Application of Soft Computing Methods for Fault Diagnosis in Power Transformers

A.Venkatasami
Department of Electrical and Electronics Engineering
Einstein College of Engineering
Tirunelveli-627012, India
venkatasami.athikkan@gmail.com

P.Latha
Department of Computer Science and Engineering
Government College of Engineering
Tirunelveli-627007, India
platha@gcetly.ac.in

***Abstract*—Power transformer is one of the more important equipment in a power system. It is very essential to keep the equipment in good health at all points of time. Dissolved Gas Analysis (DGA) is a tool used in monitoring power transformers. National and international standards provide guidelines for interpreting DGA data and discerning the nature of fault. In many cases, the suggested guidelines are not able to classify the faults precisely. Soft computing tools such as Support Vector Machine Method and Extreme Learning Machine methods seem to offer more exact classification and hence are recommended in such cases. This paper proposes to apply these relatively new methods in fault classification. In an earlier contribution, the Authors considered only one data base, reported by the IEC was considered. However, a comparison of the applicability of the computational techniques over several data bases was found necessary. Also, a need was found to include combined electrical and thermal faults as a parameter in the classification. This contribution expressly considers in this aspect in some detail. To demonstrate the application of these computational techniques, IEC TC10 database (DB1) and a local database (DB2) are considered. Fault classification based on gas concentration as the input, as also the enthalpy of the corresponding gases are compared. The fault classification based on enthalpy is found to identify the fault more precisely.**

Keywords—transformer; dissolved gas analysis; machine learning

I. INTRODUCTION

Mineral oil impregnated cellulose is an insulation of choice in power transformers. Over a period of time, the insulation undergoes slow chemical decomposition due to temperature rise and possible low energy electrical discharges within the equipment. As a consequence of these effects, the insulation evolves hydrocarbon and carbon oxide gases. The gases so generated dissolve in the transformer oil. The gases can be extracted from the oil in a noninvasive manner and analyzed using gas chromatography. Alternately, they are measured through sensors fitted on the transformer. Dissolved gas analysis is one of the oldest and reliable methods for detecting abnormalities inside a transformer. However, in view of possible uncertainties or erroneous prediction, a need arises of exploring alternate methods. As a matter of fact, DGA techniques are time intensive besides being expensive. Also, due to direct human involvement at every step, detection and interpretation of faults often become inaccurate. In some cases, due to inaccurate gas sampling methods, conventional methods are not able to provide correct fault classification [1,2].

It is in this backdrop that soft computing methods are increasingly being applied. This results in a completely automated method of classification of faults in transformer.

Artificial neural network (ANN) is one such possibility. The ANN’s are essentially 'black box' models and hence, present difficulties in interpretation. [3]. Also, a considerable size of data is required for proper training. Support Vector Machine (SVM) is an improvised method and is widely used in transformer diagnosis. Yan and Zhang [4] reports the use of online DGA data using SVM. Further to this work, the current, state of art, Extreme Learning Machine (ELM) method can give high performance in multifaceted problems and takes comparatively less computational time [5]. The model proposed by Huang et al [6] is seen to be even faster particularly when large data sizes are involved since there is no need for iterative tuning as in other neural networks. ELM based model is used with great advantage in medical applications for classification of images. [7]. The ELM is successfully applied to face recognition on two databases namely Yale face database and Carnegie Mellon University face images database [8]. State preserving extreme learning machine applied to face recognition gives best performance in comparison with other classifiers. Guang-Bin et al [9] modeled a single layer feed forward neural network with activation functions and hidden neurons. Application of Extreme learning machine with RBF (Radial Basis Function) network gives faster results. Huang et al [10] showed that a combination of single layer feed-forward neural network with randomly assigned input weights. Hidden layer biases with nonzero activation function can approximate any continuous functions.

Thermodynamic approach to fault severity estimation of transformers is suggested in [11]. It essentially relates the fault energy dissipated inside the transformer to it enthalpy. Thermo dynamic approach is based on the calculation of enthalpy of forming a substance from another substance. As is known, a change in enthalpy occurs when a physical system undergoes a thermodynamic transformation [12]. Application of ELM

Toolbox for analyzing large volumes of data is detailed by Anton Akusok [13]. IEC TC 10 database is used for Transformer fault diagnosis using adaptive neuro-inference system [14].

