



## IMPLEMENTATION OF SVM USING SEQUENTIAL MINIMAL OPTIMIZATION FOR POWER TRANSFORMER FAULT ANALYSIS USING DGA

V.Gomathy, Dr.S.Sumathi

Research Scholar, EEE Department, PSG college of Technology, India

[gomathyvelmayil@gmail.com](mailto:gomathyvelmayil@gmail.com)

Associate Professor, EEE Department, PSG college of Technology, India

[ssi@eee.psgtech.edu](mailto:ssi@eee.psgtech.edu)

### ABSTRACT

Reliable operations of power transformers are necessary for effective transmission and distribution of power supply. During normal functions of the power transformer, distinct types of faults occurs due to insulation failure, oil aging products, overheating of windings, etc., affect the continuity of power supply thus leading to serious economic losses. To avoid interruptions in the power supply, various software fault diagnosis approaches are developed to detect faults in the power transformer and eliminate the impacts. SVM and SVM-SMO are the software fault diagnostic techniques developed in this paper for the continuous monitoring and analysis of faults in the power transformer. The SVM algorithm is faster, conceptually simple and easy to implement with better scaling properties for few training samples. The performances of SVM for large training samples are complex, subtle and difficult to implement. In order to obtain better fault diagnosis of large training data, SVM is optimized with SMO technique to achieve high interpretation accuracy in fault analysis of power transformer. The proposed methods use Dissolved Gas-in-oil Analysis (DGA) data set obtained from 500 KV main transformers of Pingguo Substation in South China Electric Power Company. DGA is an important tool for diagnosis and detection of incipient faults in the power transformers. The Gas Chromatograph (GC) is one of the traditional methods of DGA, utilized to choose the most appropriate gas signatures dissolved in transformer oil to detect types of faults in the transformer. The simulations are carried out in MATLAB software with an Intel core 3 processor with speed of 3 GHZ and 2 GB RAM PC. The results obtained by optimized SVM and SVM-SMO are compared with the existing SVM classification techniques. The test results indicate that the SVM-SMO approach significantly improve the classification accuracy and computational time for power transformer fault classification.

### KEYWORDS

Support Vector Machine, Sequential Minimal Optimization, Dissolved Gas Analysis, Gas Chromatograph, Power Transformer, Classification accuracy, Computational time.

## Council for Innovative Research

Peer Review Research Publishing System

**Journal:** INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 10, No 5

[editor@cirworld.com](mailto:editor@cirworld.com)

[www.cirworld.com](http://www.cirworld.com), [member.cirworld.com](http://member.cirworld.com)



## 1. INTRODUCTION

Power transformers are normally exposed to electrical, mechanical, thermal and environmental stresses that degrade their insulation quality. To avoid the power failure, the development of periodic monitoring and accurate diagnosis system for transformers are necessary. Detecting faults in the transformer at an early stage helps in large savings of operation costs, maintenance costs and prevent premature breakdown or failure. There are several routine maintenance procedures for power transformers such as Dissolved Gas Analysis (DGA), moisture analysis in transformer oil [1, 2], oil breakdown voltage test, the tan ( $\delta$ ) test, resistivity test, acidity test, sludge test, interfacial tension test and partial discharge (PD) acoustic emission sensing. Among these methods, DGA is an important tool for the early detection and diagnosis of incipient faults for transformers [3]. It is well known that overheating, arcing, partial discharge, winding circulating currents, and continuous sparking are the main factors in deteriorating transformer condition. These phenomena develop certain dissolved gases in the insulation oil. The gases include hydrocarbons such as: methane (CH<sub>4</sub>), ethane (C<sub>2</sub>H<sub>6</sub>), ethylene (C<sub>2</sub>H<sub>4</sub>), acetylene (C<sub>2</sub>H<sub>2</sub>) and others such as: hydrogen (H<sub>2</sub>), carbon dioxide (CO<sub>2</sub>), etc. The gases are extracted from the oil under high vacuum and analyzed by Gas Chromatograph (GC), to get individual gas concentrations. By interpretation of the gas contents, the developing faults in the power transformers can be diagnosed. Many diagnostic criteria have been developed to establish relationships between the gases and the fault conditions. The gas concentrations, generation rates, specific gas ratios, and the total combustible gas are important parameters for interpreting the result of DGA. To facilitate the procedure of power transformer fault classification, algorithms like Modified Differential Evolution [3], Multiclass SVM [18], Fast algorithm [22], Self adaptive RBF NN [25], Artificial Neural Network [20], K Nearest Neighbor [26], Support Vector Machine [19] and Radial Basis Function [27] have been presented in literature.

Presently, the conventional ratio methods, statistical schemes and Artificial Intelligence (AI) methods are the major interpreting approaches for power transformer fault analysis. The conventional ratio methods mainly include Rogers Ratios [4], Duval Triangle [5], and International Electrotechnical Commission Ratios [IEC] [6]. Since conventional ratios' boundaries are sharp, they are unable to provide interpretation for every possible combination of ratio values [7]. The Artificial Neural Network methods have also been used to explore the nonlinear and complex relation between the gases concentration and the type of faults. Multilayer back propagation (MLP) [8], [9], self-organizing map network [10], Adaptive Back-propagation learning algorithm [11] and Extension NN [12, 13] are important classification algorithms used for power transformer fault classification. ANN training suffers from trapping in local minima; therefore evolutionary training algorithms have been excelling in this field [7], [14], [15]. Other methods that have been investigated are wavelet decomposition [15], SVM [15], KNN [15] and fuzzy learning vector quantization network [16], [12]. In this paper SVM and SVM-SMO are considered for fault classification in power transformer. SVM has enhanced theory, excellent performance analysis and is more effective in solving problems having small samples. It is non linear and has high dimension because of which it has attracted widespread attention towards fault analysis of transformers. The feature extraction is quadratic in the operation with dataset, while using SVM classifier. The solution for very large quadratic programming (QP) optimization problem is obtained by training the SVM. This large QP problem is broken into smallest possible series of QP problems by SMO [22]. As QP problems are analytically handled by SMO, time consuming numerical optimization is eliminated. SMO manipulates training set in utilizing the memory needed for SMO and determine the result 1000 times faster than conventional algorithms.

