Multilevel SVM and AI based Transformer Fault Diagnosis using the DGA Data

GVSSN Srirama Sarma[1], B. Ravindranath Reddy[2], Pradeep M. Nirgude[3], Vasudeva Naidu[4]

[1]Research Scholar, JNTUH & Assistant Professor, Dept. of Electrical and Electronics Engineering, Matrusri Engineering College, Saidabad, Hyderabad, India.
[2]Deputy Executive Engineer & JNTUH Univerty, Kukatpally, Hyderabad, India.
[3]Additional Director, UHV Research Laboratory, CPRI, Hyderabad, India.
[4]Associate Professor, Dept. of Electrical and Electronics Engineering, Matrusri Engineering College, Saidabad, Hyderabad, India.
[1]music.sarma2016@gmail.com, [2]bumanapalli_brreddy@yahoo.co.in, [3]pmnirgude@cpri.in, [4]pvdnaidu81@gmail.com

Abstract

The Dissolved Gas Analysis (DGA) is utilized as a test for the detection of incipient problems in transformers, and condition monitoring of transformers using software-based diagnosis tools has become crucial. This research uses dissolved gas analysis as an intelligent fault classification of a transformer. The Multilayer SVM technique is used to determine the classification of faults and the name of the gas. The learned classifier in the multilayer SVM is trained with the training samples and can classify the state as normal or fault state, which contains six fault categories. In this paper, polynomial and Gaussian functions are utilized to assess the effectiveness of SVM diagnosis. The results demonstrate that the combination ratios and graphical representation technique is more suitable as a gas signature, and that the SVM with the Gaussian function outperforms the other kernel functions in diagnosis accuracy.

1. Introduction

In power systems, transformers play a critical role. Any malfunction in the transformer could result in an outage and, as a result, a disruption in service. It's critical to spot a transformer issue early on. The key-gas ratios C_2H_2/C_2H_4, CH_4/H_2, and C_2H_4/C_2H_6 have a strong relationship with transformer failures. The ability to forecast these key-gas ratios in transformer oil is critical for detecting and identifying transformer incipient breakdowns early. Because of its nonlinearity and the limited amount

of training data, forecasting key-gas ratios in transformer oil is a difficult challenge. The use of software-based diagnosis tools to monitor the condition of transformers is critical, and DGA is considered an excellent test for detecting incipient transformer defects. The examination of certain dissolved gas concentrations in a transformer's insulating oil provides information about the state of the transformer, allowing required preventive steps to be taken. Due to the variability of gas data, standard approaches fail to identify the faulty condition, making it a difficult and time-consuming task. The key gas analysis, Rogers Ratio technique, IEC gas ratio code (IEC-60599), Doernenburg Ratio method, and Duval triangle method are all traditional ways to diagnose a transformer issue utilizing the DGA method. These diagnostic approaches do not provide much information regarding incipient defects, which can lead to a state of indecision. The support vector machine with genetic algorithm (SVMG) developed by Sheng-wei Fei et al [1] is applied to fault diagnosis of a transformer, in which the genetic algorithm (GA) is used to select appropriate free parameters of SVM, and the experimental results showed that the SVMG method provided higher diagnostic accuracy. According to A. Akbari [2] et al, condition-monitoring and software-based diagnosis tools are effective maintenance management strategies for transformers. Agent-based systems have been developed for complex systems with relatively simple individual agents, and Multiagent systems have been used to overcome complexities. An intelligent fault classification approach to power transformer dissolved gas analysis (DGA) was given by A. Shintemirov et al [3] and Y.-j. Sun et al [4], which dealt with highly varied or noise-corrupted data. To increase the interpretation accuracy for DGA of power transformers, bootstrap, and genetic programming (GP) are used. For fault classification, the features retrieved by GP are fed into artificial neural networks (ANN), support vector machines (SVM), and K-nearest neighbor (KNN) classifiers. The combined classification accuracies are compared and determined to be good. Tang [5], [6] introduced a Parzen–Windows (PW)-based transformer failure diagnostic classifier that uses a probabilistic scheme to interpret transformer dissolved gas analysis (DGA). To increase fault classification accuracies, a global optimizer called particle swarm optimizer (PSO) is used to tune the parameters of PW.

For the fault detection of power transformers, C.P Hung et al introduced a unique cerebellar model articulation controller (CMAC) neural network (NN) technique. The results of the different standards are analysed using fuzzy logic and then compared with the empirical test. The comparison is made between the suggested methods and existing methods indicates the capability of the suggested method in the on-line fault diagnosis of transformers and the results are validated in the fault diagnosis of transformers. According to Rahmatollah Hooshmand [7, 8] et al. Sheng-wei Fei [9, 10] et al presented a support vector machine with genetic algorithm (SVMG) to forecast key-gas ratios in power transformer oil, and genetic algorithm (GA) is used to determine free parameters of support vector machine. The SVMG model is shown to be a proper alternative for forecasting key-gas ratios in power transformer oil. R.N. Afiqah [11, 12] stated that "Detection of new faults in the transformer early is very important to prevent accidents and to reduce related material losses," and "Conventional methods used in the implementation of DGA are improved by using intelligent systems." In this study, the IEC ratio method, one of the classical methods is used, and an intelligent fuzzy-based analysis is performed using MATLAB. The results of fuzzy logic and genuine faults are presented in the table, along with a comparison of the IEC approaches, LV Ganyun [13 - 17] et al. SVM stands for support vector machine, and it's a new machine learning [18,19] method based on statistical learning theory (SLT). For problems with little sampling, nonlinearity, and high dimension, SVM is a powerful tool. This paper is the one to use a multi-layer SVM classifier to diagnose transformer faults. The content of five diagnostic gases dissolved in oil, as determined by dissolved gas analysis (DGA), is preprocessed, and six features are extracted for SVMs. The aforesaid data processing is then used to train the multi-layer SVM classifier with the training examples collected. The trained classifier then identifies the four different types of transformer faults. The classifier performs well in terms of training speed and reliability, according to the test findings.

