Artificial Intelligence Tools Aided-decision for Power Transformer Fault Diagnosis

Seifeddine Souahlia C3S, ESSTT 5, Taha Hussein Street – Tunis Bab Menara 1008, Tunisia Khmais Bacha C3S, ESSTT 5, Taha Hussein Street – Tunis Bab Menara 1008, Tunisia Abdelkader Chaari C3S, ESSTT 5, Taha Hussein Street – Tunis Bab Menara 1008. Tunisia

ABSTRACT

This paper presents an intelligent fault classification approach for power transformer dissolved gas analysis (DGA). Fault diagnosis methods by the DGA and artificial intelligence (AI) techniques are implemented to improve the interpretation accuracy for DGA of power transformers. The DGA traditional methods are utilized to choose the most appropriate gas signature. AI techniques are applied to establish classification features for faults in the transformers based on the collected gas data. The features are applied as input data to fuzzy logic, artificial neural network (ANN) and support vector machine (SVM) classifiers for faults classification. The experimental data from Tunisian Company of Electricity and Gas (STEG) is used to evaluate the performance of proposed method. The results of the various DGA methods are classified using AI techniques and the results are compared with the empirical test. In comparison to the results obtained from the AI techniques, the ratios DGA method has been shown to possess the most excellent performance in identifying the transformer fault type. The test results indicate that the SVM approach can significantly improve the diagnosis accuracies for power transformer fault classification ..

General Terms

Key gas, graphical representation, ratios, multi-layer perceptron, radial basis function.

Keywords

Dissolved gas analysis, support vector machine, artificial neural network, fuzzy logic, transformer fault diagnosis.

1. INTRODUCTION

Essential devices as power transformers are in a transmission and distribution system. The wide varieties of electrical and thermal stresses often age the transformers and subject them to incipient faults. Being one of the most expensive and important elements, a power transformer is a highly essential element, whose failures and damage may cause the outage of a power system. If an incipient failure of a transformer is detected before it leads to a catastrophic failure, predictive maintenance can be deployed to minimize the risk of failures and further prevent loss of services. To monitor the serviceability of power transformers, many devices have evolved, such as Buchholz relays or differential relays. But the main shortcoming of these devices is that they only respond to the severe power failures which require removal of equipment from the service. Thus, techniques for early detection of the faults would be very valuable to avoid outages.

In industrial practice, dissolved gas analysis (DGA) is a very efficient tool for such purposes since it can warn about an impendent problem, provide an early diagnosis, and ensure transformers' maximum uptime.

DGA is a non-invasive technique, which can give valuable information on the condition of a transformer. High stresses of transformer operations may result in chemical reactions of the oil or cellulose molecules constituting the dielectric insulation, which may be caused by the dielectric breakdown of the oil or hot spots [1].

To find out the incipient faults, Dissolved Gas Analysis (DGA) is a prevailing method with periodically samples which test the insulation oil of transformers to obtain the composition of the gases dissolved in the oil due to breakdown of the insulating materials inside. The DGA methods then analyse and interpret the attributes acquired: ratios of specific dissolved gas concentrations, their generation rates and total combustible gases are used to conclude the fault situations. Diverse diagnostic criteria were developed for identification of the possible fault types: e.g., the conventional key gas method used in [2], ratio method presented in [3] and [4], and graphical representation method introduced in [5].

Recently, artificial intelligence techniques have been extensively used with the purpose of developing more accurate diagnostic tools based on DGA data.

In Ref. [6] the authors present a new and efficient integrated neural fuzzy approach for transformer fault diagnosis using dissolved gas analysis. The proposed approach formulates the modelling problem of higher dimensions into lower dimensions by using the input feature selection based on competitive learning and neural fuzzy model. Then, the fuzzy rule base for the identification of fault is designed by applying the subtractive clustering method which is efficient at handling the noisy input data.

In Ref. [7] the authors describe how mapping a neural network into a rule-based fuzzy inference system leads to knowledge extraction. This mapping makes explicit the knowledge implicitly captured by the neural network during the learning stage, by transforming it into a set of rules. This method is applied to transformer fault diagnosis using dissolved gas-in-oil analysis.