This paper focuses on six segment classification model with dissolved gas analysis data of IEC TC10 database. One approach based on SVM method and the other based on ELM are used in this work.

Sections II deals with theoretical aspects of SVM and ELM methods. Fault classification based on SVM is detailed in section III. ELM based fault classification is given in section IV. Section V analyzes the results and broad conclusions are presented in section VI.

II. SOFTCOMPUTING TOOLS

The SVM method is a relatively new method for the classification of linear and non- linear data.

A. SVM Method

The basic principle of SVM is based on establishing an optimal hyper-plane to separate different classes of data. Figure.1 shows a notional optimal hyper-plane separating two classes of data. A1 and A2 are the input attributes.

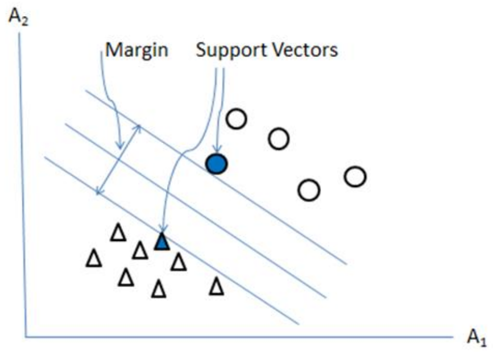

Fig. 1. Optimal Hyperplane with linearly separabe data

A typical classification problem involves separating data into training and testing sets. Each data in the training set contains class label with many “attributes”. Given the test data attributes, the goal of SVM is to create a model based on the training data to predict target values of the test data. Given a training set containing label pairs (xi , yi), i=1,--,l where xi ϵ R^n , i= 1,...., N and a class label y_i ϵ {+1,-1}for each vector. The SVM require the solution of the following optimization problem:

$$\min \phi(w,\xi) = \frac{1}{2}(w.w) + C \sum_i^l \xi_i$$

Subject to $\quad y_i\,[(w.\phi\,(x_i)) + b] \geq 1 - \xi_i$

Where slack vector $\xi_i \geq 0$, i= 1, 2,.... l

Where w is the weight vector and b is the bias term.

In many problems the data is not linearly separable. Hence a kernel function is applied for mapping data into a Vapnik-Chervonek is n- dimensional space. Within this space, a hyper-plane is identified for the separation of classes as shown in Fig.1. Here the training vectors x_i, are mapped into higher dimensional space by the function ϕ. SVM finds a linear separating hyper-plane having a maximum margin in the higher dimensional space. $K(x_i, x_j) = \phi(x_i)$ in which, $\phi(x_j)$ is the kernel function. The kernel of the Radial Basis Function, RBF, is considered as an exponential function thus.

$$K(x_i, x_j) = \exp(-\gamma\| x_i - x_j \|^2), \gamma > 0.$$

This kernel function nonlinearly allocates samples into a higher dimensional space and can solve the problems of non-linear attributes. As a matter of fact, a linear kernel can also be used, however with a loss of classification accuracy. C in Eqn.1 and γ are the SVM parameters.

In MATLAB implementation, hold out cross validation is used for validation of N observations. This returns logical index vectors during cross validation of N records by randomly selecting P X N records to hold out for the evaluation set.in which, P is a scalar with a range between 0 and 1.

B. ELM METHOD

Extreme Learning Machine is a single-layer feed -forward neural network (SLFN). It randomly selects input weights and biases of hidden neuron without training [6]. The norm least-square method and Moore-Penrose inverse method are used to get the output weights. This approach allows a significant training time reduction. Activation function used are unipolar, bipolar and radial basis function for hidden layer neuron and linear activation functions for the output neurons .The SLFN evaluated here uses additive neuron design instead of kernel based, hence random parameter selection. SLFNs are considered as a linear system.