The SVM classification is utilized in this paper to assess the state and appropriate gas name signature. The difficulty of detecting the proper fault situation and the diagnostic accuracies for transformers under fault classifications are determined by these factors. The key gas, ratios, and graphical representation approaches are all DGA methods that are used as inputs to SVM

classifiers for fault classification.

The remainder of this paper is structured as follows: Section 2 presents DGA fundamentals. Section 3 explains concepts of SVM and its mathematics, section 4 details transformer faults classification based on SVM, section 5 gives results and discussions and finally Section 6 describes conclusions.

2. DGA Fundamentals

Because transformer equipment is so expensive, it must be properly monitored while in use. DGA is first used to monitor the status of a specific transformer and has since received widespread approval among professionals. DGA can be used to determine the most likely situation inside transformers, provide early warnings and diagnoses, and boost the chance of acting appropriately, similar to how doctors check a human body with a stethoscope. DGA is a sensitive and reliable technology for locating problems in power transformers [20-25]. It is possible to distinguish fault in a wide range of oil-filled equipment using this technique. A coded list of problems identifiable by DGA is provided in IEC Publication 60599. Partial discharge (PD), low energy discharge (D1), high energy discharge (D2), thermal faults T 300^0 C (T1), thermal faults 300^0 T 700 C (T2), and thermal faults T > 700^0 C are examples of transformer faults (T3). The key gas technique, IEC ratios method, and graphical representation method are the DGA interpretation methods used to detect the incipient failure.

The most important gas technique is the key Gas method. Five key gas concentrations are used in this method: hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), and ethane (C_2H_6) accessible for consistent fault interpretation. The diagnostic interpretations of the IEC 60599 for the five gases H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6 used to detect various types of problems are shown in the table.1 below.

Table 1. Diagnostic Interpretations of The IEC 60599

Gas detected	Interpretation
Oxygen(O2)	Transformer seal fault Oxide carbon (CO)
Oxide carbon	Cellulose decomposition
Dioxide carbon	Cellulose decomposition
Hydrogen(H_2)	Electric discharge (corona effect, low partial discharge
Acetylene(C_2H_2)	Electric fault (arc, spark)
Ethylene(C_2H_4)	Thermal fault (overheating local)
Ethane(C_2H_6)	Secondary indicator of thermal fault
Methane(CH_4)	Secondary indicator of an arc or serious overheating

The ppm concentration values range in the transformers according to *IEC* 60599 are given in Table.2.

TABLE 2. ppm concentration values range in the transformers according to *IEC 60599*

Gas	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	CO	CO_2
Concentration (ppm)	60–150	40–110	50–90	60–280	3–50	540–900	5100–13,000

The IEC Ratios method employs the following five gases: H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6, as well as the following gas ratios: C_2H_2/C_2H_4, CH_4/H_2, and C_2H_4/C_2H_6. The following table correlates fault categories to gas ratios, and decisions are made such that when key-gas ratios surpass their specific limitations, an incipient fault in the transformer has occurred. Diagnosis using the ratio method (IEC 599) is given in the table 3.

TABLE 3. Diagnosis using the ratio method (IEC 599)

Fault type	C_2H_2 / C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
PD	,0.1	<0.1	<0.2
D1	>1	0.1-0.5	>1
D2	0.6- 2.5	0.1-1	>2
T1	<0.1	>1	<1
T2	<0.1	>1	1-4
T3	<0.1	>1	>4

The approach of graphical depiction utilizing Duval's triangle is discussed as below.
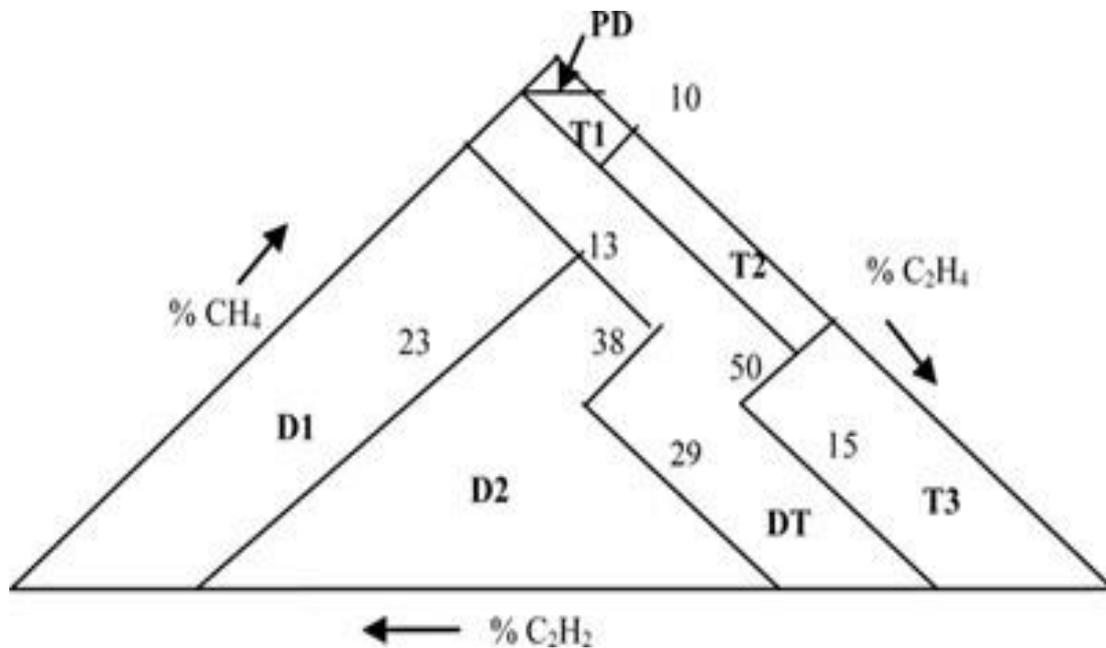

Figure 1. Coordinates and fault zones of the Duval's triangle [26].

