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, No.3, February : 2025

component therefore to avoid the needless trip by inrush current harmonic restraint logic together withdifferential logic is used in most of the faultdetection algorithm in the digital differential protection of power transformer. These methods utilize the fact that the ratio of the second harmonic to fundamental component of differential current under inrush conditions is greater in comparison to that under fault conditions. Mechanical forces build up under large inrush current condition within the transformer coils compared to those occurring at short circuit which is the reason for damage of large power transformer. Large inrush currents also affect the power quality by adding harmonics. Also the presence of large quantity of harmonics in the inrush current can cause damage to power factor correction capacitor by exciting resonant overvoltage. Hence steps are taken to mitigate the transformer inrush current by controlled switching and use of low loss amorphous core materials in modern power transformer that produce inrush current with lowsecond and fifth harmonic contents.

SYSTEM DESIGN:

Previous Faults are caused on power system components primarily when they are operated at voltages and currents higher than their rated values. Designers provide sufficiently large margins for overload, typically upto 20% fortransformers. However, the ageing Indian infrastructure is usually not upgraded at the rate at which consumption is increasing andtherefore is usually operated at the edge of the safe operating limits of their components which is typically just below the 20% derating. In day to day operation, in a grid like India, where load curve shaping is not done, it leads to operation of power system components beyond safe limits for long hours. Such long operation timesincrease the likelihood of faults[9- 11] Another case may be when the transformers in the systemare being operated at the edge of derated limits.

Load fluctuations on the consumer side as well as supply fluctuations on the transmission side are sufficient to cause mild faults like inter winding shorts and core flux overload over shortperiods of time thereby reducing the life. Also, such repetitive small faults tend to have a cumulative effect on system reliability[12, 13]. In order to predict the occurrence of faults, instantaneous power transfer by each transformer must be monitored for spikes and for continuous overload. Instantaneous load data for all transformers is to be updated once every fifteen minutes and the data is given to a trained ANN for classification as a time stepped series. Artificial neural networks are a practical way of implementing artificial intelligence with an aim to solve fitting problems generally needing herculean efforts due to the data size and complexity.

The artificial neural architecture tries to imitate the human thought process in the following ways:

- Process data as a parallel stream independently
- Identifying patterns and correlating them.
- Evolving and updating the experiences (called weights) as per the changes in the data received. Neural networks workon training and testing mechanism.
- Finally rendering an output and further storing it as an experience.

ANN basically tries to inherit this capability of human brain to self-train itself for tasks which are never been performed by it that too very efficiently. Human brain's structure consists of neurons which are interconnected with each other and there by forming a very large network which is well connected thereby helps in performing very complex task like voice and image recognition very easily. The same task when performed using normal computer won't give accurate result. Hence ANN mimics neurons structure of human brain to discover link between input and targets. Neurons havethis ability to save previous experimental data. The speed of human brain is several thousand times faster than traditional computer because inbrain unlike traditional computer as whole information is not passed from neuron to neuron they are rather encoded in the neuron network. This is reason why neural network is also namedas connectionism.

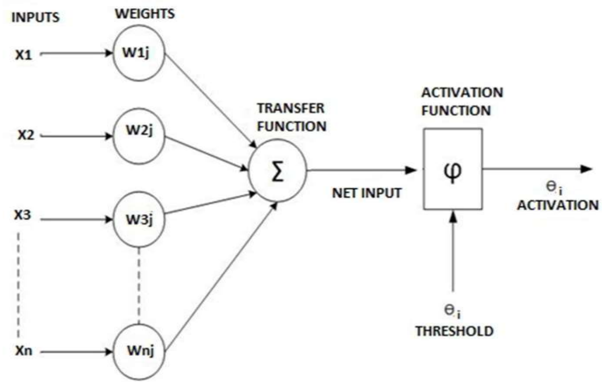

Fig 1: Mathematical Model of ANN

The mathematical conversion of the ANN can bedone by analyzing the biological structure of ANN. In the above example, the enunciated properties of the ANN that have beenemphasized upon are:

1. Strength to process information in parallelway.
2. The power to grasp and learn from weights
3. Searching for patterned sets in complexmodels of data.