For N arbitrary distinct samples (x_i,t_i), where $x_i = [x_{i1},x_{i2},\ldots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \ldots, t_{im}]T \in R^m$. SLFNs with hidden neurons and activation function g(x) are mathematically modelled as [9]:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \,.\, x_j + b_i) = o_j, \; j = 1,\ldots\ldots N \tag{1}$$

Where $w_i = [w_{i1},w_{i2}, \ldots, w_{in}]^T$ is the weight vector connecting the ith hidden neuron and the input neurons, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{in}]^T$ is the weight vector connecting the i^{th} hidden neuron and the output neurons, and b_i is the threshold of the i^{th} hidden neuron. $w_i \cdot x_j$ denotes the inner product of w_i and x_j .

The standard SLFNs with $\tilde{N}$ hidden neurons each with activation function g(x) can approximate these N samples with zero error mean $\sum_{j=1}^{n} \| o_j - t_j \| = 0$, i.e., there exists β_i,w_i, and b_i such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \,.\, x_j + b_i) = t_j, \; j = 1,\ldots\ldots N \tag{2}$$

The above N equations can be written compactly as

$$H\beta = T \tag{3}$$

Where,

$$H(w_1, \cdots, w_{\tilde{N}}, b_1, \cdots, b_{\tilde{N}}, x_1, \cdots, x_N) =$$

$$\begin{bmatrix} g(w_1.x_1+b_1 & \cdots & g(w_{\tilde{N}}.x_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(w_1.x_N+b_1 & \cdots & g(w_{\tilde{N}}.x_N + b_{\tilde{N}}) \end{bmatrix}_{NX\tilde{N}}$$

$$\beta = \begin{bmatrix} \beta_1{}^T \\ \cdot \\ \cdot \\ \cdot \\ \beta_{\tilde{N}}{}^T \end{bmatrix}_{\tilde{N}XM}, \quad T = \begin{bmatrix} t_1{}^T \\ \cdot \\ \cdot \\ \cdot \\ t_N{}^T \end{bmatrix}_{NXM} \qquad (4)$$

H is called the hidden layer output matrix of the neural network; the i^{th} column of H is the i^{th} hidden neuron output with respect to inputs $x_1, x_2 \ldots.. x_N$.

The SLFN solved by using a gradient based solution by finding suitable values of $\hat{w}_i$,$\hat{b}_i$and $\hat{\beta}(i = 1, \ldots \tilde{N})$ satisfying the model as :

$$\| H (\hat{w}_i, \ldots, \hat{w}_{\tilde{N}}, \hat{b}_1, \ldots, \hat{b}_{\tilde{N}}) \hat{\beta} - T \| = \min_{w_i, b_i, \beta} \| H (\hat{w}_i, \ldots, \hat{w}_{\tilde{N}}, \hat{b}_1, \ldots, \hat{b}_{\tilde{N}}) \hat{\beta} - T \| \qquad (5)$$

A gradient based learning algorithm can be used to minimise the Hβ = T by adjusting the parameters, w_i, b_i and β , when the H hidden layer matrix is unknown iteratively . Studies carried out by Huang et al [10] proved that single layer feed-forward neural network with randomly assigned input weights and hidden layer biases and with almost any nonzero activation function can universally approximate any continuous functions on any input data sets. Huang et al [5] suggested an alternate way to train a SLFN by finding a least square solution $\hat{\beta}$ of the linear system represented by (3).

The unique minimum norm least square (LS) solution is modelled as:

$$\hat{\beta} = H^{\dagger} T \qquad (6)$$

where H † is the MP generalized inverse of matrix H. As analysed by Huang, ELM using such MP inverse method tends to obtain good generalization performance with dramatically increased learning speed. The summarization of the ELM algorithm can be as:

Given a training set, N, then,

$$N = \{ (x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, \ldots, N \},$$

kernel function f(x), and hidden neuron .

Step 1: Select an appropriate activation function and number of hidden neurons for the given problem.

Step 2: Assign arbitrary input weights, w_i, and bias b_i , i= 1, ...H

Step 3: Calculate the output matrix H at the hidden layer
H = f. (w x+ b)

Step 4: Calculate the output weight β.