TABLE 4. Representation Method for zone limits

Type of Fault Zone	Percentage of gases			
PD	98%CH_4	100% CH_4		
D1	23%C_2H_4	13% C_2H_2	100% C_2H_2	
D2	23% C_2H_4	13% C_2H_2	38% CH_4	29% C_2H_2
T1	4% C_2H_2	10% C_2H_4		
T2	4% C_2H_2	10% C_2H_4		
T3	15% C_2H_2	50% C_2H_4	100% C_2H_4	

The concentrations (in ppm) of CH_4, C_2H_2, and C_2H_4 are stated as a percentage of the total ($CH_4 + C_2H_4 + C_2H_2$) and characterized as a point (percent CH_4, percent C_2H_4, percent C_2H_2) in a coordinate system depicted as a triangular diagram shown in figure .1 with different zones. Each zone corresponds to a specific fault category.

Partial Discharge (PD) or Corona Discharge (CD): PD occurs in the gas phase of voids or gas bubbles. It is easily identified by DGA. As it is produced over long periods of time and within large volumes of paper insulation, it frequently generates large amounts of hydrogen.

Thermal faults T <300°C (TI): TI evidenced by paper turned brownish.

Thermal faults 3000 <T<700 °C (T2): T2 evidenced when paper carbonizes.

Thermal fault, (T3), Hot Spot, T> 700 °C (T3): T3 evidenced by oil carbonization, metal coloration or fusion.

Discharges of Low energy (D1), Low Energy Arcing and Tracking: D1 such as tracking, small arcs, and uninterrupted sparking discharges are usually easily detectable by DGA, because gas formation is large enough.

Discharges of High energy (D2), High Energy Arcing: D2 is evidenced by extensive carbonization, metal fusion and possible tripping of the equipment.

Thermal & Electrical fault (DT) or Hot Spot, T>400°C

Table.4. represents the percentage of gases in different zones.

3. Concepts of SVM and its Mathematics

The Support Vector Machine (SVM) is a supervised machine learning model that learns how to divide various groups by establishing decision boundaries and is used for classifications. It is used to provide an ideal separating hyper-plane solution for linearly separable and non-linearly separable datasets by maximizing the margin between the separating data. Hyperplanes are decision boundaries that aid in data classification. Different classifications can be given to data points on either side of the hyperplane as shown in the figure 2. The hyperplane's dimension is also determined by the number of features. If there are only two input characteristics, the hyperplane is merely a line. The hyperplane becomes a two-dimensional plane when the number of input features reaches three. $T = \{X_k, Y_k\} k^m$, where X_k signifies the input vector, Y_k ranges -1 to 1 denotes the associated output values, and m denotes the total number of data patterns, and the regression approximation estimates a function based on a given data set. It starts with data separated by a hyperplane and then uses the kernel method to extend to non-linear decision boundaries.
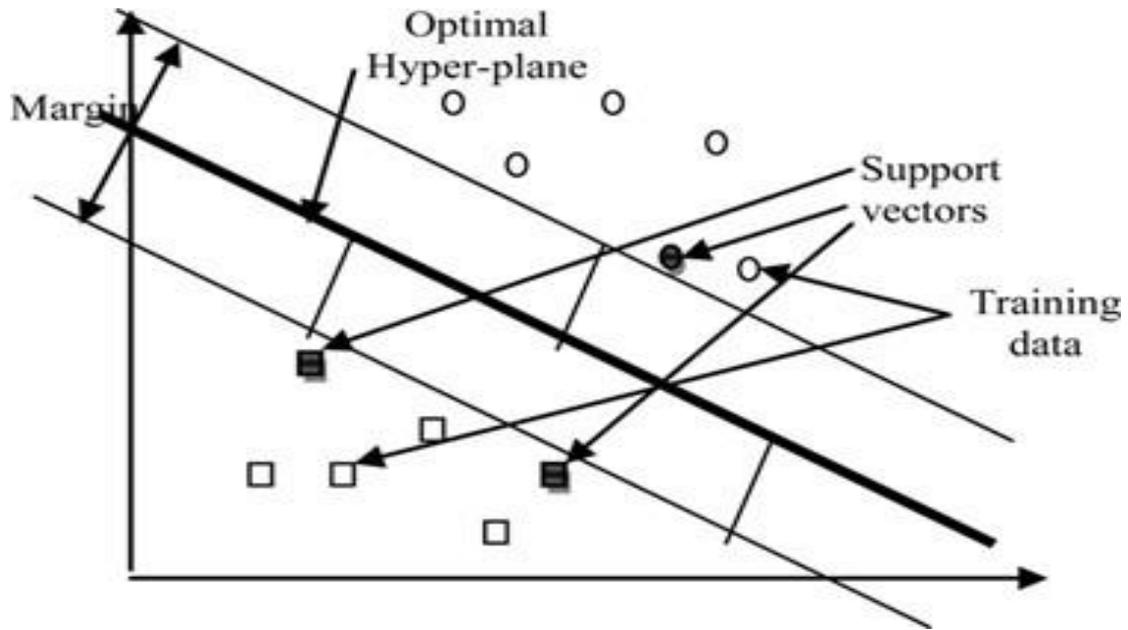
Figure 2. Hyper-plane in SVM [27].

$$W.X_k + b = 0 \qquad (1)$$

Where w denotes the weight vector and b denotes the bias term. w and b are used to define the position of the separating hyperplane by which it should satisfy the constraints.

$$Y_k\left(W.X_k + b\right) \geq 1, \; k = 1, 2, ...m$$

$$min \frac{1}{2}(\parallel W \parallel)^2 \qquad (2)$$

using the Lagrangian principle, the above equation is transformed to

$$L(W,b,\alpha) = \left(\frac{1}{2}\right)W^T W - \sum \alpha_k \left[Y_k \left(W^T.X_k + b\right) - 1 \right] \qquad (3)$$

where αk are the Lagrange coefficients ($\alpha_k > 0$)

Then the condition of optimality is applied.

$$\partial L\left(W,b,\alpha\right) / \partial W = 0, \; \partial L\left(W,b,\alpha\right) / \partial b = 0 \qquad (4)$$

