DESIGN OF MACHINE LEARNING ALGORITHM FOR FAULT DIAGNOSIS IN POWER TRANSFORMER

[(1)]A.C Shri Nandhny, Dr.R.Subasri[(2)]
[(1)] PG Scholar, [(2)] Associate Professor
[(1)] M.E Control and Instrumentation, Department of EIE, Kongu Engineering College, Erode, India

Abstract: Reliability of power system could be very essential to generate and transmit energy. Power transformer is one of the most important electric apparatus and therefore it need to be saved in good state. The incipient fault identification and classification is a major research area. The dissolved gas analysis (DGA) is a technique being extensively used to find out incipient faults but diverse strategies had been developed to analyze DGA effects, they will once in a while fail to diagnose exactly. The accurate identification of incipient fault using numerous artificial intelligence (AI) is varied with variation of input parameters. Principle Component Analysis using Rapidminer Software is applied to IEC TC10 databases and associated datasets to find out most influencing input parameters for incipient fault classification. Further, one of the Machine learning algorithm namely Extreme Learning Machine (ELM) is implemented to categorize the power transformer incipient faults accurately.

IndexTerms: Dissolved Gas Analysis, Extreme Learning Machine, Feature Selection and Classification, Power Transformer, Rapidminer.

I. INTRODUCTION

The power transformer is a major apparatus in an electrical network. Power transformer windings are insulated with multilayers of paper and immersed in oil. This paper insulation is required to withstand both electrical and mechanical stresses. The paper and oil insulation degrades over time at a rate depending on the moisture level, oxygen present and the operating temperature of the oil. The breakdown of insulation materials and its related components inside a transformer generates gases within the transformer. The identity of the gases being generated can be very useful information in any preventive maintenance program The two principle causes of gas formation within an operating transformer are electrical disturbances and thermal decomposition. The oil inside the transformer used for insulation purpose and it is the mixture of different hydrocarbon molecules, and the decomposition of these hydrocarbons in thermal or electrical faults are complex. The chemical reactions involve the breakage of carbon hydrogen and carbon-carbon bonds. During this process, active hydrogen atoms and hydrocarbons are formed. The hydrocarbon and active hydrogen atoms can combine with each other to form gases: hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), and ethane (C_2H_6). Due to thermal decomposition and breakdown of insulting materials produce methane (CH_4), hydrogen (H_2), carbon monoxide (CO), and carbon dioxide (CO_2). The above mentioned gases are considered as key gases and are generally considered it as combustible (note that CO_2 is not a combustible gas). All combustible gases may indicate the one or a combination of thermal, electrical or corona faults. Thus, preventive techniques for early fault detection in these transformers to avoid outages

Transformer oil analysis is a useful, predictive, maintenance tool for determine the condition of the transformer. Using oil sample, dissolved gas analysis (DGA) is performed and it is useful in evaluating transformer health. Dissolved Gas Analysis (DGA) is the technique for the earliest detection of incipient faults in transformers. Thus, interpretation of dissolved gas analysis (DGA) i.e.., Roger's ratio method, Key gas method, IEC ratio method, Doernenburg ratio method, Duval Triangle method is used as the preventive maintenance program to detect the incipient faults. The DGA techniques are still unable to detect the multiple faults and unable to detect the new or unknown faults owing to the lack of expert knowledge in them. So, various artificial intelligence (AI) techniques may help to solve the problems and present a better solution. Transformer fault diagnosis using artificial intelligence has the capability of making early diagnosis of the faults in the transformer. Extreme Learning Machine, one of the machine learning algorithm applied to DGA data to classify the incipient faults.

This paper is formulated as follows: Introduction and Literature review is given in Section 1. The database development is presented in Section 2. Fault classification using ELM is presented in Section 3. The results are presented and discussed in Section 4 and conclusion in Section 5.

II. DATABASE AND METHODOLOGY

2.1 Database Development

The 125 DGA instances are accumulated from publicly available online database (117 instances from IEC TC10 Database [8] and nine instances of normal condition from Table 1 to 2 [4]) for training and testing intend of ELM model. The following considerations were assumed during the development of the database: i) When concentration of a given gas is not available, consider that value is "0". ii) When a ratio is (0/0) form, consider null. iii) Consider "20" when a givenratio has infinite value assuming that $\lambda \neq 0$. iv) If a concentration represented by "<1" is assumed as 0.5.