$\hat{\beta}$= H†T.

C. DAFABASE FOR SOFT COMPUTING TOOLS

The DB1 database with 151 dataset showing training and testing is shown in Table I. The output features are partial discharge (PD), Discharge of low energy (D1), Discharge of High Energy (D2), Thermal faults ≤ 700 °C (T1 and T2), Thermal faults >700 °C (T3), Normal (NF) etc.

TABLE I. DB1 DATABASE

Sr.No	Output features (Fault types)	Training Dataset	Testing Dataset	Total Dataset
1	Partial Discharge (PD)	6	3	9
2	Discharge of low energy (D1)	18	8	26
3	Discharge of High Energy (D2)	33	15	48
4	Thermal faults ≤ 700 °C (T1 and T2)	12	4	16
5	Thermal faults >700 °C (T3)	13	5	18
6	Normal (NF)	24	10	34
	Total dataset	106	45	151

The DB2 database containing 219 dataset showing training and testing is shown in Table II. The output features considered are Partial Discharge (PD), Over Heating (OH), Arcing (A), Electrical &Thermal (E&T) etc.

TABLE II. DB2 DATASET

Sr.No	Output features (Fault types)	Training Dataset	Testing Dataset	Total Dataset
1	Partial Discharge (PD)	24	5	29
2	Over Heating (OH)	99	21	120
3	Arcing (A)	47	11	58
4	Electrical &Thermal (E&T)	9	3	12
	Total dataset	179	40	219

III. FAULT CLASSIFICATION BASED ON SVM

A. Database DB1 with gas concentration and enthalpy

The five models developed with gas concentration as input feature is shown in Fig.2. Model1 classifies fault from normal state. Electrical fault is classified from thermal fault in Model2. Low thermal fault is distinguished from high thermal fault in Model3. Model4 classifies PD from other discharges. Discharge of low energy is classified from discharge of high energy in Model5.

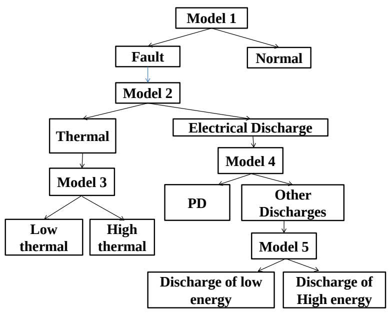

Fig. 2. Five SVM models on transformer fault classification

About 2/3rd of dataset is used for testing purpose and 1/3rd dataset for validation. The Radial Basis Function (RBF) is used as a kernel function. Performance of all the five models is evaluated and the results are given in Table III.

TABLE III. CONFUSION MATRIX WITH GAS CONCENTRATION

Model1	from \ to	Normal	Fault	Total	% correct
	Normal	8	5	13	61.54%
	Fault	1	36	37	97.30%
	Total	9	41	50	88.00%
Model2	from \ to	Thermal	Electrical Discharge	Total	% correct
	Thermal	9	2	11	81.82%
	Electrical Discharge	0	28	28	100.00%
	Total	9	30	39	94.87%
Model3	from \ to	High Thermal	Low Thermal	Total	% correct
	High Thermal	5	0	5	100.00%
	Low Thermal	2	5	7	71.43%
	Total	7	5	12	83.33%
Model4	from \ to	Other Discharges	PD	Total	% correct
	Other Discharges	24	1	25	96.00%
	PD	0	3	3	100.00%
	Total	24	4	28	96.43%
Model5	from \ to	Discharge of Low Energy	Discharge of High Energy	Total	% correct
	Discharge of Low Energy	17	0	17	100.00%
	Discharge of High Energy	3	5	8	62.50%
	Total	20	5	25	88.00%

The Radial Basis Function (RBF) is used as a kernel function. Table IV shows the results of the five models evaluated using enthalpy as input feature.