The paper is organized as follows: Section 2 presents a brief mathematical description about fault analysis and section 3 describes proposed methodology of SVM, SMO and SVM-SMO for fault classification in power transformers, Section 4 provides the experimental results, data set, data preprocessing and feature extraction, Section 5 draws the conclusion.

## 2. PROBLEM FORMULATION

The main objective of the proposed problem is to separate the normal state and fault state from the given input samples. The linear classifier separates the data, but it maximizes the distance between the given data point to the nearest data point of each class.

The training data set is given by

$$D = \{(x_1, y_1)\Lambda, (x_i, y_i)\Lambda, (x_l, y_l)\}, \quad x \in R^n, y \in \{-1,1\} \quad (1)$$

where,  $l$  – number of training data,

$X_i$  – Training data,

$y_i$  – class label as 1 or -1 for  $x_i$ .

A nonlinear function is adopted to map the original input space  $R^n$  into  $N$ -dimensional feature space.

$$\psi(x) = \varphi_1(x), \varphi_2(x), \dots, \varphi_N(x) \quad (2)$$

The separating hyper plane is developed in this  $N$ -dimensional feature space. Then the classifier is represented as,

$$y(x) = \text{sgn}(\omega \cdot \psi(x) + b) \quad (3)$$

where,  $\omega$  - weight vector and  $b$ - scalar.

In order to obtain the optimal classifier,  $\|\omega\|$  should be minimized subject to the following constraints

$$y_i[\varphi(x_i) \cdot \omega + b] \geq 1 - \xi_i, \quad i = 1, 2, \dots, l. \quad (4)$$

The variable  $\xi_i$  is the positive slack variables, necessary for misclassification.

### 3. PROPOSED METHODOLOGY

The classification of faults is effective by optimizing the kernel function of SVM with SMO. The performances of the classifiers are evaluated using classification accuracy and computational time. SVM-SMO classifier 1 separates the normal state from fault state, and then SVM-SMO classifier 2 is employed to present the difference between thermal heating and discharge type of the fault data, finally SVM-SMO classifier 3 classify the discharge data into low energy discharge and high energy discharge. Thus four types of transformer faults are identified by extracting features, training the classifiers and testing the trained networks. The overall methodology of proposed work is given in Figure 1.

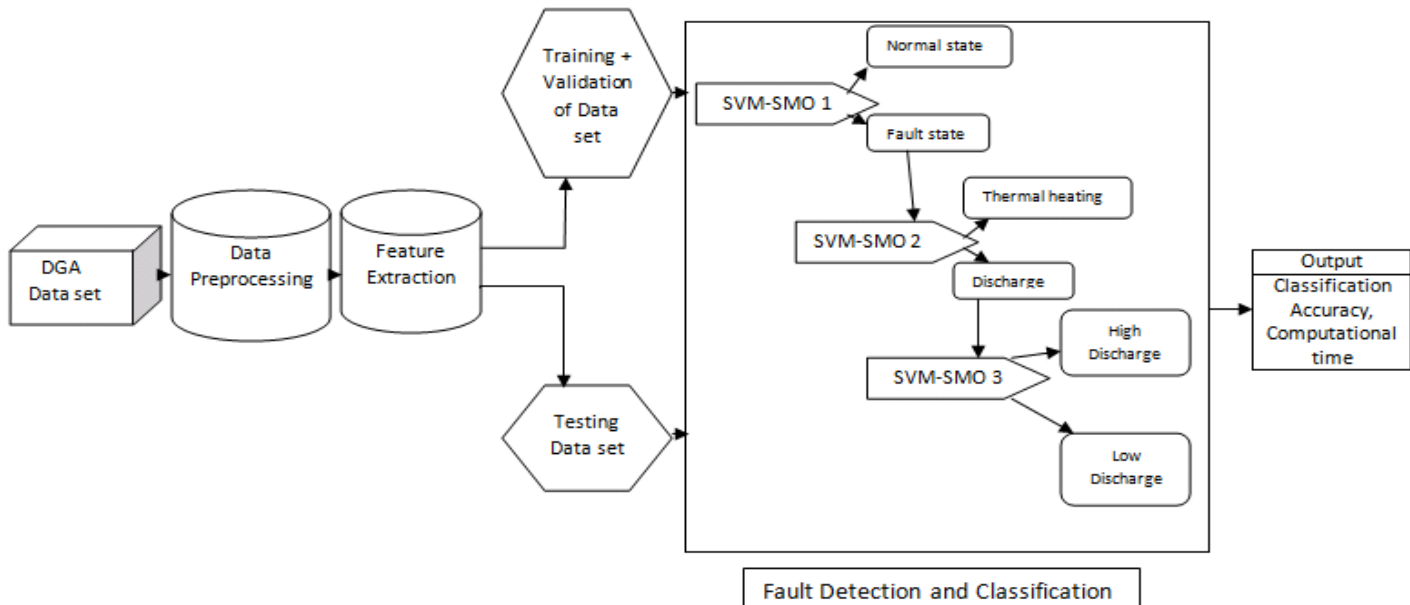


Figure 1 Block diagram of proposed method

#### 3.1 SVM classifier

Support Vector Machine is a machine learning approach that originates from statistical learning theory that is used to solve the binary classification problems. The primary SVM features utilize input data and classify them into two possible classes which form the output, making it a non-probabilistic binary linear classifier. The number of support vectors from the training samples are determined by SVM and converted into feature space using kernel functions each marked as belonging to normal state or fault state [23]. To estimate conditional expectation, kernel parameters are used in kernel density estimation and the cost factor decides weighting function between the positive data and negative data. SVM based fault diagnosis involves three steps:

- Extracting features including fault information of power transformer
- Training the SVM
- Identifying the fault types of power transformer using trained classifier.