The back propagation (BP)-based artificial neural nets (ANN) described in Ref. [2] can identify complicated relationships among dissolved gas contents in transformer oil and corresponding fault types. The BP determines the optimal connection weights and bias terms to achieve the most accurate diagnosis model for DGA.

In Ref. [1] the authors present an intelligent fault classification approach to power transformer dissolved gas analysis (DGA), dealing with highly versatile or noisecorrupted data. Bootstrap and genetic programming (GP) are implemented to improve the interpretation accuracy for DGA of power transformers. Bootstrap pre-processing is utilized to approximately equalize the sample numbers for different fault classes to improve subsequent fault classification with GP feature extraction. GP is applied to establish classification features for each class based on the collected gas data. The features extracted with GP are then used as the inputs to artificial neural network (ANN), support vector machine (SVM) and K-nearest neighbour (KNN) classifiers for fault classification.

The aim of this paper is to present a new method for detection and classification of power transformers faults by using a dissolved gas analysis and an artificial intelligence technique for decision with a maximal classification rate.

This paper is organized as follows: Section 2 introduces the faults types and DGA methods. Section 3 presents the regression arithmetic of artificial intelligence techniques. The proposed tools and the test of this performance are presented in Section 4. Finally, the conclusion is provided in Section 5.

2. DISSOLVED GAS IN THE TRANSFORMER OIL

2.1 Transformer Fault Types

IEC Publication 60599 [8] provides a coded list of faults detectable by dissolved gas analysis (DGA):

• Partial discharge (PD):

PD occurs in the gas phase of voids or gas bubbles. It is usually easily detectable by DGA, however, because it is produced over very long periods of time and within large volumes of paper insulation. It often generates large amounts of hydrogen [5].

• Low energy discharge (D1):

D1 such as tracking, small arcs, and uninterrupted sparking discharges are usually easily detectable by DGA, because gas formation is large enough [5].

• High energy discharge (D2):

D2 is evidenced by extensive carbonization, metal fusion and possible tripping of the equipment [9].

- Thermal faults T <300 $^{\circ}$ C (T1):
- T1 evidenced by paper turned brownish.
- Thermal faults 300 <T< 700 °C (T2) :
- T2 evidenced when paper carbonizes.
- Thermal faults T $>700\ ^{o}\!C\ (T3)$:
- T3 evidenced by oil carbonization, metal coloration or fusion.

2.2 Diagnosis and Interpretation Methods

The DGA methods have been widely used by the utilities to interpret the dissolved gas. According to the pattern of the gases composition, their types and quantities, the interpretation approaches below for dissolved gas are extensively followed [10]:

- Gas key method;
- Ratios method;
- The graphical representation method.

✤ key gas method

In this key gas method, we need five key gas concentrations H2, CH4, C2H2, C2H4 and C2H6 available for consistent interpretation of the fault. Table 1 shows the diagnostic interpretations applying various key gas concentrations. The results are mainly adjectives and provide a basis for further investigation [11].

Table 1.	Interpretation	gas dissolve	d in the	oil [8]
	1	0		

Gas Detected	Interpretation
Oxygen (O ₂)	Transformer seal fault
Oxide and Dioxide Carbon (CO and CO ₂)	Cellulose decomposition
Hydrogen (H ₂)	Electric discharge (corona
	effect, low partial discharge)
Acetylene (C ₂ H 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 typical values range observed in power transformers according to IEC 60599 are given in Table 2.

 Table 2. Concentration typical values observed in transformers [8]

H ₂	CH ₄	C ₂ H ₆	C ₂ H ₄	C ₂ H ₂	СО	CO ₂
60- 150	40- 110	50-90	60- 280	3-50	540- 900	5100- 13000

✤ Ratios method

Rogers, Dornenberg and IEC are the most commonly used ratio methods. They employ the relationships between gas contents. The key gas ppm values are used in these methods to generate the ratios between them. The IEC method uses gas ratios that are combinations of key-gas ratios C2H2/C2H4, CH4/H2 and C2H4/C2H6.

Table 3 shows the IEC standard for interpreting fault types and gives the values for the three key-gas ratios corresponding to the suggested fault diagnosis. When key-gas ratios exceed specific limits, incipient faults can be expected in the transformer.