Consider a signal s_1 travelling through a path p_1 from dendrites with weight w_1 to the neuron. Then the value of signal reaching the neuronwill be $s_1 \cdot w_1$ .If there are "n" such signals travelling through n different paths with weights ranging from w_1 to w_n and the neuron has an internal firing threshold value of θ_n , then the total activation function of the neuron is given by:

$$Y = \sum_{i=1}^{n} X_i \cdot W_i + \theta_i \quad (1)$$

X_i represents the signals arriving through various paths,
W_i represents the weight corresponding to the various paths and
θ is the bias.

The entire mathematical model of the neuron or the neural network can be visualized pictorially or the pictorial model can be mathematically modeled. The design of the neural network can be modeled mathematically and the more complex the neural design, more is the complexity of the tasks that can be accomplishedby the neural network. The soul of the above model lies in the fact that the system so developed tries to mimic the working of human brain in terms of the following:

- It works in a complex parallelcomputation manner
- High speed of performance due to theparallel architecture.
- It learning and adapt according to themodified link weights.

PROPOSED METHODOLOGY:

Scaled Conjugate Gradient Algorithm

The major advantages of the scaled conjugate gradient algorithm are:

1) Low space complexity
2) Relatively low time complexity suited to real time applications.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 2, No.3, February : 2025

Conjugate Gradient (CG) algorithm is a modified version of steepest descent algorithm. In CG, a hunt is done in such a direction so as to generate a faster convergence than the steepest decent direction, while saving the errorminimization attained in all previous steps. The direction in which CG moves to update weights and bias is called the conjugate direction. In linear CG algorithms, the step-size (total data processed per step) is updated in each iteration. At the start of the algorithm, the direction taken in to account is the direction of the steepest descent [20]. This is done only till the first iteration is complete.

$$p_0 = -g_0 \qquad (2)$$

The weights are updated as follows

$$x_{k+1} = x_k + \alpha_k g_k \qquad (3)$$

Where, α_k is the determined step size and

$$P_k = -g_k + \beta_k P_{k-1} \qquad (4)$$

Here p is search direction vector and g is gradient direction vector.

There are various versions of Conjugate Gradient algorithms which can be categorized by the manner in which the factor $\beta_k$is calculated. In this study, we have used Scaled Conjugate Gradient (SCG) Algorithm, in which we use LM algorithm combined with CG algorithm to calculate step size, unlike only the line search technique in the CG approach.

For SCG, β_k factor calculation and direction of the new search can be shown as in following equations.

$$\beta_K = \frac{\left(| g_{k+1} |^2 - g_{k+1}{}^T g_k\right)}{g_k^T g_k} \qquad (5)$$

$$P_{k+1} = -g_{k+1} + \beta_k P_k \qquad (6)$$

Design parameters are updated at each iteration user independently, which is crucial for the success of the algorithm. This is a majoradvantage compared to the line search based algorithms.

THE KNN APPROACH

KNN stands for the K-Nearest Neighbor. KNN can be used for both classification and regression predictive problems. However, it is more widely used in classification problems in the industry. To evaluate any technique we generally look at 3 important aspects:

1. Ease to interpret output
2. Calculation time
3. Predictive Power

The KNN works on the principle of finding the Euclidean distance of the present data sample to a particular class. The steps involved in KNN are:

1) Receive an unclassified data;
2) Measure the distance from the new data to all others data that is already classified;
3) Gets the K smaller distances;
4) Check the list of classes had the shortest distance and count the amount of each class that appears;
5) Takes as correct class the class that appeared the most times;
6) Classifies the new data with the class that wasdone in step 5.