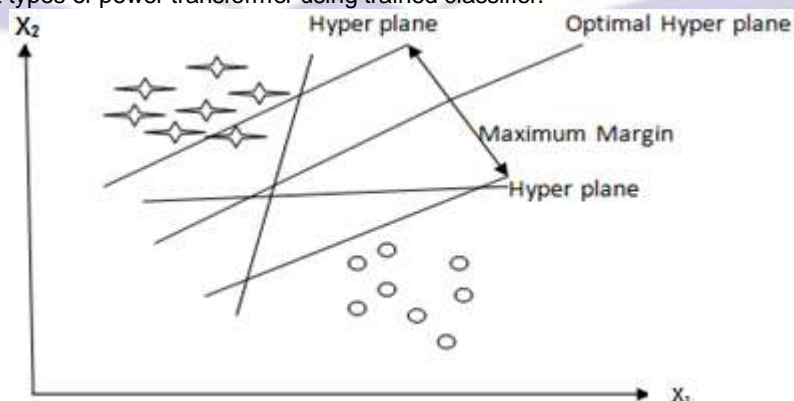


Figure 2 Structure of optimal separating hyper plane of SVM

Consider the classification of two classes of patterns that are linearly separable, *i.e.*, a linear classifier can perfectly separate the classes. The operation of SVM algorithm is based on finding the hyper planes that gives the distance



between the training samples. The definition of Optimum Separation Hyper plane (OSH) is the linear classifier with the maximum margin for a given finite set of learning patterns. Figure 2 shows the optimal hyper plane (i.e., distances between hyper planes) is the linear classifier with the maximum margin for exact classification of the input patterns. The objective function is given in Eq. (5),

$$\min J(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i \tag{5}$$

we have,

$$y_i [\varphi(x_i) * \omega + b] \geq 1 - \xi_i \tag{6}$$

where, C – margin parameter,  $\omega$  - weight vector,  $x_i$  - training data,  $y_i$  - class label (1 or -1) for  $x_i$ ,  $\xi_i$  - positive slack variables;  $\xi_i \geq 0, i = 1, \dots, l$ , b – scalar, l – number of training data.

Objective function obeys the principle of structural risk minimization in order to obtain the optimal solution [25]. The objective function in Eqn (5) can be re-modified by following Lagrangian principle as,

$$L(\omega, b, \xi, a, \gamma) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l a_i (y_i [\varphi(x_i) * \omega + b] - 1 + \xi_i) - \sum_{i=1}^l \gamma_i \xi_i \tag{7}$$

Where,  $a_i \geq 0, \gamma_i \geq 0 (i = 1, 2, \dots, l)$ ,  $a_i, \gamma_i$  - Lagrange multipliers

The optimality conditions are given in Eq. (8)

$$\left. \begin{aligned} \frac{\partial L}{\partial \omega} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi} = 0 \\ \sum_{i=1}^l a_i y_i = 0 \\ \omega = \sum a_i y_i \varphi(x_i) \\ C - a_i - \gamma_i = 0 \end{aligned} \right\} \tag{8}$$

On substituting Eq. (8) in Eq. (7), the dual problem becomes,

$$\max W(a) = -\frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j (\varphi(x_i), \varphi(x_j)) + \sum_{i=1}^l a_i \tag{9}$$

$$\max W(a) = -\frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j k(x_i, x_j) + \sum_{i=1}^l a_i \tag{10}$$

$$\text{Kernel function, } k(x_i, x_j) = \varphi(x_i), \varphi(x_j) \tag{11}$$

$$\text{Also, } \sum_{i=1}^l a_i y_i = 0, 0 \leq a_i \leq C, i = 1, \dots, l. \tag{12}$$

The Eq. (12) is a quadratic programming problem subject to the constraints of convex set and the unique solution is obtained.

$$\sum_{sv} a_i y_i k(x_i, x) + b = 0 \tag{13}$$

The optimal separating hyper plane is given in Eq. (13)

Where, sv – support vectors, k – kernel parameter.

The non-linear SVM classifier is

$$Y = \text{sgn}(\sum_{sv} a_i y_i k(x_i, x) + b) \tag{14}$$

On close observation of all the analysis, the training samples and kernel function are important factors that determine classifications using SVM [18]. In order to train SVM classifier, cost factor (C) and the kernel parameter (k) are utilized to





factorize kernel function. The most common kernel functions are Gaussian kernel, linear kernel and Radial Basis function kernel.

- The Gaussian basis function is defined as

$$K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (15)$$

where,  $\sigma$  - standard deviation.

- The Linear kernel function is a type of polynomial kernel function that is defined as

$$K(x, y) = (x^T y + 1)^d \quad (16)$$

where,  $d = 1$ , since the function is linear.

- The Radial basis kernel function is defined as,

$$K(x, y) = -e^{-\sigma\|x-y\|^2} \quad (17)$$

where,  $\sigma$  - standard deviation,  $x$  - attribute value,  $y$  - label value [26].

Only the best kernel function yields minimum error and highest classification accuracy. Support vector classification with Gaussian RBF kernel is sensitive to the kernel width. Small kernel width may cause over-fitting, and large one causes under-fitting, so optimal kernel width is merely selected based on the tradeoff between under-fitting loss and over-fitting loss. To reduce the coexisting over-fitting and under-fitting loss in support vector classification, Gaussian RBF kernel is adopted to obtain good feature space distribution than other kernel functions. Hence Gaussian kernel function is accepted as the kernel function in this work for better fault classification applications [23, 24].

### 3.2 Sequential Minimal Optimization

Sequential Minimal Optimization (SMO) solves the SVM QP problem without any extra matrix storage and without invoking an iterative numerical routine for each sub-problem. Unlike traditional methods, SMO chooses to solve the smallest possible optimization problem. For the standard SVM QP problem, the smallest possible optimization problem involves two Lagrange multipliers, which is used for optimization of SVM to obey a linear equality constraint. At every step, SMO has been employed to resolve the QP problems by classifying them as sub problems thus the entire iteration due to numerical QP optimization is avoided. Execution of SMO Algorithm thus solves overall QP problems.

Consider a binary classification problem with a dataset  $(x_1, y_1) \dots (x_n, y_n)$ , where  $x_i$  is an input vector and  $y_i \in \{-1, +1\}$  is a binary label corresponding to it. A soft-margin support vector machine is trained by solving a quadratic programming problem, which is expressed in the dual form as follows:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j K(x_i, x_j) \alpha_i \alpha_j, \quad (18)$$

Subject to:

$$0 \leq \alpha_i \leq C, \text{ for } i=1, 2, \dots, n, \quad (19)$$

$$\sum_{i=1}^n y_i \alpha_i = 0 \quad (20)$$

Where  $C$  is an SVM hyper parameter and  $K(x_i, x_j)$  is the kernel function, both supplied by the user and the variables,  $\alpha_i, \alpha_j$  are Lagrange multipliers.