Table 3. Diagnosis using the ratio method (IEC 599) [8]

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 graphical representation

The graphical representation method is used to visualize the different cases and facilitate their comparison. The coordinates and limits of the discharge and thermal fault zones of the Triangle are indicated in Figure 1. Zone DT in Figure 1 corresponds to mixtures of thermal and electrical faults.



Figure 1: Coordinates and fault zones of the Triangle

The Triangle coordinates corresponding to DGA results in ppm can be calculated as follows:

% $C_2H_2 = 100x / (x + y + z)$, % $C_2H_4 = 100y / (x + y + z)$ and % $CH_4 = 100z / (x + y + z)$, where $x = (C_2H_2)$, $y = (C_2H_4)$ and $z = (CH_4)$.

You can translate the previous figure in a painting that gives the limits of each fault which are summarized in Table 4.

PD	98 % CH ₄	100 % CH ₄		
D1	23 % C ₂ H ₄	13 % C ₂ H ₂	100 % C ₂ H ₂	
D2	23 % C ₂ H ₄	13 % C ₂ H ₂	38 % C ₂ H ₄	29 % C ₂ H ₂
T1	4 % C ₂ H ₂	10 % C ₂ H ₄		
T2	4 % C ₂ H ₂	10 % C ₂ H ₄		
Т3	15 % C ₂ H ₂	50 % C ₂ H ₄	100 % C ₂ H ₄	

 Table 4. Graphical representation method zone limits [8]

3. BASIC CONCEPTS OF ARTIFICIAL INTELLIGENCE

3.1 Fuzzy Logic Approach

In the standard set, an element either belongs to or does not belong to a set; so the range of the standard set is , which can be used to solve a two-valued problem. In contrast to the standard set, the fuzzy set enables the description of concepts where the boundary is not explicit. It concerns not only whether an element belongs to the set but also to what degree it belongs to. The range of a fuzzy set is [0, 1]. The comparisons of the standard sets and fuzzy sets are shown in Table 5 [12].

Table 5. Two different sorts of mathematical sets

Mathematical sets	Standard set	Fuzzy set
Research objects	Data variables	Linguistic variables
Model	Mathematics model	Fuzzy mathematics model
Descriptive function	Transfer function	Membership function
Descriptive property	Precision	Ambiguity
Range of set	$C_A(x) \in (0,1)$	$\mu_{\rm A}(x) \in \left[0, 1\right]$

The fuzzy logic analysis consists of three parts: fuzzification, fuzzy inference and defuzzification.

Fuzzification is the process of transforming crisp input values into grades of membership for linguistic terms of fuzzy sets. The membership function is used to associate a grade to each linguistic term. A chosen fuzzy inference system (FIS) is responsible for drawing conclusions from the knowledgebased fuzzy rule set of If (X is A) then (Y is B) linguistic statements.

Defuzzification then converts the fuzzy output values back into crisp output actions [13].

There are several methods for calculating the output set representative value. The main ones are: defuzzification based on the sets gravity center and the maximum average methods.

3.2 Neural Network Approach

The neural network technique is used to recognize and classify complex fault patterns without much knowledge about the process, the used trials or the fault patterns themselves. A neural network consists of many simple neurons which are connected with each other.

The principal neural networks that we will use for the classification are:

• Multi-Layer Perceptron (MLP): Is a network organized in layers. A layer is a uniform neurons group without connection with each other and makes a transformation vector.

The architecture of the MLP is composed of an input layer, a variable number of hidden layers and by an output layer which is fully connected with them. In particular, as outlined in Figure 2, a three-level fully connected network, using a sigmoid output function, has been considered because it is known that this number of levels allows building decision regions of any shape.



Figure 2: The multi-layer perceptron architecture

• Radial Basis Function (RBF): RBF nets belong to the group of kernel function nets that utilize simple kernel functions and that are distributed in different neighborhoods of the input space whose responses are essentially local in nature. The architecture consists of one hidden and one output layer. This shallow architecture has great advantage in terms of computing speed compared to multiple hidden layer nets.

Each hidden node in an RBF net represents one of the kernel functions. An output node simply computes the weighted sum of the hidden node outputs. A kernel function is a local function and the range of its effect is determined by its center and width. Its output is high when the input is close to the center and it decreases rapidly to zero as the input's distance from the center increases. The Gaussian function is a popular kernel function and will be used in this algorithm [14].