From the collected database, input space vector (X) of twenty one variables is prepared, which includes: the 7 key gases in ppm (C_2H_2, H_2, C_2H_4, CO, C_2H_6, and CH_4, CO_2), the 7 ratio between the key gases (CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, C_2H_2/CH_4, C_2H_4/CH_4, C_2H_2/H_2, CO_2/CO) and the corresponding percentage value of 7 gases ($\%CH_4=100X/(X+Y+Z)$, $\%C_2H_4=100Z/(X+Y+Z)$, $\%H_2=100H_2/(H_2+C_2H_6+CO+CO_2)$, $\%C_2H_2=100Y/(X+Y+Z)$, $\%C_2H_6=100C_2H_6/(C_2H_6+X+Y+Z)$, $\%CO=100CO/(CO+CO_2+C_2H_6+X+Y+Z$ & $\%CO_2=100CO_2/(CO+CO_2+C_2H_6+X+Y+Z)$, where $X=CH_4$, $Y=C_2H_2$, and $Z=C_2H_4$). Most relevant input variable selection is the foremost step for designing an AI based approach for incipient fault classification. Our aim is to identify the most relevant input variable to the ELM model which has higher classification accuracy as explained in following section.

2.2 Input Variable Selection

The input space vector X consists of 21 input variables are given as follow

$$X = \begin{bmatrix} H_2, CO_2, C_2H_2, CO, C_2H_4, C_2H_6, CH_4, \frac{CH_4}{H_2}, \frac{C_2H_2}{C_2H_4} \\ \frac{C_2H_4}{C_2H_6}, \frac{C_2H_2}{CH_4}, \frac{C_2H_4}{CH_4}, \frac{C_2H_2}{H_2}, \frac{CO_2}{CO}, \%CH_4, \%C_2H_2, \\ \%C_2H_4, \%CO, \%CO_2, \%C_2H_6, \%H_2 \end{bmatrix}_{125\times 21}$$

Using Rapidminer explorer [7], the calculated 21 attributes from 125 data samples are reduced. Inorder to reduce the dimensionality in the input space vector X, Principle Component Analysis is used. In this Software, Weight by PCA and Select by weight methods are used. Weight by PCA is used as ranking technique and Select by weight is used as selection technique and analyse the classification accuracy. . The order of attributes C_2H_2, H_2, C_2H_4, CO, C_2H_6,CH_4, CO_2 , CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, C_2H_2/CH_4, C_2H_4/CH_4, C_2H_2/H_2, CO_2/CO , $\%CH_4$,$\%C_2H_4$, $\%H_2$, $\%C_2H_2$,$\%C_2H_6$, $\%CO$ and $\%CO_2$ have been discarded one-by-one from the input variable vector X and we found the higher classification accuracy than the accuracy of pruneless input variable vector. The problem of input variable selection is completed, one can reduce the training data to include only ten significant attributes, namely $\%C_2H_4$, $\%C_2H_2$, $\%H_2$, C_2H_4/C_2H_6, C_2H_2/CH_4, C_2H_4/CH_4 , CH_4/H_2, C_2H_2/C_2H_4, C_2H_2/H_2 that is recognized as the most informative input variables. This combination of selected variables will give higher fault classification accuracy with AI models of Rapidminer.

III. FAULT CLASSIFICATION USING ELM

3.1 Extreme Learning Machine

ELM is a recently developed learning method for Single Layer Feedforward Neural Network architecture (SLFNs). The SLFNs based ELM network architecture used in this paper is presented in Figure-1. ELM is utilized to resolve a different kind of problems such as real time learning & prediction, feature selection, classification, time series, biometrics, signal processing, disease prediction & eHealthcare, security & data privacy, biomathematics, image processing, human action recognition, brain computer interface and human computer interface location positioning system.

