SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I2P107

Volume 11 Issue 2, 58-64, February 2024

Original Article

Power Transformer Classification through Dissolved Gas Analysis Utilizing Least-Square Support Vector Machine

Zuhaila Mat Yasin[1,2], Fathiah Zakaria[2], Nur Ashida Salim[2], Nur Fadilah Ab Aziz[3]

[1]*Solar Research Institute (SRI), Universiti Teknologi MARA, Shah Alam, Malaysia.*
[2]*School of Electrical Engineering, College of Engineering, University of Technology MARA, Shah Alam, Malaysia.*
[3]*Faculty of Engineering, University Tenaga Nasional, Kajang, Malaysia.*

[1]*Corresponding Author : zuhai730@uitm.edu.my*

Received: 17 November 2023 Revised: 07 December 2023 Accepted: 17 January 2024 Published: 16 February 2024

Abstract *- This article proposes the utilization of the Least-Square Support Vector Machine (LS-SVM) approach to ascertain the presence of a fault in power transformers. Power transformers are essential elements of electrical power systems. The failure of a power transformer can cause a disturbance in the functioning of power distribution and transmission systems. This situation will result in an increase in operating expenses due to the need for repairs and maintenance. The reliability of the electrical grid may be compromised. Therefore, it is crucial to identify any flaws in the power transformer at an early stage. In this paper, the LS-SVM utilizes Dissolved Gas Analysis (DGA) data as its input. The DGA methodology is widely accepted as the prevailing method for identifying the early stages of defects that arise in power transformers by analyzing the ratio of essential gases. The simulation data acquired from the industry comprises a standard state and six distinct fault types of transformers, which are utilized as input for the LS-SVM models. The suggested model underwent testing in multiple scenarios, yielding a maximum accuracy of 97.37%.*

Keywords *- Dissolved Gas Analysis, Least-Square Support Vector Machine, Incipient fault, Power transformer, ANN.*

1. Introduction

Power transformers are a necessary, costly, and crucial component of the electrical power system, and their failure would be catastrophic [1]. A transformer failure or damage can cause the suspension or interruption of electrical distribution and transmission activities, as well as significant repair costs. The primary role of transformer oil is to insulate, prevent arcing and corona discharge, and disperse heat from the transformer windings and core [2]. As a result, recognizing and fixing any faults in power transformers is crucial for improving system efficiency. Detecting incipient problems in power transformer oil via an intelligent system technology has become an intriguing research topic.

Several studies on monitoring and analyzing approaches to assess the state of the health of the power transformer for subsequent preventative and maintenance operations have been offered. Dissolved Gas Analysis (DGA) is a dependable and widely used technology for detecting early problems in power transformer-immersed oil [3]. The DGA requires regular oil testing as well as modern technology for online analysis and monitoring. Analyzing the vapors recovered from the transformer oil can uncover early-stage faults within the transformer [4]. The quantity and types of these gases are determined by the amount of energy exposed by the oil [5].

Several sources have offered fault detection in power transformers utilizing the Dissolved Gas Analysis (DGA) technique, taking into account both normal and faulty situations. The analysis of dissolved gas can be conducted using several ways for interpreting faults, such as the Roger ratio method, Doernenburg ratio method, key gas method, IEC ratio method, and Duval triangle method [6]. References [7-11] have described how to identify a power transformer fault by taking into account both the normal and the defective conditions. The Multilayer Perceptron Neural Network (MLPNN) identification of fault types in a power transformer is described in reference [7]. Combinations of Roger's ratio, Doernenburg's ratio, and Roger's and Doernenburg's ratios are used with the MLPN model. The suggested method can increase the accuracy of power transformer incipient defects by 85.31%.

Artificial Neural Network (ANN) technology was utilized in [8] to determine the type of power transformer fault. Levenberg Marquardt's backpropagation technique as multilayer reverse diffusion is the ANN training algorithm that is utilized. To identify different kinds of faults, the DGA approach incorporated the key gas method and Duval triangle method into consideration. The performance of the forecast was evaluated for accuracy using Mean Absolute Error.

RadBas, Tan-Sigmoid, Log-Sigmoid, and Purelin linear transfer functions are utilized. The results demonstrate that the mean absolute error produced using ANN is less than that obtained using the manual fault prediction technique.