We have the following equation along with the constraint:

$$W = \sum_{k=1}^{m} \alpha_k.X_k.Y_k = 0 \; , \; \sum \alpha_k.Y_k = 0 \qquad (5)$$

Hence the dual problem is expressed as the following equations

$$max \sum \alpha_k - \frac{1}{2}\sum \alpha_k.\alpha_j X_k.X_j Y_k.Y_j \; \text{ for all k and } \alpha_k \geq 0 \; \& \; \sum \alpha_k.Y_k = 0 \qquad (6)$$

The ranking function $class\left(x\right)$ is defined as

$$class\ (x) = sign\left[\left[(w_0 \cdot x)\ +\ b_0\right]\right] = sign\sum \alpha_i Y_i (X_i, X) + b) \tag{7}$$

we define that if $class\ (x)$ is less than 0, X is the $C = class = -1$,else it is a $C = class = 1$

For the nonlinear cases, a margin parameter C is added

$$\max \sum \alpha_k - \frac{1}{2} \sum \alpha_k . \alpha_j X_k . X_j Y_k . Y_j \ \text{ for all k and } 0 \le \alpha_k \le C \ \& \sum \alpha_k . Y_k = 0 \tag{8}$$

Where C is the margin parameter.

A kernel is a method of placing a two-dimensional plane into a higher dimensional space, so that it is curved in the higher dimensional space. In a simple way a kernel is a function from the low dimensional space into a higher dimensional space. A Kernel Trick is a simple method where a non-Linear data is projected onto a higher dimension space so as to make it easier to classify the data where it could be linearly divided by a plane. The selection of kernel function important to the support vector machine. The common kernel functions are linear kernel, polynomial kernel, Gaussian radial function kernel and sigmoid kernel function. The kernel function is defined as $K(X_k, X_j) = \phi(X_k).\phi(X_j)$, considered as positive kernel.

4. Transformer Faults Classification based on SVM

DGA training and testing data: It is based on the transformer oil's dissolved gas content data. The training data set and the testing data set are separated from the rest of the data. The training data set is assessed using several DGA methods, with judgments assigned to seven classes i.e normal, partial discharge, low energy discharge, high energy discharge, low temperature overheating, middle temperature overheating, and high temperature overheating. The transformer ratings and characteristics used for the taking of mineral oil samples are given in the table.5.

Because traditional fault detection cannot achieve high diagnosis accuracies, new methods using AI approaches are being explored. Different forms of defects can arise in transformers, necessitating a diagnosis of the transformer's individual state. A binary classifier, a single SVM classifier [28] can only classify input into two unique and opposite states. This classifier can be used to determine if a nonzero real integer is negative or positive. A multilayer SVM is used to categorize the various errors, with each layer's SVM being a binary classifier. The multilayer SVM is a "one-against-all" type of model. If the binary SVM in the preceding layer erroneously identifies the condition, the other SVMs will categorize the condition incorrectly, and the final classification result will be incorrect. It's difficult to detect less severe issues inside transformers with limited data and known conditions, thus it's always necessary to thoroughly inspect and evaluate conditions inside transformers for alarming or failure-prone conditions.

TABLE.5. Transformer Ratings

No	MVA rating	Voltage rating
1	150	235/15.5
2	4	15.5/6.8
3	100	2255/90
4	42.5	230/11
5	20	90/30

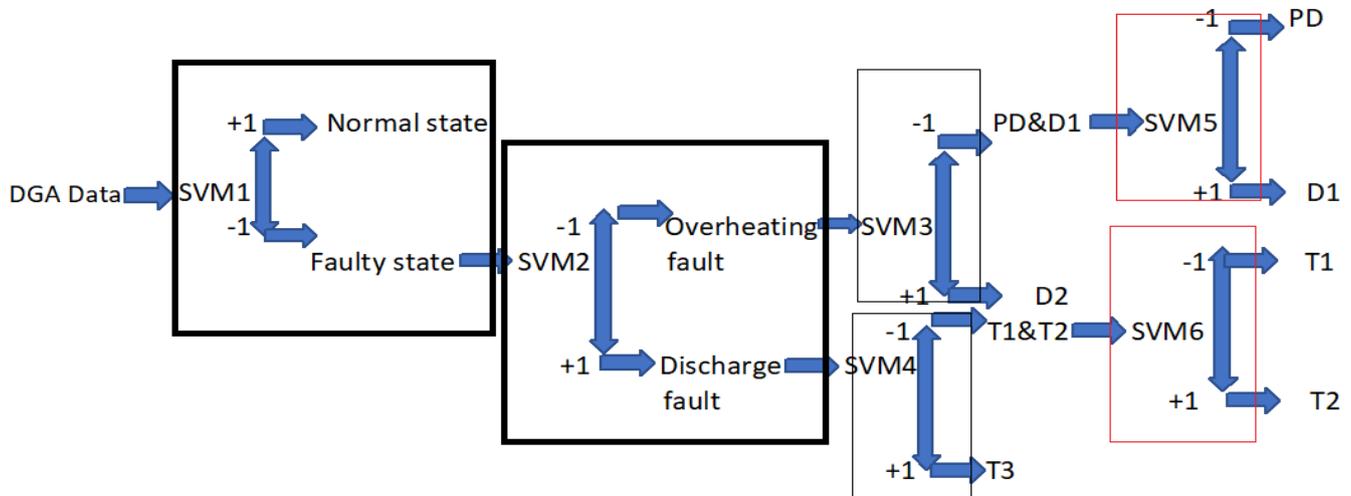## 4.1 The Functional Concept of Multi-Level SVM Classifier:

Figure 3. Functional concept of multi-level SVM classifier

Six SVM classifiers are utilized to identify the seven states, normal state (NS) and the six faults i.e partial discharge (PD), D1, D2, T1, T2, and T3. SVM1 is trained to distinguish between normal and fault states using all of the state training data shown in figure 3. When SVM1 receives a sample representing the normal condition as input, output is set to +1; otherwise, output is set to 1. The data is then passed to SVM2, which has been trained to distinguish between the discharge and overheating faults. The output of SVM2 is set to +1 when the input of SVM2 is a discharge fault; otherwise, it is set to 1. This information is passed to SVM3, which has been trained to distinguish between high-energy discharge (D2), partial discharge (PD), and low-energy discharge (D1) faults. The output of SVM3 is set to +1 when the input of SVM3 is a D2 fault; otherwise, it is set to 1. SVM4 is trained to distinguish between high temperature overheating (T3) and low and moderate temperature overheating (T1 and T2). When the input of SVM4 is a sample representing the T3 fault, the output of SVM5 is set to +1; otherwise −1. Then the it is fed to SVM5 is trained to separate the middle temperature overheating (T2) fault the low temperature overheating (T1) fault and when the input of SVM5 is a sample representing the T2 fault, the output of SVM5 is set to +1; otherwise −1. Then it is fed to SVM6 which is trained to separate the partial discharge (PD) fault from the low energy discharge (D1) fault. The output of SVM6 is set to +1 when the input of SVM6 is a sample indicating the D1 fault; otherwise, it is set to 1. The parameters of the SVM model, the activation function and C, the margin parameter, are optimized and chosen as the most appropriate parameters. The SVM model is trained using the optimum parameters. The kernel functions of all six SVMs are polynomial and Gaussian. The SVMs are run with a large number of epochs, which is defined as training the main SVM and hidden-layer SVM on their respective datasets only once. The bias values of all SVMs are determined by averaging over the mistakes, and this is employed with the gradient ascent technique, which uses a limited learning rate and a fixed number of iterations. The matrix condition output of SVMS is given in table.6.

TABLE 6. The matrix condition output of SVMS

Fault condition	SVM1	SVM2	SVM3	SVM4	SVM5	SVM6
Norma state	+1					
PD	-1	+1	-1			-1
D1	-1	+1	-1			+1
D2	-1	+1	+1			
T1	-1	-1		-1	-1	
T2	-1	-1		-1	+1	
T3	-1	-1		+1		

The Algorithm for the multi-layer SVM algorithm:
- *Initialize output SVM*
- *Initialize hidden-layer SVMs*
- *Compute kernel matrix for hidden-layer SVMs*
- *Train hidden layer SVMs on perturbed dataset*
- *repeat*
  - *Compute kernel matrix for output-layer SVM*
  - *Train output-layer SVM*
  - *Use backpropagation to create training sets for hidden-layer SVMs*
  - *Train hidden layer SVMs*
- *until maximum number of epochs is reached*

For classification, the training technique, and the selection of SVM parameters for training are critical. The process of optimizing SVM parameters with the cross-validation approach is depicted in the flow diagram given in the figure 4.
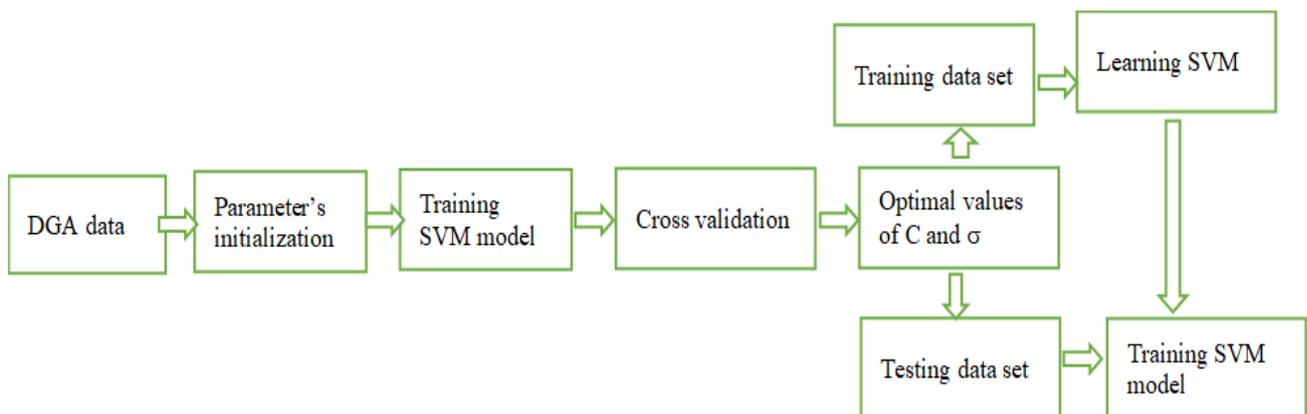

Figure 4. Functional concept of multi-level SVM classifier

Data generation: The data is generated by considering both a normal and a failed state mode.

• Load the 192-sample training data as training data.

• Polynomial and Gaussian kernel functions, with Kernel=1 and C=1

• The SVM training data set with the SVM class function.

• Display of the results as well as the most important support vectors.

• Determine the value of the hyper-parameter

• Using a logarithmic scale spanning from 1 to 1000, determine the C parameter.

• The cross-validation approach is used to select the C and parameters.

• Using the suitable values for these parameters, the SVMs are trained and tested.

5.   Results and Discussions

The classification of SVM errors is done using DGA algorithms as a gas signature. Statistical diagnosis results of the ratio methods is shown given table 7.

Table 7. Statistical diagnosis results of the ratio methods

Type	Samples	No. of correct diagnosis		
		Doernenburg	Roger	IEC
T1	20	10	7	0
T2	25	15	10	15
PD	9	4	3	2
D1	28	10	15	13
D2	45	37	32	42
NF	65	43	35	55
Total	192	119	102	127

5.1. SVM key gas

The defects are classified using the gas as input data. The words false alarm and non-detection rate are used to examine it. Polynomial and Gaussian kernels are put to the test, shown in table.8.