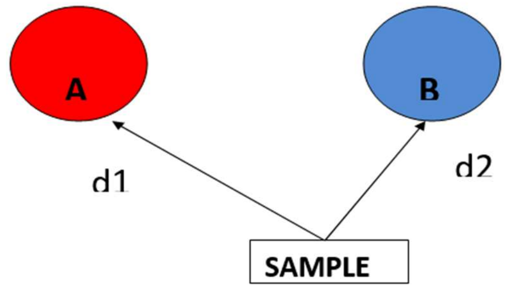

Fig 2: KNN Approach

The KNN evaluated the distance d1 and d2 and classifies the sample to belong to a dataset basedon

$$z = \min(d1, d2) \quad (7)$$

The proposed approach uses an **ada-boost** approach for android malware detection. In this approach, the output of one neural network isfed as the input to another neural network. The characteristic of such an approach is the fact thatit can achieve higher effectiveness of classification accuracy compared to a single neural architecture for classification since the parameters which distinguish malwares and non- malwares are very similar and often makes the classification accuracy plummet.

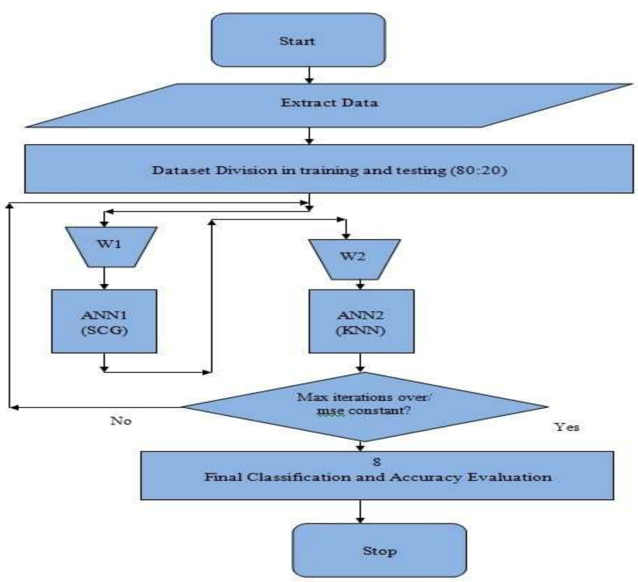

Fig 2: Flowchart for proposed approach

EVALUATION PARAMETERS:

Often an artificial intelligence based approach is used for the classification and renders the following issues.[10]

1. **True Positive (TP):** It is indicative of the true or correct cases of the data to be in a particular class.
2. **True Negative (TN):** It is indicative of the true or correct cases of the data not to be in a particular class.
3. **False Positive (FP):** It is indicative of the false or incorrect cases of the data to be in a particular class.

4. **False Negative (FN):** It is indicative of the false or incorrect cases of the data not to be in a particular class.

Sensitivity (Se): It is indicative the ratio in which a data set is categorized Mathematically itcan be defined as:

$$S_e = \frac{TP}{TP + F} \qquad \textbf{(8)}$$

Accuracy (Ac): It is an indicative of the accuracy of classification of the algorithm for data classification, Mathematically its defined as:

$$A_C = \frac{TP + TN}{TP + TN + FP + FN} \qquad \textbf{(9)}$$

RESULTS

The system is designed on Matlab.
The training algorithm used is the scaled conjugate gradient (SCG) approach. The finalclassification is done based on the KNN approach. The confusion a=matrix renders the average error and accuracy.