#### 3.2.1 Algorithm:

SMO is an iterative algorithm for solving the optimization problem described through Lagrange multipliers that confirm to linear equality constraints. A series of smallest possible sub-problems are solved analytically by SMO involving two multipliers. For two multipliers  $\alpha_1$  and  $\alpha_2$ , the constraints are reduced to:

$$0 \leq \alpha_1, \alpha_2 \leq C, \quad (21)$$

$$y_1 \alpha_1 + y_2 \alpha_2 = k \quad (22)$$

Eqs. (21) and (22) can be solved analytically.

The algorithm proceeds as follows:

1. Find a Lagrange multiplier  $\alpha_1$  that violates the Karush–Kuhn–Tucker (KKT) conditions given by

$$L[x, \lambda, \eta] = F(x) + \lambda h(x) + \xi \lambda (g_i(x)) \tag{23}$$

Recall the geometry of the Lagrange multiplier conditions: The gradient of the objective function must be orthogonal to the tangent plane of the (active) constraints i.e., the projection of the gradient of  $f$  onto the space of directions tangent to the constraint "surface" is zero.

2. Pick a second multiplier  $\alpha_2$  and optimize the pair  $(\alpha_1, \alpha_2)$ .
3. Repeat steps 1 and 2 until convergence.

When all the Lagrange multipliers satisfy the KKT condition, the problem has been solved. Although this algorithm is guaranteed to converge, decision making heuristics are used to choose the pair of multipliers so as to accelerate the rate of convergence.

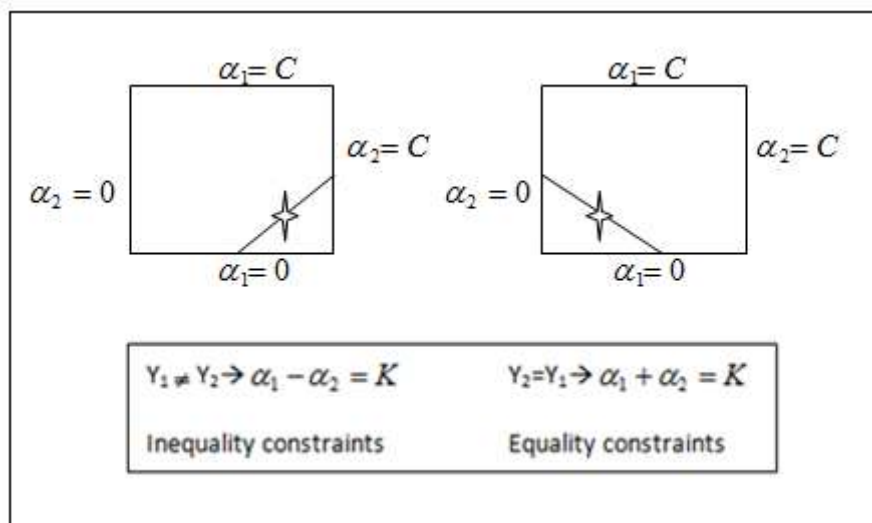
### 3.3 Optimization of SVM using SMO

The fault data sets are the combination of five types of diagnostic gases content obtained by overheating of the transformer oil used for fault analysis. Initially these gases are preprocessed by data processing methods and the features are extracted to identify the fault types through classifiers. This paper describes the effective fault diagnosis performance of SVM, by employing optimization on SVM is known as SVM-SMO classifier. The SVM-SMO technique increases classification accuracy and reduces computational time by using DGA data set. Optimization problem has two Lagrange multipliers which must obey linear equality constraint according to Eq (24) to establish an optimized SVM,

$$\sum Y_i \alpha_i = 0 \tag{24}$$

At every step, SMO chooses two Lagrange multipliers jointly to optimize and find the optimal values of the multipliers to update the input data by satisfying KKT conditions [22]. SMO does not require additional space, i.e., for storing the previous  $\alpha_1 \alpha_2$  and current  $\alpha_1 \alpha_2$  needs only 2\*2 matrices so very large SVM training problems can fit inside the memory. The kernel parameter of SVM can be optimized by SMO by updating the two Lagrange multipliers at every step. The

kernel function (k) measures the similarity or the distance between the input vectors  $\bar{x}$  and stored training vector  $\bar{x}_j$ .



**Figure 3 Constraints for Lagrange multipliers**

The two Lagrange multipliers of SMO must satisfy the linear equality constraints on updating the kernel parameters. Figure 3 presents the linear equality constraints which cause the two Lagrange multipliers to lie on a diagonal line. SMO finds an optimum of the objective function on a diagonal line segment [22].

## 4. EXPERIMENTAL RESULTS

In this section, transformer fault analysis are carried out using DGA dataset performed with Intel core processor of speed 3GHZ and 2GB RAM PC. The features are extracted from the dataset; these features are trained, validated and tested using MATLAB for the classification of types of fault states in the transformers. The performances of the classifier are experimented and presented in terms of relative speed and the computational time for the 75 sets of fault data of power transformer.

### 4.1 Dissolved Gas Analysis



The four types of fault states observed from the power transformer are as follows: normal state, thermal heating, low energy discharge and high energy discharge. Due to overheating the insulating transformer oil will degrade and emit gases, the ratios of this emission gases are closely related to the fault types that contain H<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub> and C<sub>2</sub>H<sub>6</sub>[17, 21]. These combustible gases occur due to fault in the transformer and are monitored regularly to determine the degree, pattern and the abnormality conditions using SVM and optimized SVM classification methods [25].

#### 4.1.1 Data set

The proposed method has 75 historical data sets of a 500kV main transformer, obtained from Pingguo Substation of South China Electric Power Company, and is utilized for training and testing the SVM and optimized SVM classifiers [24]. The data samples are then divided, where 50 data sets are used in training and validation and the remaining 25 data sets are considered for testing. The 50 data sets involved in training and validation are divided into 5 sets for normal state, 25 sets for thermal heating, 5 sets for low energy discharge and 15 sets for high energy discharge. The 25 data sets involved in testing are divided into 4 sets for normal state, 13 sets for thermal heating, 6 sets for low energy discharge and 2 sets for high energy discharge.