3.3 Support Vector Machine (SVM) Approach

The purpose of SVM is to find an optimal separating hyperplane by maximizing the margin between the separating hyper-plane and the data (Vapnik, 1995) [15].

Applying the Lagrange principle, we obtain the quadratic programming problem of dimension m (examples number) as follows in Eq. 1:

$$\begin{cases} \max \sum_{i=1}^{m} \alpha_{i} - \frac{1}{2} \sum_{i,j} \alpha_{i} \alpha_{j} y_{i} y_{j} (x_{i} x_{j}) \\ \forall i, \quad \alpha_{i} \ge 0 \\ \sum_{i=1}^{m} \alpha_{i} y_{i} = 0 \end{cases}$$
(1)

Where x is the input vector, $y \in \{-1, 1\}$ is the output class and α_i are the Lagrange coefficients.

We define the supports vectors VS any vector xi as $yi.[(w_0, x_i) + b_o] = 1$. This is equivalent to Eq. 2:

$$VS = \{x_i | \alpha_i > 0\} \text{ for } i = 1, ..., m$$
 (2)

The ranking function class(x) is defined by Eq. 3:

$$class(x) = sign\left[(w_0.x) + b_0\right] = sign\left[\sum_{x_i \in VS} \alpha_i y_i(x_i.x) + b_0\right]$$
(3)

If class (x) is less than 0, x is the class -1 else it is a Class 1.

However for nonlinear cases, there is insufficient space for classifying the inputs. So, we need a larger space. We must therefore resolve Eq. 4:

$$\begin{cases} \max \sum_{i=1}^{m} \alpha_{i} - \frac{1}{2} \sum_{i,j} \alpha_{i} \alpha_{j} y_{i} y_{j} \phi(x_{i}) \phi(x_{j}) & (4) \\ \forall i, \quad 0 \le \alpha_{i} \le C \\ \sum_{i=1}^{m} \alpha_{i} y_{i} = 0 \end{cases}$$

With C is a margin setting.

 $K(x_i, x_j) = \phi(x_i).\phi(x_j)$ is a positive kernel function definite on R. The main kernel functions used for classification are given in Eq. 5, 6, 7 and 8:

W The linear Kernel function :

$$K(x, x') = x \cdot x' \qquad (5)$$

4 The polynomial Kernel function:

$$K(x, x') = (x.x')^{d}$$
 or $(c + x.x')^{d}$ (6)

Gaussian radial basis function :

$$K(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2})$$
(7)

Sigmoïd Kernel Fonction :

 $K(x, x') = \tanh(\alpha_0(x.x') + \beta_0)$ (8)

4. TRANSFORMERS FAULTS CLASSIFICATION BY ARTIFICIAL INTELLIGENCE TECHNIQUES

4.1 Training and Testing Data

This study employs dissolved gas content data in power transformer oil from chemistry laboratory of the Goulette central of Tunisian Company of Electricity and Gas (STEG). The data is divided into two data sets: the training data sets (94 samples) and the testing data sets (30 samples). The extracted DGA data contain not only the five concentrations of key gas, three relatives percentages and three ratios but also the diagnosis results from on-site inspections. The training data sets have been evaluated using various methods DGA and the corresponding judgments related to seven classes have been provided: normal unit (46 samples), Partial Discharge (2 samples), low energy discharge (3 samples), high energy discharge (17 samples), low temperature overheating (4 samples), middle temperature overheating (7 samples) and high temperature overheating (15 samples).

4.2 Classification by Fuzzy Logic

For The fuzzy logic faults classification is performed using several DGA methods as gas signature.

4.2.1 Fuzzy key gas

Firstly, we will classify the faults using key gas as input data with:

•5 linguistic variables are the 5 gas: H_2 , CH_4 , C_2H_2 , C_2H_4 and C_2H_6 ;

•3 linguistic values: small, medium and high;

•5 sets of reference: U = [0, 600] for H_2 , U = [0, 500] for CH_4 , U = [0, 400] for C_2H_2 , U = [0, 700] for C_2H_4 and U = [0, 310] for C_2H_6 ;

•7 outputs, the reference sets are : U = [0, 1] for the non fault, U = [0, 2] for the PD, U = [1, 3] for the D1, U = [2, 4] for the D2, U = [3, 5] pour for the T1, U = [4, 6] for the T2 and U = [5, 7] for the T3 ;

•3 membership functions: triangular, trapezoidal and Gaussian;

 $\cdot 3^5 = 243$ fuzzy rules;

•Defuzzification by the centroid method.