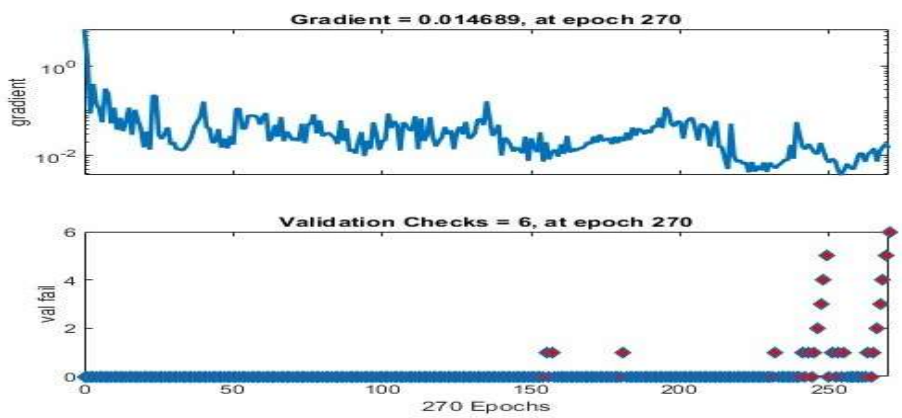

Fig 3: Variation of gradient and validationswith respect to epochs.

The figure above depicts the variation of the gradient and validations as per the variations in the number of iterations.

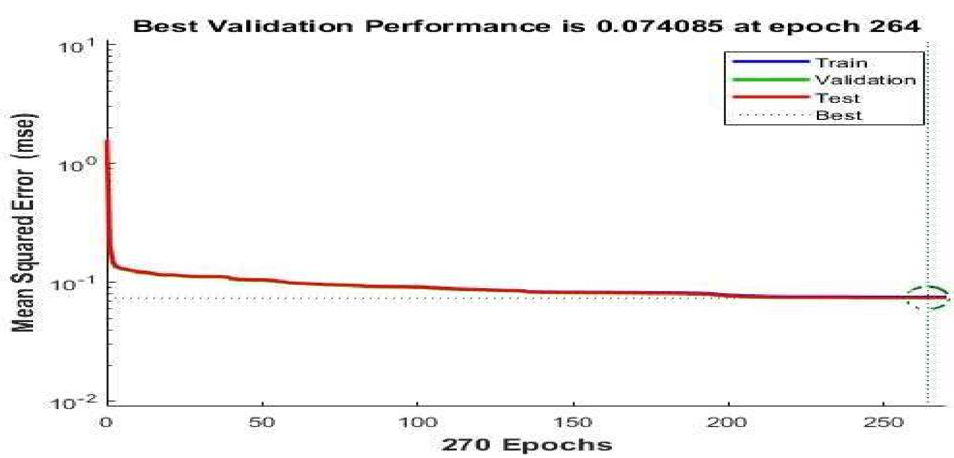

Fig 4: Variation of gradient and validationswith respect to epochs.

The figure above depicts the variation of themean square error as per the variations in the number of iterations.

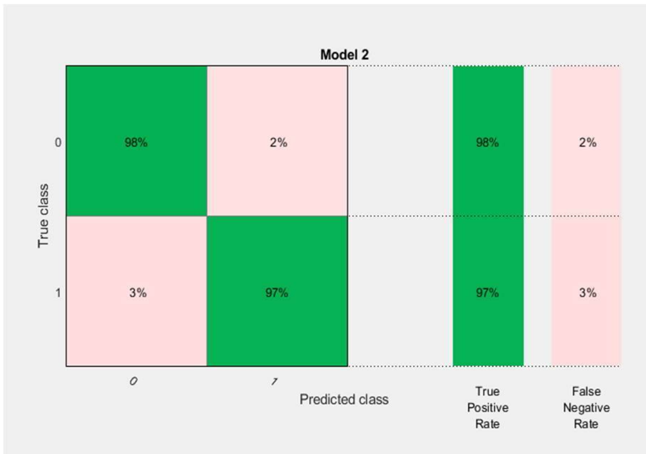

Fig 5: Confusion matrix for KNN approach.

The figure above depicts the confusion matrix for the used data set.

CONCLUSION:

From the present study, we have concluded that the signals for different cases for a power transformer are to be analyzed using wavelet transform for extraction of feature vector (containing statistical data) to train the ANN. The performance of trained ANN is successful for the classification of various cases. From the study and analysis carried out in this dissertation, the performance of neural networks has been found to surpass the performance of conventional methods, which need accuratesensing devices, costly equipment and an expert operator or engineer. It has been found that the proposed system achieves worst case accuracy of 98% which is higher than previously existing techniques.