#### 4.1.2 Feature extraction

The dissolved diagnostic gases obtained from oil transformer are pre-processed by data processing and the features are extracted for fault analysis using SVM and SVM-SMO classifiers [24].

Feature extraction:

- (1) The data sets obtained by DGA methods having five features in each sample such as, y<sub>1</sub>, y<sub>2</sub>, y<sub>3</sub>, y<sub>4</sub> and y<sub>5</sub>, are extracted by the equation as follows:

$$y_k = \frac{c^k}{\max_{i=1}^5(c^i)} \tag{25}$$

where k=1, 2, . . . , 5,

c<sup>k</sup> – Absolute gas content of five diagnostic gases per one sample.

- (2) The absolute information is calculated for different samples given by,

$$Y_6 = \log_{10} \left( \max_{m=1}^5 c^m \right) \tag{26}$$

Thus, six features are extracted, gives a feature vector

Y = [y<sub>1</sub> y<sub>2</sub> y<sub>3</sub> y<sub>4</sub> y<sub>5</sub> y<sub>6</sub>]<sup>T</sup> for fault diagnosis by SVM-SMO.

Eq. (25) reflects the information about the classified four fault types of transformer.

### 4.2 Parameters

The parameters of SVM, and SMO used for classification of faults with DGA data set are discussed in this section. The performance of the classifiers mainly depends on the parameter values.

#### 4.2.1 SVM

The challenging nature of SVM is based on the best choice of kernel function, kernel parameter and margin parameter. The Gaussian RBF kernel function is the universally accepted kernel function. The kernel parameter varies as the classification accuracy increases and the values are varied between [0.3, 1.0]. Margin parameter has to set according to the kernel value increased from 75 to 100, order of the polynomial kernel is set as 3, parameter learning rate is chosen as 0.01 and scaling factor in the radial basis function kernel is taken as 1 which are provided in Table 1.

**Table 1 Parameters of SVM for transformer fault analysis**

Parameters	Notations used	Values
Kernel parameter	K	0.5
Margin parameter	C	100
Order of the polynomial	Order	3
Learning rate	λ	0.01
Scaling factor	σ	1

#### 4.2.2 SMO

The parameters of SMO are Karush-Kuhn-Tucker, Kernel cache limit, number of iterations, order of the polynomial and scaling factor are mentioned in Table 2. The Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient conditions for an optimal point of a positive definite QP problem. The selection of values between 0 and 1 specifies the

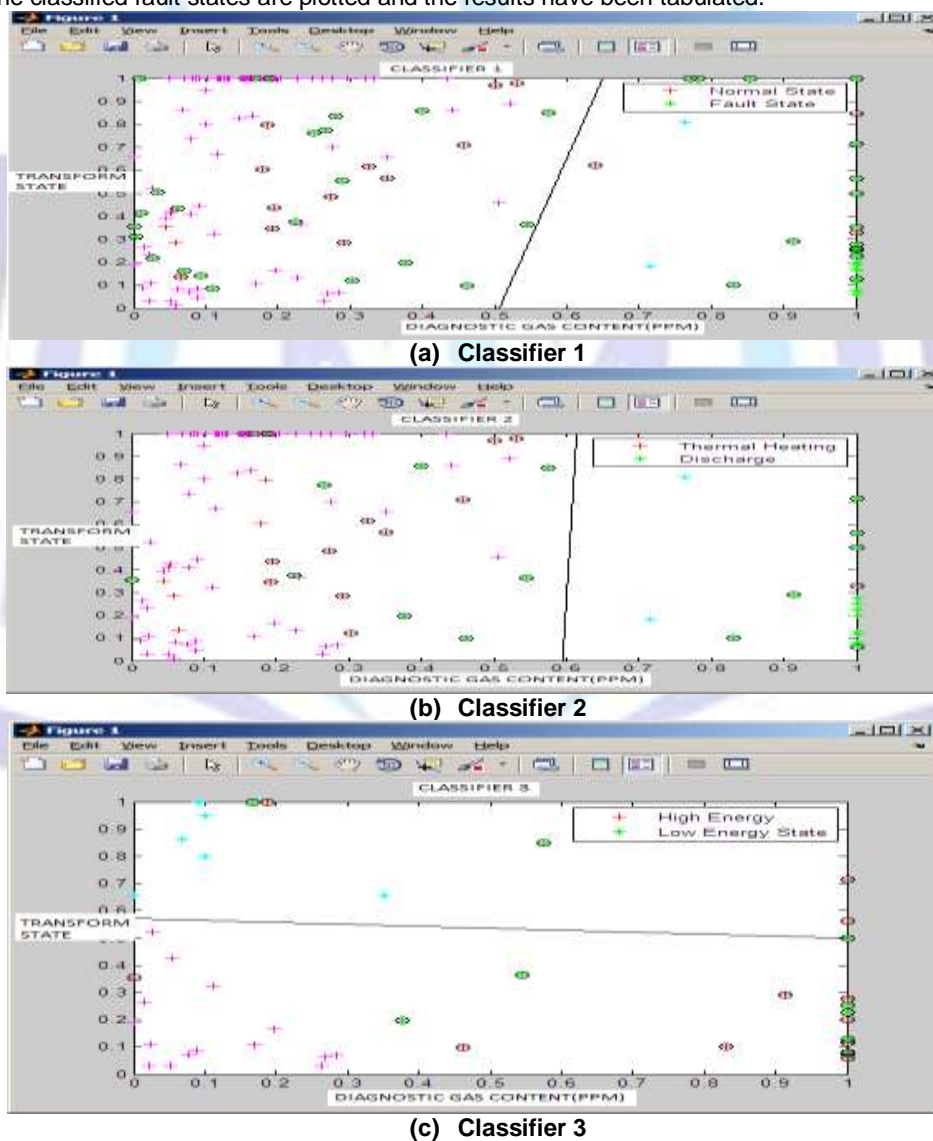
fraction of variables allowed to violate the KKT conditions for the SMO training method. For Gaussian RBF kernel function order of the polynomial is 3 and scaling factor is set as 1. Kernel cache limit specifies the size of the matrix has been fixed in advance as 5000 otherwise; the SMO will lead to overflow condition. Max iter is an integer set as 15000 that specify number of iterations of the main loop, if this limit is exceeded before the algorithm converges, then the algorithm stops and returns with an error.