TABLE IV. CONFUSION MATRIX WITH ENTHALPY

Model1	from \ to	Normal	Fault	Total	% correct
	Normal	8	4	12	66.67%
	Fault	3	35	38	92.11%
	Total	11	39	50	86.00%
Model2	from \ to	Thermal	Electrical Discharge	Total	% correct
	Thermal	9	1	10	90.00%
	Electrical Discharge	0	29	29	100.00%
	Total	9	30	39	97.44%
Model3	from \ to	High Thermal	Low Thermal	Total	% correct
	High Thermal	6	1	7	85.71%
	Low Thermal	1	4	5	80.00%
	Total	7	5	12	83.33%
Model4	from \ to	Other Discharges	PD	Total	% correct
	Other Discharges	23	1	24	95.83%
	PD	0	4	4	100.00%
	Total	23	5	28	96.43%
Model5	from \ to	Discharge of Low Energy	Discharge of High Energy	Total	% correct
	Discharge of Low Energy	18	2	20	90.00%
	Discharge of High Energy	0	5	5	100.00%
	Total	18	7	25	92.00%

From the table, fault classification accuracies are 86%, 97.44%, 83.33%, 96.43% and 92.0% for model1, model2, model3, model4 and model5 respectively.

B. *Database DB2 with gas concentration and enthalpy*

The three models developed with gas concentration as input feature is shown in Fig.3.

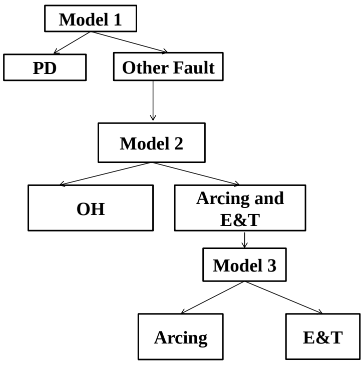

Fig. 3. Three SVM models on transformer fault classification

The SVM model1 classifies PD from other faults. The model2 classifies over heating from arcing, electrical and thermal. The Arcing fault is classified from electrical and thermal in SVM model3.

TABLE V. CONFUSION MATRIX WITH GAS CONCENTRATION

	from \ to	Other Fault	PD	Total	% correct
Model1	Other Fault	63	1	64	98.44%
	PD	2	7	9	77.78%
	Total	65	8	73	95.89%
Model2	from \ to	A and E&T	OH	Total	% correct
	A and E&T	24	0	24	100.00%
	OH	3	36	39	92.31%
	Total	27	36	63	95.24%
Model3	from \ to	E&T	A	Total	% correct
	E&T	3	2	5	60.00%
	A	0	18	18	100.00%
	Total	3	20	23	91.30%

DB2 has 219 dataset. The confusion matrix in Table V shows fault classification accuracies of 95.89%, 92.31% and 91.3% for model1, model2 and model3 respectively with gas concentration as input feature.

From the gas concentration, Enthalpy is calculated using Table VI.

TABLE VI. ENTHALPY OF TRANSFORMER FAULT GASES

Sr.No	Fault Gas	Chemical formula	Enthalpy (kJ/mol)
1	Methane	CH_4	77.7
2	Ethane	C_2H_4	93.5
3	Ethylene	C_2H_6	104.1
4	Hydrogen	H_2	128.5
5	Acetylene	C_2H_2	278.3

All the three models are evaluated with enthalpy of the gas as input feature and the results are tabulated in Table VII.

TABLE VII. CONFUSION MATRIX WITH ENTHALPY

	from \ to	Other Fault	PD	Total	% correct
Model1	Other Fault	64	1	65	98.46%
	PD	1	7	8	87.50%
	Total	65	8	73	97.26%
Model2	from \ to	A and E&T	OH	Total	% correct
	A and E&T	21	0	21	100.00%
	OH	1	41	42	97.62%
	Total	22	41	63	98.41%
Model3	from \ to	E&T	A	Total	% correct
	E&T	3	2	5	60.00%
	A	0	18	18	100.00%
	Total	3	20	23	91.30%

The Table VII shows the confusion matrix for all the three models. The fault classification accuracies obtained are 97.26%, 98.41% and 91.30% for model1, model2 and model3 respectively.