In general, ELM model has only one hidden slayer with parameters of input weight (w) and biases (b) of the hidden nodes (1, 2,....., L) . These parameters need not to be tuned. Even if the number of hidden neurons (L) < the number of distinct samples (G), ELM can still assign random parameters to the hidden nodes and calculate the output weight using pseudo inverse of H giving only a small error (□□ o). All hidden nodes parameters as assigned randomly, which are independent upon the target function (t) and the training data (G number of sample). Afterwards, the output weights (β) (linking the hidden layer to the output layer) are determined and analytically using a Moorse-Penrose generalised inverse (H^{ψ}). ELM provides much better generalisation results at extremely fast learning speed as compared to traditional learning algorithm for ANN, because of its simple and efficient learning algorithm. The formulation of ELM will be summarized as given below,

Consider a pre-processed training data set D = (x_n ,t_n), where n =1,2,....,G The classification of ELM with h hidden nodes and activation function can be mathematically represented as,

$$f_i\left(x_p\right) = \sum_{i=1}^{h} w_i x_p + b_i \ldots\ldots(1)$$

Where,

$$g\left(w_i x_p + b_i\right) = \frac{1-\exp\left(-w_i x_p + b_i\right)}{1+\exp\left(-w_i x_p + b_i\right)} \ldots\ldots\ldots(2)$$

x_p= unseen case features data pre-processing; w_i = weight vector between input nodes and i^{th} hidden nodes, β_i = weight vector between the output nodes and i^{th} hidden nodes; b_i =biases (i.e., threshold or centers and impact factors of the i^{th} hidden node); G = training data with zero error.

If an ELM with h hidden nodes can approximate G training data with zero error, it implies that there exist β_i ,w_i and b

$$f_i\left(x_p\right) = \sum_{i=1}^{h} w_i x_p + b_i = H\beta \ldots\ldots\ldots(3)$$

Where, H is the Hidden layer matrix is given by

$$\begin{bmatrix} h\left(x_1\right) \\ \vdots \\ h\left(x_G\right) \end{bmatrix} = \begin{bmatrix} g\left(w_1 x_1 + b_1\right) \ldots g\left(w_h x_1 + b_h\right) \\ \vdots \\ g\left(w_1 x_G + b_1\right) \ldots g\left(w_h x_G + b_h\right) \end{bmatrix} \ldots\ldots\ldots(4)$$

$$\beta = \begin{bmatrix} \beta_{1,1} \ldots \beta_{1,m} \\ \vdots \\ \beta_{h,1} \ldots \beta_{h,m} \end{bmatrix} \ldots\ldots\ldots(5)$$

The ith column of H is the ith hidden node output with respect to inputs x_1 x_2 , ,....., x_G. The nth row of H is the hidden layer feature mapping w.r.t. the nth input x_n .The output weight vector β can be calculated by

$$\beta = H^{\psi}T \ldots\ldots\ldots (6)$$

Where H^{ψ} is the Moore-Penrose pseudo inverse of the hidden layer output matrix H and can be calculated using several methods including iterative method (IM), orthogonal projection methods (OPM), singular value decomposition(SVD), orthogonalization method (OM) etc. The OPM can be used only when H^T H is non-singular and $H^{\psi} = (H^TH)^{-1}H^T$. Due to the use of searching and iteration, OM and IM have some limitations. Implementation of ELM utilizes SVD to find out the H^{ψ} , since it can be used in all situations.

T is Output Matrix which is given below

$$\begin{bmatrix} t_{1,1} \cdots t_{1,m} \\ \vdots \\ t_{h,1} \cdots t_{h,m} \end{bmatrix} \ldots\ldots\ldots(7)$$

According to above learning algorithm, training and testing time for ELM is extensively fast due to only three calculation steps are needed.

Step 1: Randomly assigns the input weight w_i and bias b_i according to any continuous sampling distribution, i = 1, 2,...h

Step 2: Calculate the hidden layer output matrix H

Step 3: Calculate the output weight β ($\beta = H^{\psi}$ T)

3.2 DGA Training and Testing Data

Database is bifurcated into 2 sets: randomly chosen 85 cases for training and for the testing 40 cases. The complete datasets are analyzed using numerous DGA methods and therefore the corresponding judgments associated with six categories are provided: normal unit-NF (9 cases), partial discharge-PD (9cases), discharge of low energy-D1 (26 cases), discharge of high energy-D2 (47 cases), low and medium temperature overheating-T1/T2 (<700^oC) (16 cases), and high temperature overheating-T3 (>700^oC) (18 cases).

ELM based fault classification is performed using Key gas, IEC ratio and Roger's Ratio. The Table1 shows that the number of nodes, number of data samples used for training and testing for each method.