REFERENCES:

[1] I. B. M. Taha, S. Ibrahim and D. -E. A. Mansour, "Power Transformer Fault Diagnosis Based on DGA Using a Convolutional Neural Network With Noisein Measurements," in IEEE Access, vol. 9, pp. 111162-111170, 2021

[2] Preeti, Mrs. Sweety Sharma “Review on Transformer Protection through Artificial Neural Network”, International Journal of All Research Education and Scientific Methods (IJARESM),Volume 4, Issue 9, September- 2021

[3] AAP Bíscaro, RAF Pereira, M Kezunovic,“Integrated fault location and power-quality analysis in electric power distribution systems”, IEEE 2020

[4] E. Abdallah A. Helal “Review of Protection Approaches for Discrimination between Internal Fault and Inrush Current in Electric PowerTransformers” Journal of Electrical Engineering,2019.

[5] Santosh Kumar Nanda “Virtual Instrument based Fault Classification in Power Transformers using Artificial Neural Networks” 1st International Conference on Condition Assessment

Techniques in Electrical Systems, IEEE, 2018.

[6] Ms. A. Pavithra "Artificial Neural Network Approach for Discriminating Various faults in Transformer Protection" International Journal ofResearch and Engineering, 2017.

[7] ManojTripathy, "Power transformer differential protection using neural network Principal Component Analysis and Radial Basis Function Neural Network", Simulation Modelling Practice andTheory, Elsevier, 2017.

[8] Sendilkumar Subramanian "Wavelet Packet Transform and Support Vector Machine Based Discrimination of Power Transformer Inrush Current from Internal Fault Currents" Modern Applied Science, Vol. 4, No. 5; May 2016.

[9] S. N. Deepa, S.S. Sumathi, S.N. Sivanandam. "Neural Network using MATLAB", 1sted.TataMcgraw Hill :2015.

[10] P. Arboleya, G. Diaz, J.G. Aleixandre, "A solution to the dilemma inrush/fault in transformer relaying using MRA and wavelets", Electric Power Compo. Syst. 34, 2006.M. Tripathy, R.P. Maheshwari, H.K. Verma, "Advances in transformer protection: a review", Electric Power Compo. Syst. 33 (11) 1203–1209 2005.

[11] S.A. Saleh, M.A. Rahman, "Modeling and protection of a three-phase transformer using waveletpacket transform", IEEE Transactions on Power Delivery. 20 (2),1273–1282, 2005.

[12] M.C. Shin, C.W. Park, J.H. Kim, "Fuzzy logic based relaying for large power transformer protection", IEEE Transactions on Power Delivery. 18 (3),718–724, 2003.

[13] P.L.Mao and R.K. Aggarwal. "A wavelettransform based decision making logic method for discrimination between internal faults and inrush currents in power transformers" International journal of Electric power and Energy systems, Vol.22, Oct 2000.

[14] H.K. Verma, G.C. Kakoti, "Algorithm for harmonic restraint differential relaying based on the discrete Hartley transform", Electric Power Syst. Res.18 (2) , 125–129, 1990.

[15] B.Golomb and T.Sejnowski,"Sexrecognitionfromfaces usingneural networks," in Applications of Neural Networks, A. Murray, Ed. Norwell, MA: Kluwer,1995, pp. 71–92.

[16] B. B. Thompson, R. J. Marks, II, J. J. Choi, M.A. El-Sharkawi, M. Y. Huang, and C. Bunje, "Implicit learning in autoencoder noveltyassessment," in Proc. IEEE Int. Joint Conf. Neural Netw. World Congr. Comput. Intell., Honolulu, HI, May 12–17, 2002, pp. 2878–2883.

[17] T. D. Sanger, "Optimal unsupervised learning ina single-layer linear feed forward neural network,"NeuralNetw.,vol.2,pp.45 9–473,1989