IV. FAULT CLASSIFICATION BASED ON ELM

Extreme Learning Machine (ELM) is used for the classification of faults in power transformers.

A. *Extreme Learning Machine Method*

ELM is a single layer with feed forward neural network (SLFNN). It randomly selects input weights and hidden layer biases without training. The output weights are obtained analytically using the norm least square solution and Moore-Penrose inverse of a general linear system. Four models mentioned earlier, are analyzed using Matlab packages The flow charts shown in Figs. …….give the methodology in detail.

The gas data are obtained from gas sensors fitted in the power transformer or from gas chromatograph. Out of seven gases as shown in Fig. 4, five gases are selected as input features for ELM-I and ELM-II models.

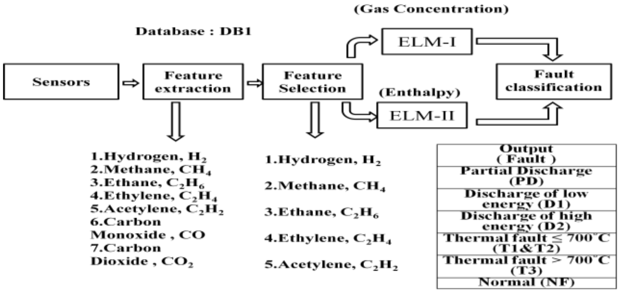

Fig. 4. hree ELM models on transformer fault classification

ELM-I classifier displays a confusion matrix showing the actual and predicted class with gas concentration as input feature is shown in Table VII.

TABLE VIII. CONFUSION MATRIX OF GAS CONCENTRATION

Actual Class	Predicted Class					
	PD	D1	D2	T1&T2	T3	NF
PD	2	0	0	1	0	0
D1	0	8	0	0	0	0
D2	0	0	13	1	0	1
T1&T2	0	0	0	1	2	1
T3	0	0	1	0	4	0
NF	0	0	0	0	0	10

In the table showing actual 3 cases of PD class, the computer classified 2 cases as PD and one case is classified as T1&T2 class. The actual and predicted class for other fault classes is shown in the table. The classification accuracy of EML-I model achieved with DB1 database using gas concentration as input feature is 84.44 %.

The Table IX shows the confusion matrix calculated with enthalpy of gas as input feature. It shows the actual and the predicted results of PD, D1, D2, T1&T2, T3 and NF class.

TABLE IX. CONFUSION MATRIX FOR ENTHALPY

Actual Class	Predicted Class					
	PD	D1	D2	T1&T2	T3	NF
PD	3	0	0	0	0	0
D1	0	8	0	0	0	0
D2	0	2	13	0	0	0
T1&T2	0	0	0	4	0	0
T3	0	0	0	0	5	0
NF	0	0	1	0	0	9

In the confusion matrix shown above, the classifier has predicted all three cases correctly as PD class. The classifier predicted all D1. D2, T1&T2, T3 class correctly. For 10 cases of NF class, 9 cases are classified correctly with 1 case wrongly classified. The classification accuracy of EML-II model achieved with DB1 database using enthalpy as input feature is 95.55 %.

B. DB2 with gas concentration and enthalpy

Database DB2 contains five gases as input features and the output features are PD, OH, A, E&T and NF respectively. Two models are developed namely ELM-III and ELM-IV as shown in Fig.5.

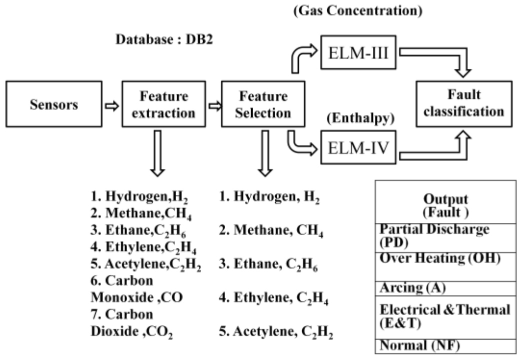

Fig. 5. Three ELM models on transformer fault classification

Dataset used for training and testing are shown in Table IX. DB2 database has 29, 120, 58, 12 dataset of PD, OH, A, and E&T class respectively.