The performance of key gas method is analyzed in terms false alarm rate and non-detection rate for triangular, trapezoidal and Gaussian membership functions as shown in Table 6.

According to the results, we find that the triangular membership function is more efficient for system fault diagnosis, but this method does not give excellent results. So, we must propose an alternative method.

Membership	False alarm rate	Non-detection rate
function	(%)	(%)
Gaussian	16.7 (5 / 30)	53.3 (16 / 30)
Trapezoid	6.7 (2 / 30)	26.7 (8 / 30)
Triangular	3.3 (1 / 30)	26.7 (8 / 30)

Table 6. The key gas method classification performance

4.2.2 Fuzzy graphical representation

Secondly, we will classify the faults using graphical representation as input data with:

•3 linguistic variables are the 3 relatives percentages: % CH₄, % C_2H_2 and % C_2H_4 ;

•4 linguistic values: very low, low, medium and high;

•1 set of reference: U = [0, 100] for every linguistic variable;

•The same 7 outputs;

•The same 3 membership functions;

•33 fuzzy rules;

The test set is again fed to the fuzzy logic in order to verify if it classifies and estimates correctly. Table 7 tabulates the results by representing the false alarm rate and the nondetection rate for three membership functions.

Table 7. The key gas method classification performance

Membership	False alarm rate	Non-detection rate	
function	(%)	(%)	
Gaussian	20 (6 / 30)	3.3 (1 / 30)	
Trapezoid	16.7 (5 / 30)	6.7 (2 / 30)	
Triangular	16.7 (5 / 30)	3.3 (1 / 30)	

From Table 7, we note that the triangular membership function is more efficient system fault diagnosis.

4.2.3 Fuzzy ratios

Now, we will classify the faults using ratios as input data with:

•3 linguistic variables are the 3 ratios: $C_2H_2/C_2H_4,\,CH_4/H_2$ and $C_2H_4/C_2H_6\,;$

•4 linguistic values: very low, low, medium and high ;

•3 sets of reference: U = [0, 10] for C_2H_2/C_2H_4 , U = [0, 25] for CH_4/H_2 and U = [0, 12] for C_2H_4/C_2H_6 ;

•The same 7 outputs;

•The same 3 membership functions;

 $\bullet 4^3 = 64$ fuzzy rules;

The diagnosis results for test set with the triangular, trapezoidal and Gaussian membership functions are listed in Table 8.

According to test results, we find that the triangular membership function gives a better diagnosis compared to the Gaussian and trapezoidal membership functions.

Table 8. Th	e ratios me	thod classifica	tion performance
-------------	-------------	-----------------	------------------

Membership function	False alarm rate (%)	Non-detection rate (%)
Gaussian	16.7 (5 / 30)	53.3 (16 / 30)
Trapezoid	6.7 (2 / 30)	10 (3 / 30)
Triangular	6.7 (2 / 30)	10 (3 / 30)

4.2.4 Comparative investigation of fuzzy logic classification

According to test results of the four inputs data classified by the fuzzy logic, we conclude that the triangular membership function gives the best results.

To select the most significant gas analysis method, we compare the false alarm rate and non-detection rate of four inputs data types which are given in Table 9.

 Table 9. False alarm rate and non-detection rate of the three DGA methods

Membership function	False alarm rate (%)	Non-detection rate (%)
Key gas	3.3 (1 / 30)	26.7 (8 / 30)
Graphical representation	16.7 (5 / 30)	3.3 (1 / 30)
Ratios	6.7 (2 / 30)	10 (3 / 30)

The actual result indicates the classification accuracies obtained by using the ratios method is higher than those of gas signature for the classification by fuzzy logic.

To conclude, faults classification by the fuzzy logic technique can be achieved by the ratios method using the triangular membership function.

4.3 Classification by Neural Networks

An ANN-based power transformer fault diagnostic system includes input features, network topology, fault outputs as well as training patterns. In the current study, we used:

> Two types of neural networks: the MLP and RBF are used for transformers faults classification.