**Table 2 Parameters of SMO for transformer fault analysis**

Parameters	Notations used	Values
Karush-Kuhn-Tucker	KKT	0.2
Kernel cache limit	Cache	5000
Number of iteration	Max Iter	15000
Order of the polynomial	Order	3
Scaling factor	$\sigma$	1

### 4.3 Result Analysis

The proposed optimization approach is made effective by optimizing the SVM parameter has been proven in the result analysis where the classified fault states are plotted and the results have been tabulated.

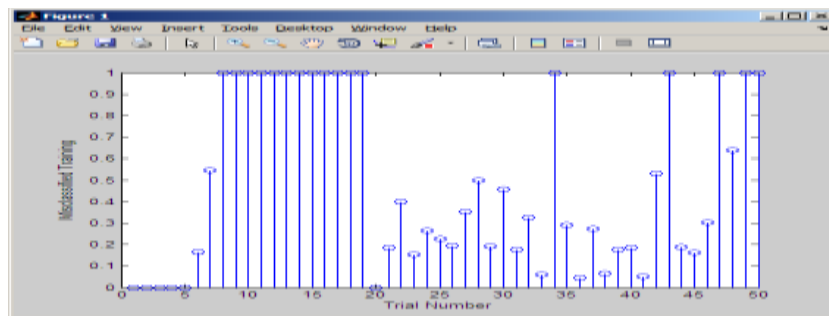


**Figure 4 Principal component projection maps for normal state and various fault state conditions using SVM SMO as a classifier**

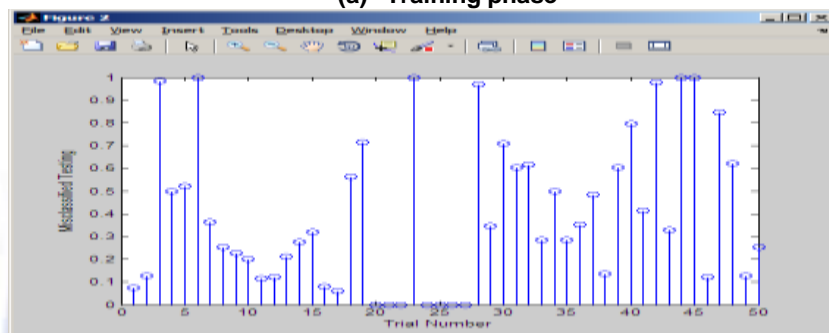
In Figure 4, original input data sets are projected over an n-dimensional feature space through objective function. The four types of fault have been differentiated through three classifiers: classifier 1, classifier 2 and classifier 3. First classifier classifies DGA data sets as Normal state and Fault state, the second classifier classifies fault states as Thermal heating



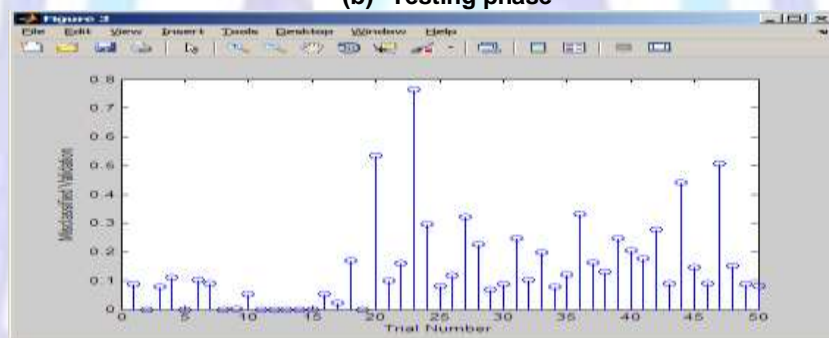
and Discharge, finally the third classifier classifies Discharge data as High energy discharge and Low energy discharge spread over given space. Thus, four types of faults are divided by three classifiers each one representing the plot drawn between various transform states and Diagnostic gas content.



(a) Training phase



(b) Testing phase



(c) Validation phase

Figure 5 Graph of power transformer fault data set for performance comparison using SVM SMO

Figure 5 shows the results of 50 trials observed, that represent the performance comparison of power transformer fault diagnosis, carried out using SVM-SMO techniques. Misclassification of input data during training, validation and testing are generated and predicted to show the performance evaluation of classifier in terms of classifier accuracy.

Table 3 Classification performance of power transformer fault data set using SVM

Training set(%)	Testing set (%)	Training+Validation			Testing		
		Accuracy (%)	Standard Deviation	Training time(sec)	Accuracy (%)	Standard Deviation	Testing time(sec)
30	70	87	0.079	3.78	90.13	0.037	4.00
40	60	89	0.09	3.50	93	0.045	4.02
70	30	92.05	0.081	3.78	94.58	0.039	4.00
80	20	93.12	0.085	3.71	95.13	0.042	3.96
90	10	94.12	0.09	3.71	96.15	0.045	3.10

Table 3 presents the results of proposed SVM methods tabulated in terms of training accuracy, standard deviation and training time, for various sets of training and testing percentages. Training accuracy and testing accuracy of standard SVM are 86% and 82% for 90% training and 10% testing of input datasets [25]. The increased training accuracy and testing accuracy of proposed SVM are 94.12%, 96.15% for same percentages of training and testing of input datasets. The tabulation shows improved performance with increased training accuracy and testing accuracy of 87% and 90.13%, even for 30% training of input datasets.



#### 4.4 Performance of SVM-SMO:

Table 4 depicts effective execution of SVM-SMO classifier, in detection and diagnosis stages, of the fault data set. The improved training accuracy and testing accuracy of SVM-SMO than SVM are 90.12% and 91.10% for 30% training and 70% testing of input datasets, whereas training accuracy and testing accuracy of proposed SVM are 87% and 90.13% for the same number of training and testing samples. Therefore for different range of sample selection tests, SVM-SMO provides high performance in terms of training accuracy, training time and standard deviation when compared to proposed SVM.