Table X shows the confusion matrix with gas concentration as input feature. The ELM-III classifier classifies the PD, OH and E&T class correctly. With 10 dataset containing A class, the classifier predicted 10 cases correctly and 1 case wrongly classified.

TABLE X. CONFUSION MATRIX OF GAS CONCENTRATION

Actual Class	Predicted Class			
	PD	OH	A	E&T
PD	5	0	0	0
OH	0	21	0	0
A	0	0	10	1
E&T	0	0	0	3

The confusion matrix with enthalpy of gas as input feature is shown in Table XI. All the classes namely PD, OH, A and E&T are predicted correctly. The classification accuracy of EML-III model achieved with DB2 database using gas concentration as input feature is 97.50 %.

TABLE XI. CONFUSION MATRIX OF ENTHALPY

Actual Class	Predicted Class			
	PD	OH	A	E&T
PD	5	0	0	0
OH	0	21	0	0
A	0	0	11	0
E&T	0	0	0	3

The classification accuracy of EML-IV model achieved with DB2 database using enthalpy as input feature is 100%.

V. ANALYSIS OF RESULTS

SVM and ELM methods are applied to databases DB1 and DB2. Input features used are gas concentration and enthalpy. Output features used in DB1 database are PD, D1, D2, T1&T2, T3 and NF. PD, OH, A and E&T are used as output features in DB2 database. The performance of the SVM and ELM classifiers are evaluated.

A. SVM method

The comparison of five models applied to DB1 database and three models used in DB2 database with gas concentration and enthalpy as input features are shown in Fig.6.

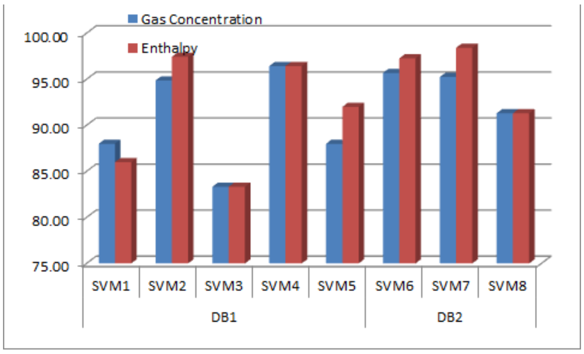

Fig. 6. Classification accuracy (in %) of SVM models

Classification accuracy with DB1 database with gas concentration and enthalpy is 90.90% and 91.55% respectively. Hence, SVM classifier shows more or less accuracy using DB1 database.

With DB2 database, classification accuracy with gas concentration and enthalpy is 94.96% and 96.85% respectively. Hence, SVM classifier shows more or less same accuracy using DB2 database.

B. ELM method

Fig.6 shows the comparison of SVM models applied to DB1and DB2 database using gas concentration and enthalpy as input feature is shown in Fig,7.

Classification accuracy with DB1 database using gas concentration and enthalpy is 84.44% and 95.55% respectively. Hence, ELM classifier shows an improved performance on DB1 database.

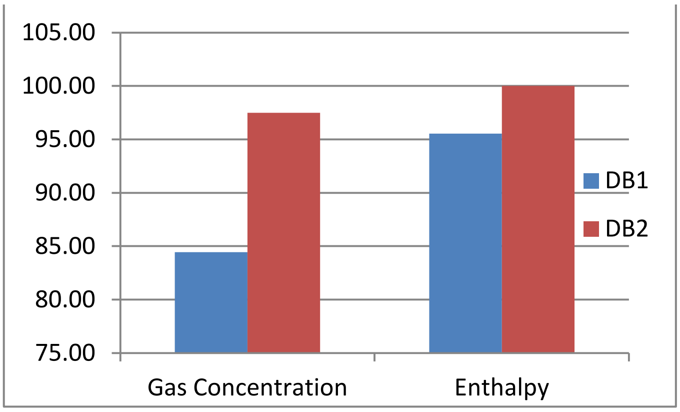

Fig. 7. Classification accuracy (in %) of ELM models

With DB2 database, the accuracy of classification with gas concentration and enthalpy is 97.50% and 100.00% respectively. Hence, there is an Improvement with ELM classifier on DB2 database.