TABLE 8. The matrix condition output of SVMS

SVM kernel function	False alarm rate (%)	Non-detection rate (%)
Polynomial	0	45 (11/25)
Gaussian	0	16.7 (6/25)

The Gaussian kernel function is more efficient for system problem diagnostics, however it does not produce outstanding results. As a result, another method, SVM ratios, is proposed. The flaws in this method are dependent on the ratios that serve as input data. The test data is given back into the SVM to see if it correctly classifies and return results. For the two kernel functions polynomial and Gaussian kernels, the table.9 shows the results of the same false alarm rate and non-detection rate.

TABLE 9. Results of the false alarm rate and non-detection rate

SVM kernel function	False alarm rate (%)	Non-detection rate (%)
Polynomial	0	16.7 (6/25)
Gaussian	0	17.5 (7/30)

The Gaussian kernel function is found to be a more efficient approach for defect diagnostics.

5.2. Graphical representation of SVM

The defects are categorized using the same Polynomial and Gaussian kernel functions as input and the same graphical representation. The classification performance of the SVM graphical representation method is shown in the table.10.

TABLE 10. SVM graphical representation method

SVM kernel function	False alarm rate (%)	Non-detection rate (%)
Polynomial	0	23.3 (7/30)
Gaussian	0	16.7 (6/25)

In comparison to the Polynomial kernel function, the Gaussian kernel function provided a good diagnosis. The SVM combination method classification performance is given in the table.11.

TABLE 11. SVM combination method classification performance

SVM kernel function	False alarm rate (%)	Non-detection rate (%)
Polynomial	0	16.7 (6/25)
Gaussian	0	11(2/30)

5.3. Analysis on SVM Classification

The Gaussian kernel appears to be producing good results based on the four inputs of the data classified by the SVM. For the gas analysis approach, the false alarm rate and non-detection rate of four input data types are used. We analyse the false alarm rate and non-detection rate of four input data types to determine the most significant gas analysis method. The real result shows that the classification accuracies obtained by combining ratios and graphical representation methods are higher than those produced by SVM for gas signature classification. The performance of the multilayer SVM is compared to AI methods such as fuzzy logic and MLP. Though the MLP has the advantage of a quick learning process and no iteration for updating weights, it requires a significant quantity of training data and requires adjusting the hidden activation function's parameters. The Fuzzy logic method requires linguistic variables, membership functions for each gas signature with "low," "medium," and "high" descriptions, and an inference rule basis. Finally, the SVM technique is used to classify errors using a combination of ratios and graphical representation. Gaussian, trapezoidal, and triangle functions are the most utilized membership functions

to convert judgments into numerical expressions. MLP is based on a four-layer design with six input nodes, seven output nodes, and 19 hidden nodes. 6 input variables with 3 membership functions each, 7 output variables with 3 membership functions, and 353 inference rules make up the Fuzzy method. The results demonstrate that the multilevel SVM technique provides a reliable and accurate fault diagnosis transformer. Comparison between various DGA methods with false alarm and non-detection rate is given in table 12.

It can be observed that the IEC ratio method gives correct diagnosis and finds No Fault (NF) cases but T1 fault cases of cannot be detected. The other two methods Doernenburg ratio and the Roger ratio methods are able to detect the T1 cases. Out of the Doernenburg ratio and the Roger ratio methods, the Roger ratio method finds good number of discharge faults, and the Doernenburg ratio method is the best to detect the thermal faults.

TABLE 12. Comparison between various DGA methods

DGA methods	False alarm rate (%)	Non-detection rate (%)
Key gas	0	16.7(6/25)
Ratios method	0	11(2/30)
Graphical method	0	16.7 (6/25)
Combination of ratios and graphical representation	0	11(2/30)

The test of results for the AI-based methods, based on 45 total sets of testing data shown in the in table 13 and based on the results the diagnosis accuracies of these methods are compared. The accuracies from the traditional ratio methods are based on all 192 datasets, whereas the accuracies from the AI-based methods are only based on 88 testing datasets.

Table 13. Diagnosis results from ANN, SVM and MLSVM methods

Method applied	Scenario1	Scenario2	Scenario3	Scenario4	Scenario5	Diagnosis accuracy
MLSVM	64	65	65	69	71	66.8
SVM	55	64	59	68	59	61
ANN	58	48	64	58	66	58.8

Comparison between Multi-layer SVM model and other methods is shown in figure 5. The comparison results given table 13 shows that the multi-layer SVM model can diagnose an average of 66.8 cases, the SVM models can correctly identify an average of 61 cases and ANN method identify 58.8 cases and this comparison confirms that of the multi-layer SVM is superior to all methods considered and this explains the effectiveness of the Multilevel SVM for the fault diagnosis of transformers.
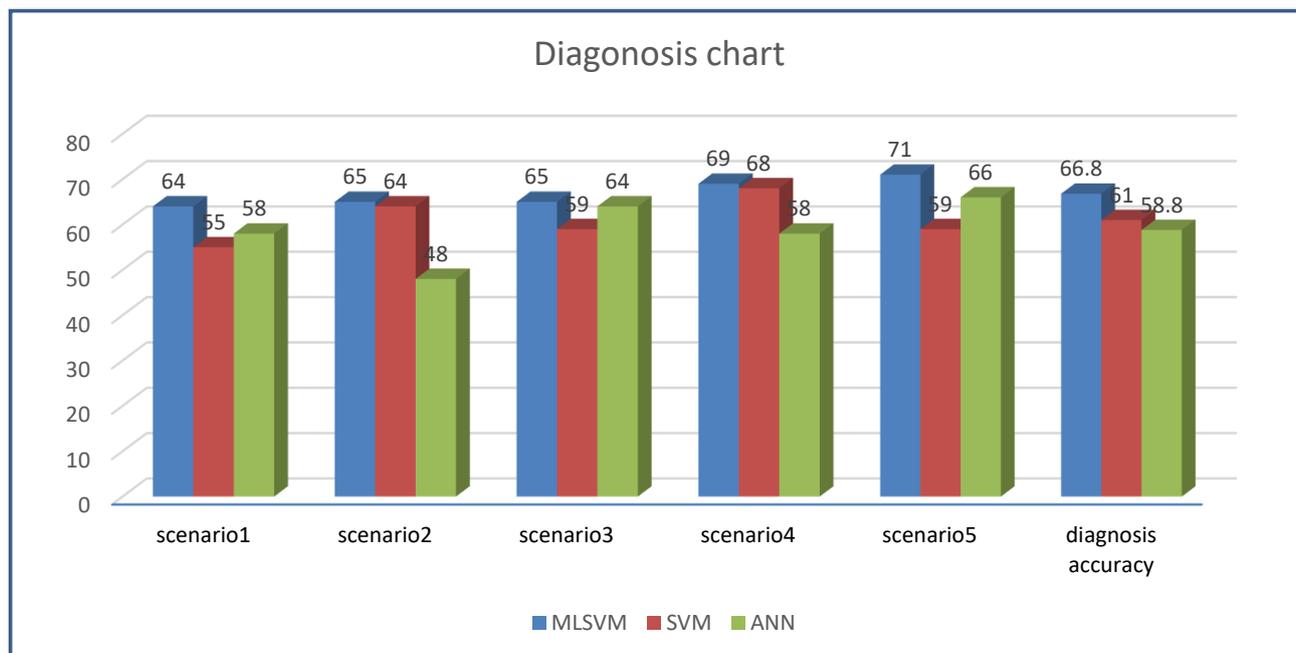
Figure 5. Comparison Between Multi-layer SVM Model and other methods