**Table 4 Classification performance of power transformer fault data set using SVM-SMO**

Training set (%)	Testing set (%)	Training+Validation			Testing		
		Accuracy (%)	Standard Deviation	Training time(sec)	Accuracy (%)	Standard Deviation	Testing time(sec)
30	70	90.12	0.019	4.01	91.10	0.011	4.25
40	60	94	0.00	3.86	93.98	0.013	4.50
70	30	94.98	0.029	4.01	96.48	0.03	4.25
80	20	95.87	0.031	4.10	96.99	0.039	4.10
90	10	96.13	0.041	4.15	97.59	0.041	3.98

**Table 5 Comparative results of power transformer fault analysis of the classifiers**

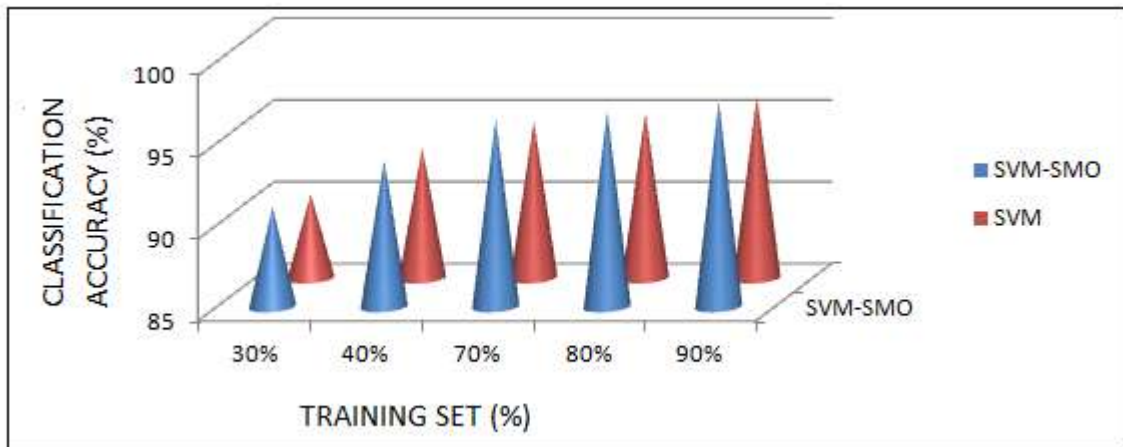
	SVM			SVM-SMO		
	Testing Accuracy (%)	Testing time (sec)	Error (%)	Testing Accuracy (%)	Testing time (sec)	Error (%)
CLASSIFIER 1	94	0.98	6	95.13	1.05	4.87
CLASSIFIER 2	93	0.90	7	96.12	1.15	3.88
CLASSIFIER 3	92.01	0.91	7.99	97.12	1.12	2.88
MEAN	93	0.93	7.0	96.12	1.12	3.88

Table 5 provides overall results of classification accuracy in terms of trial time and error rate that compares the performance of classifiers SVM and SVM-SMO. The mean computational time for SVM and SVM-SMO are calculated and presented. SVM-SMO takes larger testing time than SVM because for optimizing the parameter the code execution period is vast to get improved classification rate. The results indicate that the mean accuracy of SVM-SMO is high (96.12%) with reduced error rate than SVM (93%).

**Table 6 Results comparison of training and testing samples (90% and 10%)**

Method	Training + validation accuracy (%)	Testing accuracy (%)	Testing time (sec)	Classifiers mean accuracy (%)
Existing SVM[25]	86.00	82.00	0.66	90.10
Proposed SVM	94.12	96.15	3.10	93
SVM – SMO	96.13	97.59	3.98	96.12

Table 6 depicts the comparative results of existing SVM [25], proposed SVM and SVM-SMO of DGA dataset, for training (90%) and testing (10%) samples of the data set. SVM-SMO has highest accuracy of 3% which is higher than proposed SVM, also is 6% higher than existing SVM is achieved by optimizing the parameter of SVM.



**Figure 6 Comparison of classification accuracy of various classifiers for a power transformer fault data set with variable training sets.**

The classification accuracies of SVM and SVM-SMO against various percentages of training samples and testing samples of the DGA data are plotted in Figure 6. The improved performance of SVM-SMO than SVM classifier in terms of classification accuracy is shown in Figure 7.



**Figure 7 Comparative results showing the performance of classifiers during fault classification of power transformer.**

## 5. CONCLUSION

In this paper, the kernel function is optimized by continuously updating the two Lagrange multipliers to solve SVM's QP problem using a SVM-SMO based approach. The gas chromatograph method is utilized to choose the most appropriate gas signature from the faulty power transformer; these data sets are used to present the performance of optimized SVM. The features are applied as input data to the classifiers for faults classification in terms of accuracies and computational time. Test results of the proposed approach explain the effectiveness of the classifiers in terms of high efficiency than the standard approach. The classification accuracy of SVM-SMO (96.13) is better than that of SVM (94.12), achieves excellent performance in identifying the transformer fault type i.e. nature of the Fault. Experiments are carried out using



real world data derived from the power transformer DGA data, results of proposed optimization techniques are demonstrated with the effectiveness of classification rate. In future, the optimization steps to SVM are also applied to solve non-linear QP problem with reduced computational time of the classifiers, this problem can also be extended for advanced optimization methods like stochastic methods, heuristic methods and response surface methodology based approaches to get high accuracy and improvement in speed.