VI. CONCLUSIONS

This paper presents the application of SVM and ELM approaches to transformer fault classification. There are 5 different models under SVM and 3 models in ELM are developed and validated on DB1 and DB2 databases. Conventional transformer fault classification based on gas concentration is compared with fault classification with enthalpy of gas.

ELM models perform better than the SVM models for transformer fault classification.

From the results, Fault classification based on enthalpy of gas as input feature presents a better performance of fault classification in power transformers.

ACKNOWLEDGMENT

The authors would like to acknowledge the Principal Dr.K. Ramar, Einstein College of Engineering, for supporting this Research. The authors are also thankful to Mr.A. Amudhavanan, Managing Trustee, Einstein College of Engineering, for granting permission to publish this paper.

REFERENCES

[1] IEC publication 60599," Mineral oil impregnated electrical equipment in service - Guide to the interpretation of dissolved gas and free gases analysis",March 1999

[2] IEEE Guide for the interpretation of gases in oil-immersed transformers, IEEE Std.C57.104-2008.

[3] Kloppel, S.,Stonnington, CM.,Dragnasski, Chu C.,and Scahill, R I. (2008) Automatic classification of MR scans in Alzheimer's disease. Brain. 2008;131(3):681-689. http://dx.doi.org/10.1109/9.402235

[4] Yanhang.,Bhide Zhang. (2010) Transformer fault diagnosis based on support vector machine. IEEE International conference on computer science and information technology, **2** , 681-689.DOI: 10.1109/ICCSIT.2010.5563911

[5] Huang, GB., Zhu, QY., and Siew, CK. (2006) Extreme learning machine Theory and applications. Neuro Computing, **70(1)**, 489-501.

[6] Gulloway, MM. Texture analysis using Grey Level run lengths. Copyright © 1975 Published by Elsevier Inc.doi:10.1016/S0146-664X(75)80008-6

[7] Deepa, SN., and Arunadevi, B. (2013) Extreme learning machine for classification of brain tumor in 3D MR images. Informatol, **46**, 111-121.

[8] Zahangir Alam, M., Sidike, VK.., and Taka, TM. (2015) State preserving extreme learning machine for face recognition. International joint conference on Neural Networks, 1-7. DOI: 10.1109/IJCNN.2015.7280788

[9] Guang Bin., Huang Ang., and Chee Kheong Siew. (2004) Extreme learning Machine with randomly assigned RBF Kernels. The proceedings of the Eighth International Conference on Control, Automation, Robotics and Vision.

[10] Huang,G.B., Zhu, Q.Y., and Siew,C.K,Extreme learning machine : A new learning scheme of feedforward neural networks, Proc. Int, conf. Neural Networks , Budapest, Hungary, 985-990 (2004).

[11] Fredijakob., and James , JD. (2015) Thermodynamic estimation of Transformer fault severity. IEEE Transactions on Power Delivery, 39, 1941 - 1948. DOI: 10.1109/TPWRD.2015.2415767

[12] Jacob, F., and Noble, P., and Dukarm, J. (2012) A thermodynamic approach to evaluation of the severity of transformer faults. IEEE Transactions on Power Delivery, **27**, 554 - 560. DOI: 10.1109/TPWRD.2011.2175950

[13] Anton Akusok, Kaj-Mikael Bjork, Yoan Miche and Amaury Lendasse, High Performance Extreme Learning Machines: A Complete Toolbox for Big Data Applications, IEEE IEEE Access (Volume: 3) ,Page(s): 1011 - 1025, 30 June 2015 , DOI: 10.1109/ACCESS.2015.2450498

[14] A. Venkatasami , P. Latha and K. Kasirajan, Application of Adaptive Neuro-Fuzzy Inference System Based on IEC Method for Transformer Fault Diagnosis, International Journal of Soft Computing , 2014, Vol.9, Issue : 5, pp. 333-337, 2014.

[15]