The procedure followed for the comparison is as follows. Randomly selected the required number of datasets from the 180 samples to develop models for the ANN method, the SVM method, and the proposed Multilayer SVM   method. The remaining samples are the testing sets. Randomly selected 10 datasets from the testing sets are used in this case study. Obtained the diagnosis results from the models developed in the first step. Fed the 10 datasets into the ratio method to get diagnosis results. Summarized and compared the results. Comparison of diagnosis results is shown in table 14. Selected gas concen*tration is in ppm.*

Table 14. Comparison of diagnosis results

S. No.	Source	H_2	CO	CH_4	C_2H_4	C_2H_6	C_2H_2
1	Duval	0.05	3900	18900	540	410	330
2	Duval	960	15800	4000	1560	1290	6
4	Duval	1100	0.05	1600	2010	221	26
5	Duval	3910	1800	4290	6040	626	1230
6	Duval	92600	6400	10200	0.05	0.05	0.05
7	Duval	26788	704	18342	27	2111	0.05
8	Duval	60	780	10	4	4	4
9	Duval	6870	29	1028	900	79	5500
10	Duval	5100	117	1430	1140	0.05	1010

Table 15. Diagnosis results from the different methods for the selected data

S. No.	Actualfault	Doernenburg	Rogers	IEC	ANN	SVM	MLSVM
1	T1	T1&T2	ND	ND	NF	T2	T1
2	T1	T1&T2	T1	T1	T1	T1	T1
3	T1	T1&T2	T1	ND	T1	NF	T1
4	T2	T1&T2	T2	T2	T2	T2	T2
5	T2	T1&T2	ND	ND	D2	T2	T2
6	PD	ND	ND	ND	PD	PD	PD
7	PD	ND	NF	ND	NF	T1	T1
8	D1	ND	ND	ND	D1	D1	D1
9	D1	D1&D2	D1	D1	D2	D1	D1
10	D2	D1&D2	D2	D2	T2	D2	D2

Form the table 15, it should be noted that as in case 6 is an example, although the Doernenburg ratio method is not able detect the PD cases, the MLSVM can identify PD. Similar to the cases 1, 3 and 8 accurately, while the SVM model failed to do so, which shows the effectiveness of the MLSVM.

6. Conclusion

The MLSVM technique is used and compared with basic SVM and conventional IEC methods in this paper for fault classification in transformers utilizing dissolved gas measurements. The key gas, ratios, graphical representation, and combination ratios and graphical representation approach are the DGA methodologies investigated. Polynomial and Gaussian functions are used to examine the effectiveness of SVM diagnosis. The real data sets are utilized to test the DGA algorithms' capability in forecasting transformer oil. According to test results, the combination, ratios and graphical representation approach is more suitable as a gas signature and the MLSVM with the Gaussian function performs better in diagnosis accuracy than the other kernel functions. Due to their extensive study capabilities, the accuracy of multi-layer SVM for fault identification is comparable to that of conventional approaches. It can be observed that conventional ratio methods are not able to detect the internal faults and hence given Not Detected (ND) and some cases which cannot be diagnosed are mentioned as No Fault (NF) cases. It is observed that the ratios methods are inferior to the MLSVM method. The MLSVM outperforms previous AI approaches when it comes to fault diagnostics. The MLSVM method can be used to diagnose incipient defects in transformers in real time. The results show that the MLSVM method has the ability to anticipate the DGA method in transformer oil. The SVM method has a greater precision but a poorer recall over positive classes, whereas the MLP approach has a better classification performance, correctly predicting two of the four failures that happened within the chosen period. Future work will focus on improving the overall performance metric by incorporating a more robust data set as well as a different set of features to improve and lessen the potential for bias in the results. It will be highly promising in the future to develop new intelligent comprehensive fault diagnostic systems by adding new ML theories and frameworks, as well as a new ML based on multi-layer ANN to transformer fault diagnosis based on DGA. In some circumstances, such systems can detect and discard bad data automatically, and they have better real-time capabilities and self-adaptation.

References

[1] S.-W. Fei and X.-B. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Syst. Appl.*, vol. 36, no. 8, pp. 11352–11357, 2009.

[2] A. Akbari, A. Setayeshmehr, H. Borsi, E. Gockenbach, and I. Fofana, "Intelligentagent based system using dissolved gas analysis to detect incipient faults in power transformers," *IEEE Electrical Insulation Magazine*, vol. 26, no. 6, pp. 27–40, 2010.

[3] A. Shintemirov, W. Tang, and Q. H. Wu, "Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming," *IEEE Trans. Syst. Man Cybern. C Appl. Rev.*, vol. 39, no. 1, pp. 69–79, 2009.