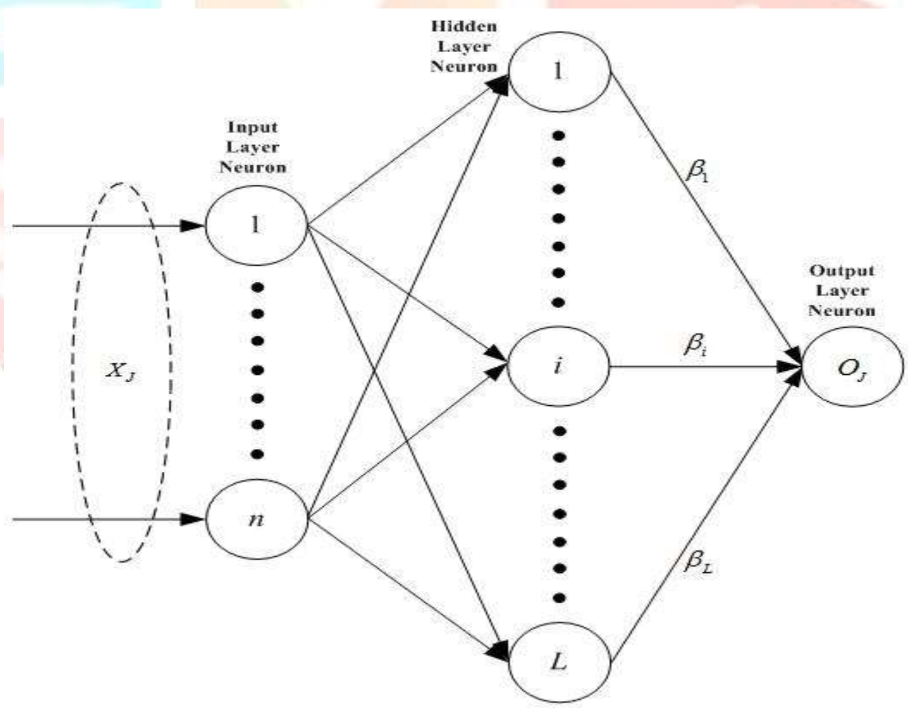

Figure-1: Schematic Diagram of ELM

Table-1 Structure of Various DGA Methods

DGA Methods	No of Input	No of Output	No of Training	No of Testing

	Nodes	Nodes	Data	Data
Key gas	5	6	85	40
Roger's Ratio	4	6	85	40
IEC Ratio	3	6	85	40

3.3 ELM Based Fault Classification

The diagnostic model which is utilized to distinguish the six states: normal state (NF) and the five fault conditions (FC). The EFIM has been designed according to the 125 data samples of DGA with associated fault condition. 85 training records for the ELM are stored in Data Base, and the 40 records are used for testing purpose.

The performance of proposed model for incipient fault identification is found out by computing the necessary measures after the modifications.

Imbalance fault identification accuracy (IFIA),

$$IA = \frac{\text{Correctly Classified Samples}\big|_{\text{by model}}}{\text{Total Number of Samples}\big|_{\text{by dataset}}} \quad \ldots\ldots\ldots (8)$$

Mean Square Error (MSE) is given by,

$$MSE = \frac{1}{n}\sum_{q=1}^{n}\left(E_q{}^2\right);\ \text{where } E_q = \left|T_q - OA_q\right| \ldots\ldots. (9)$$

Where n = number of samples in data set, T_q = target value and OA_q= actual model output obtained from the trained ELM based classifier

IV. RESULT AND DISCUSSION

Five ELM models are designed using different input variables. Five key-gases are used as input in model1 and and Roger's & IEC's ratios of gases are used in model2 and model3 respectively. In model 4, percentage values of three gases are used as input variable. Finally, model5 uses selected input variable by PCA method as an input. After successful training of each model, testing is performed and then performance analysis of each model is analyzed using Eq. (8) and (9) as shown in Table 2. The testing data is not included with training data set.

V. CONCLUSION

Power Transformer being the vital and essential device in electric network has gained the maximum importance. In operating condition of power transformer it experiences the electrical, electromagnetic, thermal and mechanical stress which leads to an incipient fault. These incipient fault leads to failure of power transformer and failure leads to unplanned outage or sudden breakdown of power transformer. This may reduce the designed life span of power transformer and hence cost effectiveness.