➤ Three types of input data: five key gases (key gas method), three ratios (ratios method) and three relatives percentages(graphical representation method) are chosen as input features;

➤ 3 binary outputs in order to minimizing the neurons number in output layer. The outputs codification are presented in the following:

- \checkmark 0 0 1 : no fault (normal working);
- \checkmark 0 1 0 : partial discharge fault (DP) ;
- \checkmark 011: low energy discharge fault (D1);
- ✓ 100: high energy discharge fault (D2);
- \checkmark 101: low temperature overheating (T1);
- \checkmark 1 1 0 : middle temperature overheating (T2) ;
- \checkmark 1 1 1: high temperature overheating (T3).

4.3.1 MLP classification

Firstly, we will classify the faults by MLP network. After the test, with several parameters, we obtained the appropriate MLP architecture with the minimal error rate. Then, the optimal parameters are utilized to train the MLP model. So, we used:

• Four layers: an input layer, two hidden layers, and an output layer;

• Transfer function tangent sigmoid for 4 layers which is given in Figure 3.



Figure 3: Tangent sigmoid transfer function

 $\hfill \black$ The neurons and iterations numbers are tabulated in Table 10.

Gas signature	Ne	eurons num	Epoch number	
Gussignuture	Input layer	Hidden layers	Output layer	number
Key gas	5	4	3	400
Ratios	3	3	3	400
Graphical representation	3	3	3	250

Table 10. Neurons and epoch numbers for MLP

The performance of MLP network is analyzed in terms false alarm rate and non-detection rate for three gas signature methods as can be seen in Table 11.

Table 11. The MLP classification performance

Gas signature	False alarm rate (%)	Non-detection rate (%)
Key gas	0	26.7 (8 / 30)
Graphical representation	0	26.7 (8/30)
Ratios	0	23.3 (7 / 30)

The actual result indicates the superiority of the ratios method for the MLP network classification.

4.3.2 RBF classification

Now, we will classify the faults by RBF network. The parameters of RBF model are optimized by many tests. The adjusted parameters with maximal classification accuracy are selected as the most appropriate parameters using:

- \Im Mean squared error goal = 0.01
- Spread of radial basis functions = 1
- \Im Maximum number of neurons = 40
- \Im Number of neurons to add between displays = 1

The performance of RBF network is analyzed in terms false alarm rate and non-detection rate for four gas signature methods as shown in Table 12.

Gas signature	False alarm rate	Non-detection rate
	(%)	(%)
Key gas	0	23.3 (7 / 30)
Graphical	0	20 (6 / 30)
representation		
Ratios	0	16.7 (5 /30)

From Table 12, the actual result indicates that the ratios as a gas signature has higher diagnostic accuracy than two other methods for the RBF network classification.

4.3.3 Comparative investigation of MLP and RBF classification

According to test results of the MLP and RBF networks, we conclude that the ratios method present the best performance; it can be used as a neural network input vector.

To select the appropriate neural network between the MLP and RBF networks, we compare the false alarm rate and nondetection rate given in Table 13.

Fable	13.	The	RBF	classification	performance
--------------	-----	-----	-----	----------------	-------------

Neural network	False alarm rate (%)	Non-detection rate (%)
MLP	0	23.3 (7 / 30)
RBF	0	16.7 (5 / 30)

The test samples examined by two neural networks shows that the types of faults detected by the RBF network are almost identical to the real fault types. The other network MLP doesn't have to correct the faults. The table above shows that RBF network has more excellent performance than MLP network. In this case, we adopt the RBF network.