[4] Y.-J. Sun, S. Zhang, C.-X. Miao, and J.-M. Li, "Improved BP neural network for transformer fault diagnosis," *J. China Univ. Min. Technol. (Eng. Ed.)*, vol. 17, no. 1, pp. 138–142, 2007.

[5] Z. J. Richardson, J. Fitch, W. H. Tang, J. Y. Goulermas, and Q. H. Wu, "A probabilistic classifier for transformer dissolved gas analysis with a particle swarm optimizer," *IEEE trans. power deliv.*, vol. 23, no. 2, pp. 751–759, 2008.

[6] C.-P. Hung and M.-H. Wang, "Diagnosis of incipient faults in power transformers using CMAC neural network approach," *Electric Power Syst. Res.*, vol. 71, no. 3, pp. 235–244, 2004.

[7] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," in IEEE Electrical Insulation Magazine, vol. 18, no. 3, pp. 8-17, May-June 2002, doi: 10.1109/MEI.2002.1014963..

[8] R. Hooshmand, M. Banejad, "Application of fuzzy logic in fault diagnosis in transformers using dissolved gas based on different standards," World Academy of Science, Engineering and Technology, 17,pp.157–161,2006.

[9] V. Miranda and A. R. G. Castro, "Improving the IEC table for transformer failure diagnosis with knowledge extraction from neural networks," *IEEE trans. power deliv.*, vol. 20, no. 4, pp. 2509–2516, 2005.

[10] S.-W. Fei, C.-L. Liu, and Y.-B. Miao, "Support vector machine with genetic algorithm for forecasting of key-gas ratios in oil-immersed transformer," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6326–6331, 2009.

[11] Standard IEC 60599, Guide for the interpretation of dissolved gas analysis and gas-free, 2007.

[12] R.N. Afiqah, I. Musirin, D. Johari, M.M. Othman, T.K.A. Rahman, Z. Othman, Fuzzy logic application in DGA methods to classify fault type in power transformer, Selected Topics in Power Systems and Remote Sensing, Malaysia ,2008.

[13] W. C. Flores, E. E. Mombello, J. A. Jardini, G. Rattá, and A. M. Corvo, "Expert system for the assessment of power transformer insulation condition based on type-2 fuzzy logic systems," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8119–8127, 2011.

[14] R. Ziani, A. Felkaoui, R. Zegadi, Performances de la classification par les Séparateurs à Vaste Marge (SVM): application au diagnostic vibratoire automatisé, In the Proceedings of the 4th International Conference on Computer Integrated Manufacturing CIP'2007, 03–04 November, 2007.

[15] L. V. Ganyun, C. Haozhong, Z. Haibao, and D. Lixin, "Fault diagnosis of power transformer based on multi-layer SVM classifier," *Electric Power Syst. Res.*, vol. 74, no. 1, pp. 1–7, 2005.

[16] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.

[17] W.-M. Lin, C.-H. Lin, and M.-X. Tasy, "Transformer-fault diagnosis by integrating field data and standard codes with training enhancible adaptive probabilistic network," *IEE Proc. - Gener. Transm. Distrib.*, vol. 152, no. 3, p. 335, 2005.

[18] S. Ghosh and S. Dutta, "Ensemble machine learning methods for better dynamic assessment of transformer status," *J. Inst. Eng. (India) Ser. B*, vol. 102, no. 5, pp. 1113–1122, 2021., https://doi.org/10.1007/s40031-021-00599-1.

[19] B. Zeng *et al.*, "Prediction model for dissolved gas concentration in transformer oil based on modified grey wolf optimizer and LSSVM with grey relational analysis and empirical mode decomposition," *Energies*, vol. 13, no. 2, p. 422, 2020, https://doi:10.3390/en13020422.

[20] R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis," *IEEE trans. electr. insul.*, vol. EI-13, no. 5, pp. 349–354, 1978, https://doi.org/10.1109/TEI.1978.298141.

[21] A.-M. Aciu, C.-I. Nicola, M. Nicola, and M.-C. Niṭu, "Complementary analysis for DGA based on Duval methods and furan compounds using artificial neural networks," *Energies*, vol. 14, no. 3, p. 588, 2021. https://doi.org/10.3390/en14030588.

[22] E. T. Mharakurwa, G. N. Nyakoe, and A. O. Akumu, "Power transformer fault severity estimation based on dissolved gas analysis and energy of fault formation technique," *J. Electr. Comput. Eng.*, vol. 2019, pp. 1–10, 2019. https://doi.org/10.1155/2019/9674054.

[23] A. Abu-Siada, "Improved consistent interpretation approach of fault type within power transformers using dissolved gas analysis and gene Expression Programming," *Energies*, vol. 12, no. 4, p. 730, 2019. https://doi.org/10.3390/en12040730.

[24] N. Poonnoy, C. Suwanasri, and T. Suwanasri, "Fuzzy logic approach to dissolved gas analysis for power transformer failure index and fault identification," *Energies*, vol. 14, no. 1, p. 36, 2020. https://dx.doi.org/10.3390/en14010036.

[25] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 2, pp. 1239–1248, 2017, https://dx.doi.org/ 10.1109/TDEI.2017.005959.

[26] J. Faiz and M. Soleimani, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 25, no. 5, pp. 1798–1806, 2018, doi: https://dx.doi.org/10.1109/TDEI.2018.007191.

[27] E. García-Gonzalo, Z. Fernández-Muñiz, P. J. García Nieto, A. Bernardo Sánchez, and M. Menéndez Fernández, "Hard-rock stability analysis for span design in entry-type excavations with learning classifiers," *Materials (Basel)*, vol. 9, no. 7, p. 531, 2016. doi: https://dx.doi.org/10.3390/ma9070531.

[28] J. Li, Q. Zhang, K. Wang, J. Wang, T. Zhou, and Y. Zhang, "Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 23, no. 2, pp. 1198–1206, 2016, doi: https://dx.doi.org/10.1109/TDEI.2015.005277.