Power transformer is major power system equipment. Dissolved Gas Analysis is an efficient tool for diagnosing incipient faults in oil filled electrical equipment. From the analysis of the power transformer, gases such as ethane (C_2H_6), methane (CH_4), hydrogen (H_2), acetylene (C_2H_2), ethylene (C_2H_4), carbon dioxide (CO_2) and carbon monoxide (CO) are produced during its operation due to electrical and thermal stresses and they get dissolved in transformer oil. For early detection, three methods have been applied for the interpretation of fault types from the DGA data namely Roger's ratio, IEC ratio and Key gas. In order to increase the efficiency of fault classification and reduce the dimensionality of input parameters, data analysis software i.e.., Rapidminer is used. Extreme Learning Machine is applied to the shortlisted input parameters to improve the diagnosis of incipient fault in a power transformer.

Table-2: Fault Diagnosis of ELM model

ELM model using input as	ELM model processing phase	MSE	RMSE	Accuracy (%)
Key gas	Training phase	0.133	0.364	96.66
	Testing phase	0.147	0.470	76.92
Roger's Ratio	Training phase	0.208	0.4560	95.83
	Testing phase	0.115	0.339	88.46
IEC Ratio	Training phase	0.534	0.7280	95.27
	Testing phase	0.153	0.392	83.33
Proposed Model	Training Phase	0	0	100
	Testing Phase	0.008	0.091	94.44

REFERENCES

[1] M. Duval, "A Review of Faults Detectable by Gas-in-Oil Analysis in Transformers", IEEE Electr. Insul. Mag. vol. 18, no.3, pp. 8-17, 2002.
[2] Hasmat Malik, A. Azeem and RK Jarial, "Application Research Based on Modern-Technology for Transformer Health Index Estimation", in: Proc. IEEE Intl. Multi-Conf. on Systems, signals and Devices (SSD), vol. 9, pp. 1-7, 2012.
[3] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An Artificial Neural Network Approach to Transformer Fault Diagnosis," IEEE Trans. Power Deliv. Vol. 11, No. 4, pp. 1836-1841, 1996.
[4] L.V. Ganyun, Cheng Haozhong, Zhai Haibao and Dong Lixin, "Fault diagnosis of power transformer based on multi-layer SVM classifier," Int. J. Electr. Power Syst. Res., Vol. 74, pp. 1–7, 2005.
[5] Hasmat Malik, Tarkeshwar and R.K. Jarial, "An Expert System for Incipient Fault Diagnosis and Condition Assessment in Transformers," Proc. IEEE Intl. Conf. on Computational Intelligence and Communication Systems, pp. 138-142, 2011.
[6] R. Naresh, Veena Sharma, and Manisha Vashisth, "An Integrated Neural Fuzzy Approach for Fault Diagnosis of Transformers," IEEE Trans. Power Deliv., Vol. 23, No. 4, pp. 2017-2024, 2008.
[7] RAPID MINER User's Manual. Rapid-I GmbH, 2010. <www.rapid-i.com>.
[8] Michel Duval and Alfonso dePablo, Interpretation of Gas-In-Oil Analysis Using New IEC Publication 60599 and IEC TC 10 Databases IEEE Electrical Insulation Magazine, Vol. 17, No. 2, pp. 31-41, 2001.
[9] MATLAB User's Guide. The MathWorks, Inc., Natick, MA 01760, 1994–2001 <http://www.mathworks.com>.
[10] Hasmat Malik, Amit Kr Yadav, Sukumar Mishra, Taekeshwar Mehto, Application of Neuro-Fuzzy Scheme to investigate the winding insulation paper deterioration in oil-immersed power transformer. Int J Electr Power Energy Syst., vol. 53: pp. 256–271, 2013.
[11] Hasmat Malik and Sukumar Mishra, "Feature Selection using RapidMiner and Classification through Probabilistic Neural Network for Fault Diagnostics of Power Transformer," In Proc. 2014 Annual IEEE India Conference (INDICON), pp. 1-6, 2014.
[12] G.-B. Huang, Q.Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," 2004 International Joint Conference on Neural Networks, (Budapest, Hungary), July 25-29, 2004
[13] Li, X., and Wu., "DGA Interpretation Scheme derived from case study", IEEE Trans. Power Deliv., Vol. S26(2), pp. 1292–1293, 2011.
[14] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, vol. 42(2), pp. 513-529, 2012.
[15] IEEE guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, ANSI/IEEE std. C57.104, 1991.
[16] IEC guide for the interpretation of dissolved and free gases analysis, IEC std. IEC/CEI 60599, 2007.