4.4 Classification by SVM

As shown in Figure 4, the diagnostic model includes six SVM classifiers which are used to identify the seven states: normal state and the six faults (PD, D1, D2, T1, T2 and T3). With all the training samples of the states, SVM1 is trained to separate the normal state from the fault state. When input of SVM1 is a sample representing the normal state, output of SVM1 is set to +1; otherwise -1. With the samples of single fault, SVM2 is trained to separate the discharge fault from the overheating fault. When the input of SVM2 is a sample representing discharge fault, the output of SVM2 is set to +1; otherwise-1. With the samples of discharge fault, SVM3 is trained to separate the high-energy discharge (D2) fault from the partial discharge (PD) and low energy discharge (D1) fault. When the input of SVM3 is a sample representing the D2 fault, the output of SVM3 is set to +1; otherwise -1. With the samples of overheating fault, SVM4 is trained to separate the high temperature overheating (T3) fault from the low and middle temperature overheating (T1 and T2) fault. When the input of SVM4 is a sample representing the T3 fault, the output of SVM5 is set to +1; otherwise -1. SVM5 is trained to separate the middle temperature overheating (T2) fault from the low temperature overheating (T1) fault. When the input of SVM5 is a sample representing the T2 fault, the output of SVM5 is set to +1; otherwise -1. SVM6 is trained to separate the partial discharge (PD) fault from the low energy discharge (D1) fault. When the input of SVM6 is a sample representing the D1 fault, the output of SVM6 is set to +1; otherwise -1.



Figure 4: Diagnostic model of power transformer based on SVM classifier

All the six SVMs adopt polynomial and Gaussian as their kernel function. In SVM, the parameters σ and C of SVM model are optimized by the cross validation method. The adjusted parameters with maximal classification accuracy are selected as the most appropriate parameters. Then, the optimal parameters are utilized to train the SVM model.

So the output codification is presented in Table 14.

 Table 14. Codification output of SVM

	svm1	svm2	svm3	svm4	svm5	svm6
No	+1					
fault						
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		

Firstly, we will classify the faults by SVM with the polynomial kernel. The diagnosis results for test set with four gas signature methods are listed in Table 15.

Table 15. The polynomial kernel classification
performance

Gas signature	False alarm rate (%)	Non-detection rate (%)
Key gas	0	40 (12 / 30)
Graphical	0	23.3 (7 / 30)
representation		
Ratios	0	20 (6 / 30)

The actual result, which indicates the ratios gas signature, has a significantly high degree of diagnosis accuracy for the classification by SVM with polynomial kernel.

Secondly, we will classify the faults by SVM with the Gaussian kernel. Table 16 lists the diagnosis results for test set with four gas signature methods.

Table 16. The Gaussian kernel classification performance

Gas signature	False alarm rate (%)	Non-detection rate (%)
Key gas	0	26.7 (8 / 30)
Graphical	0	20 (6 / 30)
representation		
Ratios	0	13.3 (4 / 30)

As shown in Table 16, we note that the ratios gas signature with Gaussian kernel presents the best result in comparison with the other three gas signature.

According to faults classification test results by the SVM with polynomial and Gaussian kernel functions, we find that the ratios method is the most appropriate gas signature.

To select more efficient kernel between the two cores used (polynomial and Gaussian), we compare the false alarm rate and non-detection rate given in Table 17.

Table 17. False alarm and non-detection rates of SVM for different kernels

Kernel type	False alarm rate (%)	Non-detection rate (%)	
Polynomial	0	20 (6/30)	
Gaussian	0	13.3 (4/30)	

The results in Table 17 show that the Gaussian kernel gives the best performance for the test. This is aided by a proper choice of the kernel parameter σ by the cross validation method, because this parameter determines the hypersphere radius which encloses the data in multidimensional space.

So, for comparison with other classification techniques, we adopt the SVM with Gaussian kernel SVM as the most efficient.

4.5 Comparative investigation of diagnosis tools

After determining the classification tools of transformers faults, most processors meet the diagnosis system. In this paragraph, we will determine the most appropriate gas signature and the most efficient classification technique.

Regarding the most significant gas analysis method, all the three techniques (fuzzy logic with triangular membership function, RBF and SVM with Gaussian kernel function) indicate that the ratios method is a proper alternative. To choose the best tool possible, we will compare their performance using the test samples. Table 18 compares the proposed applications performance for identifying the transformers faults.

Gas signature	False alarm rate (%)	Non- detection rate (%)	Test performance (%)
Fuzzy logic (triangular membership function)	6.7 (2 / 30)	10 (3 / 30)	83.3
Neural networks RBF	0	16.7 (5 / 30)	83.3
SVM (gaussian kernel function)	0	13.3 (4/30)	86.7

Table 18. Performance of several classifications tools

The results presented in Table 18 show that the highest percentage of performance test is obtained with the SVM technique with a Gaussian kernel function (86.7 %).