## REFERENCES

- [1] T.K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," IEEE Trans. Dielect. Electr. Insul., vol. 10, no. 5, pp. 903–917, Oct. 2003.
- [2] V.G. Arakelian, "The longway to the automatic chromatographic analysis of gases dissolved in insulating oil," IEEE Elect. Insul. Mag., vol. 20, no. 6, pp. 8–25, Nov./Dec. 2004.
- [3] G. Y. Yang, Z. Y. Dong, and K. P. Wong, "A modified differential evolution algorithm with fitness sharing for power system planning," IEEE Trans. Power Syst., vol. 23, no. 2, pp. 514–522, May 2008.
- [4] R.R. Roger, "IEEE and IEC Codes to Interpret Incipient Faults in Transformers Using Gas in Oil Analysis," IEEE Transactions on Electrical Insulation, vol. 13, no. 5, pp. 348-354, 1978.
- [5] M. Duval, "Dissolved gas analysis: It can save your transformer," IEEE Electrical Insulation Magazine, vol. 5, no. 6, pp. 22-27, 1989.
- [6] T. O. Rose, "Mineral Insulating Oil in Transformers," IEEE Electrical Insulation Magazine, vol. 14, No. 3, May/June, pp. 6-28, 1998.
- [7] X. Hao and S. Cai-xin, "Artificial Immune Network Classification Algorithm for Fault Diagnosis of Power Transformer," IEEE TRANS. ON POWER DELIVERY, vol. 22, no. 2, APRIL 2007, pp. 930-935.
- [8] Z. Wang, Y. Liu, and P. J. Griffin, "A combined ANN and expert system tool for transformer fault diagnosis," IEEE Trans. Power Del., vol. 13, no. 4, pp. 1224–1229, Oct. 1998.
- [9] A. Akbari, A. Setayeshmehr, H. Borsi and E. Gockenbach, "Intelligent Agent-Based System Using Dissolved Gas Analysis to Detect Incipient Faults in Power Transformers," IEEE Electrical Insulation Magazine, Vol. 26, No. 6, pp.27-40, November/December, 2010.
- [10] K.F. Thang, R.K. Aggarwal, A. J. McGrail, and D. G. Esp, "Analysis of power transformer dissolved gas data using the self-organizing map," IEEE Trans. Power Del., vol. 18, no. 4, pp. 1241–1248, Oct. 2003.
- [11] N.K. Patel and R.K. Khubchandani, "ANN Based Power Transformer Fault Diagnosis," IE (I) Journal. EL, vol 85, pp. 60-63, June 2004.
- [12] D.V.S.S. Siva Sarma and G.N.S. Kalyani, "Application of AI Techniques for Nondestructive Evaluation of Power Transformers Using DGA," IJIESP, Vol. 2, no. 1 pp.1-7, 2007.
- [13] M.H. Wang, "Extension Neural Network for power transformer incipient fault diagnosis," IEE Proceedings, Generation, Transmission, Distribution, vol.150, issue 6, pp. 679-685, Nov 2003.
- [14] 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. On System, Man, and Cybernetics-Part C: APPLICATIONS AND REVIEWS, vol. 39, no. 1, pp. 69-79, Jan 2009.
- [15] W. Chen, C. Pan, Y. Yun, and Y. Liu, "Wavelet Networks in Power Transformers Diagnosis Using Dissolved Gas Analysis," IEEE Transactions on Power Delivery, Vol. 24, no, pp. 187-194, Jan. 2009.
- [16] H.T. Yang, C.C. Liao, and J.H. Chou, "Fuzzy learning vector quantization networks for power transformer condition assessment," IEEE Trans. Dielect. Electr. Insul., vol. 8, no. 1, pp. 143–149, Feb. 2001.
- [17] M. Dong, D.K. Xu, M.H. Li, et al. "Fault Diagnosis Model for Power Transformer Based on Statistical Learning Theory and Dissolved Gas Analysis," in proc. IEEE International Symposium on Electrical Insulation, USA, 2004, pp.85-88.
- [18] C.W. Hsu, C.J. Lin, "A Comparison of Methods for Multiclass Support Vector Machines," IEEE Transactions on Neural Networks, vol 13, no 2, pp. 415-425, March 2002 .
- [19] S.F. Yuan and F.L. Chu, "Support vector machines based fault diagnosis for turbo-pump rotor," Mechanical Systems and Signal Processing, vol 20, no 4, pp. 939–952, 2006.





- [20] R. Wehrens, H. Putter, and L.M.C. Buydens, "The bootstrap: a tutorial," Chemometrics and Intelligent Laboratory Systems, vol 54, pp. 35–52, 2000.
- [21] S. Qaedi, and S. Seeyedtabaii, "Improvement in Power Transformer Intelligent Dissolved Gas Analysis Method", World Academy of Science, Engineering and Technology vol. 61, pp. 1144-1147, 2012.
- [22] John C.Platt, "A Fast Algorithm for Training Support Vector Machines," Microsoft Research, Technical Report MSR-TR-98-14, pp. 1-21, April 21, 1998.
- [23] D.Chandrakala, S.Sumathi, and S.Karthi, "Optimized code matrix generation for classification of multi-class pattern recognition problems using machine learning techniques" International Journal of Knowledge-based and Intelligent Engineering Systems 15 (2011) 227–245.
- [24] G. Y. Lv, H.Z. Cheng, H. B. Zhai, and L. X. Dong, "Fault diagnosis of power transformer based on multi-layer SVM classifier," Electric Power Systems Research, vol. 75, no. 1, pp. 9-15, 2005.
- [25] Ke Meng, Zhao Yang Dong, Dian Hui Wang, and Kit Po Wong, "A Self-Adaptive RBF Neural Network Classifier for Transformer Fault Analysis," IEEE TRANSACTION ON POWER SYSTEMS, VOL. 25, NO.3, pp. 1350-1360, AUGUST 2010.
- [26] Seifedeline souahlia, khmais Bacha, Abdelkader chaari, "Artificial Intelligence Tools Aided-decision for power transformer Fault Diagnosis", International Journal of Computer Applications (0975-8887). Volume 38- No.3, January 2012.
- [27] Surekha P, Dr.S.Sumathi, "A Self Adaptive Fuzzy C-means based Radial Basis function to solve Economic Load Dispatch problems", International Journal of Computer Applications (0975-8887), Volume-25-No.4, pp. 50-59, July 2011.

### Author' biography



**V.Gomathy** is currently pursuing her Ph.D in Electrical Engineering from PSG College of Technology, Coimbatore. Her research areas include Power Electronics, Fault Diagnosis, Soft Computing and Intelligent Controllers



**Dr.S.Sumathi** is currently working as Associate Professor in department of Electrical and Electronics Engineering, PSG College of Technology, Coimbatore. Her research areas include Applied Electronics, Neural Network and Data Mining.