Another study in [9] also proposed a multilayer perceptron type of ANN with DGA to determine the transformer conditions. Monitoring transformer conditions is critical for improving transformer dependability and efficiency. It can also extend the life of the power transformer. The results show that the ANN using DGA data provides an accurate solution. Multilayer Artificial Neural Network (MANN) and Support Vector Machine (SVM) models have been presented in [10] to classify the types of faults occurring in power transformers using DGA data. The accuracy achieved with the SVM classifier is 81.4%, which is higher than the performance with the multilayer ANN, which is only 76% accurate. In [11], SVM was used to determine the types of defects using DGA data. Power transformer faults were classified into four types: low-intensity discharge, high-intensity discharge, thermal fault, and no fault.

References [12-15] examined power transformer fault diagnosis based solely on fault condition. Reference [12] advocated using ANN to identify the types of issues that may occur in power transformers. DGA data was generated using the Duval triangle method, Roger ratio method, and Doernenburg method ratio. The Duval triangle method has the best accuracy (85%), followed by the Doernenburg ratio method (58.97%) and the Roger ratio method (52.27%). In [13], two SVM models are presented: Fine Gaussian for SVM1 and Kernel Linear for SVM2. Power transformer fault types were classified into four categories: partial discharge, low energy discharge, high energy discharge, and thermal fault. SVM1 has a better accuracy than SVM 2, which is 97.9% and 91%, respectively.

In their study, the authors of reference [14] have introduced a multistage Support Vector Machine (SVM) approach for accurately categorizing the different failure types in power transformers. The faults were categorized into four stages: SVM1, SVM2, SVM3, and SVM4. The SVM1 algorithm was employed to detect the presence of discharge fault and thermal fault. The SVM2 algorithm was employed to detect and classify partial discharge and discharge faults. The SVM3 algorithm was employed to detect thermal faults that are below 700°C and those that are above 700°C. The SVM4 algorithm was employed to detect the discharge of low-energy and high energy.

The SVM1, SVM2, SVM3, and SVM4 achieved accuracies of 91.45%, 98.80%, 91.18%, and 83.78% correspondingly. Different research conducted in [15] has introduced a Hybrid Genetic Algorithm and Artificial Neural Network (GA-ANN) to detect various forms of faults in power transformers. The performance of GA-ANN is being compared to a conventional approach. The results indicate that the proposed method achieved an accuracy of 95%, whereas the conventional method only achieved an accuracy of 72%.

Most previous researchers categorize various types of faults that occur in power transformers. Identifying power transformer failures in real-life scenarios is known for its inherent difficulty. Furthermore, determining the sort of fault occurring in a power transformer and how to avoid it is quite challenging. Therefore, this research has suggested a technique for determining whether a transformer requires maintenance or is in a normal state.

Based on DGA analysis data, the classification technique LS-SVM is provided to identify and predict the type of faults that occur. The DGA data in this research is analyzed using the IEC ratio approach, which is based on key-gas ratio concentrations. The DGA data obtained from the power industry is employed as a training and testing data set for LS-SVM. Seven distinct categories of incipient defects are considered: partial discharge, discharge of low energy, discharge of high energy, thermal fault below 300ºC, thermal fault between 300ºC and 700ºC, thermal fault over 700ºC, and normal state.

2. Methodology

2.1. Dissolved Gas Analysis (DGA)

This work introduces the adoption of LS-SVM for classifying different types of faults in power transformers using data received from Dissolved Gas Analysis (DGA). The DGA data is acquired according to the IEC 60599 (2007) standard. The IEC ratio method employs five gases: Hydrogen (H_2), Methane (CH_4), Acetylene (C_2H_2), Ethylene (C_2H_4), and Ethane (C_2H_6). The gas ratios C_2H_2/C_2H_4, CH_4/H_2, and C_2H_4/C_2H_6 are determined from these gases [11]. Table 1 displays the elucidation of gas that is dissolved in the oil [1]. Table 2 displays the IEC standard that is utilized for interpreting the various forms of defects in power transformers. It includes three gas ratios that correlate to the optional fault analysis.