The SVM classifiers offer very considerable results and has demonstrated its efficiency for power transformers fault diagnosis.

5. CONCLUSION

In this paper, the artificial intelligence techniques are implemented for the faults classification using the dissolved gas analysis for power transformers. The DGA methods studied are key gas, graphical representation and ratios method. The fault diagnosis models performance was analyzed with fuzzy logic (using Gaussian, trapezoidal and triangular membership functions), neural networks (MLP and RBF) and Support Vector Machine (with polynomial and Gaussian kernel functions). The real data sets are used to investigate the performance of the DGA methods in power transformer oil.

The experimental results show that the fuzzy logic classifier with triangular membership presents the best result in comparison with the other two membership functions. The classification accuracies of RBF are superior to MLP NN and the SVM with Gaussian kernel function has more excellent diagnostic performance than the SVM with polynomial kernel function.

According to test results, it is found that the ratios method is more suitable as a gas signature. The SVM with the Gaussian kernel function has a better performance than the other AI methods in diagnosis accuracy. The accuracy of the SVM for fault detection is comparable to conventional methods due to their great facilities for study. The proposed method can be applied to online diagnosis of incipient faults in transformers.

6. REFERENCES

[1] A. Shintemirov, W. Tang and Q. H. Wu, "Power Transformer Fault Classification Based on Dissolved Gas Analysis by Implementing Bootstrap and Genetic Programming", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, Vol. 39, NO. 1, January 2009, 69-79.

- [2] SUN Yan-jing, ZHANG Shen, MIAO Chang-xin, LI Jingmeng, "Improved BP Neural Network for Transformer Fault Diagnosis", Journal of China University of Mining & Technology Vol.17 No.1, Mars 2007, 138-142.
- [3] W. H. Tang, J. Y. Goulermas, Q. H.Wu, Z. J. Richardson, and J. Fitch, "A Probabilistic Classifier for Transformer Dissolved Gas Analysis With a Particle Swarm Optimizer", IEEE transactions on power delivery, Vol. 23, No. 2, April 2008, 751-759.
- [4] Chin-Pao Hung, Mang-Hui Wang, "Diagnosis of incipient faults in power transformers using CMAC neural network approach", Electric Power Systems Research 71, 2004, 235–244.
- [5] Michel Duval, "A Review of Faults Detectable by Gas-in-Oil Analysis in Transformers". IEEE Electrical Insulation Magazine, Vol.18, No. 3, June 2002, 8-17.
- [6] R. Naresh, Veena Sharma and Manisha Vashisth, "An Integrated Neural Fuzzy Approach for Fault Diagnosis of Transformers", IEEE transactions on power delivery, Vol. 23, No. 4, October 2008, 2017-2024.
- [7] V. Miranda, A. Rosa Garcez Castro, "Improving the IEC Table for Transformer Failure Diagnosis With Knowledge Extraction From Neural Networks", IEEE transactions on power delivery, Vol. 20, No. 4, October 2005, 2509-2516.
- [8] Standard IEC 60599, 2007, "Guide for the interpretation of dissolved gas analysis and gas-free".
- [9] Tapan K. Saha, "Review of Modern Diagnostic Techniques for Assessing Insulation Condition in Aged Transformers", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 10, No. 5, October 2003, 903-917.
- [10] Hong-Tzer Yang, Chiung-Chou Liao, "Adaptive Fuzzy Diagnosis System for Dissolved Gas Analysis of Power Transformers", IEEE Transactions on Power Delivery, VOL. 14, NO. 4, October 1999, 1342-1350.
- [11] IEEE Standards, "Terminology for Power and Distribution Transformers", ANSI/IEEE Standard, C57.12.80, 1992.
- [12] Mang-Hui Wang, "A Novel Extension Method for Transformer Fault Diagnosis", IEEE transactions on power delivery, Vol. 18, No. 1, January 2003, 164-169.
- [13] 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, 2009, 83-88.
- [14] Z. Moravej, D.N. Vishwakarma, S.P. Singh, "Application of radial basis function neural network for differential relaying of a power transformer", Computers and Electrical Engineering 29, 2003, 421–434.
- [15] V.N. Vapnik, 1995, The Nature of Statistical Learning Theory, Springer-Verlag, New York.