Table 1. Interpreting dissolved gases in transformer oil [1]

Gas Detected	Interpretation
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

Table 2. Classification of gas ratios and fault types according to IEC60599 (2007)

Fault Type	Fault Type Code	C_2H_2 / C_2H_4	CH_4 / H_2	C_2H_4 / C_2H_6
Partial Discharge	PD	<0.1	<0.1	<0.2
Discharge of Low Energy	D1	>0.1	0.1-0.5	>0.1
Discharge of High Energy	D2	0.6-2.5	0.1-1	>2
Thermal Fault below 300°C	T1	<0.1	>1	<1
Thermal Fault between 300°C and 700°C	T2	<0.1	>1	1-4
Thermal Fault above 700°C	T3	<0.1	>1	>4

2.2. Least Square-Support Vector Machine (LS-SVM)

Least-Square Support Vector Machines (LS-SVM) are highly effective for addressing nonlinear classification, function estimation, and density estimation problems. LS-SVM is a type of Support Vector Machine (SVM) that solves linear Karush-Kuhn-Tucker (KKT) problems. The goal of LS-SVM is to find the optimal hyperplane that separates the outcomes by decreasing the margin between the hyperplane and the data points [16].

The default configuration of the LS-SVM classifier is specifically designed to handle binary classification issues. In these situations, the data is divided by an ideal hyperplane that a group of support vectors determines. Support vectors are a subset of the training set that determines the border values separating the two classes [17].

This paper utilized LS-SVM as a multi-classifier to develop a model for categorizing power transformers into non-faulty and various fault types. The LS-SVM function is expressed in (1), where y represents the output vector and x represents the input vector. The function Ø(x) is a nonlinear mapping function utilized to transform input data into a space with a higher number of dimensions. The revised weight vector is represented as w, whereas the scalar threshold value is represented as c.

$$y(x) = w^T \emptyset(x) + c \quad (1)$$

Equation (2) represents the LS-SVM classification model, which is governed by (3). In (3), γ denotes the cost function, i refers to the input instance, and ei represents the error value. This study uses the radial basis kernel function. The Gaussian Radial Basis Kernel (RBF) function, as depicted in (4), utilizes the kernel function K(xi,xj) to enable the computation of dot products in high-dimensional feature spaces using low-dimensional data. The parameter σ represents the standard deviation.

$$\min \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{n} e_i^2 \quad (2)$$

$$e_i = y_i - (w^T \emptyset(x) + b), \quad i = 1,2,3 \dots . n \quad (3)$$

$$K(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma}\right) \quad (4)$$

In order to achieve the best level of accuracy in classification, it is necessary to fine-tune two parameters: gamma (γ) and sigma (σ^2). The optimal values of the gamma and sigma parameters in this research were determined by a mix of Coupled Simulated Annealing (CSA) and the conventional complex approach. The formula for computing classification accuracy is depicted in (5). N_p represents the count of accurate predictions, while T_p is the overall count of prediction data.

$$Accuracy = (N_p/T_p) \times 100\% \quad (5)$$

2.3. LS-SVM for Classification of Fault Types

The LS-SVM model utilizes gas ratios as input data and generates seven various forms of incipient failure output. The LS-SVM model was developed and tested to accurately categorize the faulty state of a power transformer into several categories. The LS-SVM model categorized the outputs into the following categories: “1” for partial discharge, “2” for discharge of low energy, “3” for discharge of high energy, “4” for thermal fault with temperature below 300°C, “5” for thermal fault with temperature ranging from 300°C to 700°C, “6” for thermal fault with temperature exceeding 700°C, and “7” for normal state. Table 3 displays the class code and the quantity of data utilized, as seen in Table 3.

The LS-SVM model’s methodology is illustrated in Figure 1. Prior to proceeding, it is necessary to load the input and output data. This article utilizes seven distinct permutations of training and testing data. Gamma and sigma values were thereafter selected optimally by the utilization of the cross-validation technique. The data will then undergo training using a classification algorithm. The mean accuracy will be computed once the training operation is finished. Should the convergence conditions fail to be met, the hyperparameter values will be altered. The procedure is iterated until convergence is achieved. Subsequently, the testing procedure commences. Ultimately, the correctness of the testing data will be computed.

Table 3. Representation of class codes and number of samples for the LS-SVM classifier

Fault Types	Class Code	No. of Data
Partial Discharge (PD)	1	17
Discharge of Low Energy (D1)	2	3
Discharge of High Energy (D2)	3	12
Thermal Fault below 300 °C (T1)	4	154
Thermal Fault between 300°C and 700°C (T2)	5	3
Thermal fault above 700°C (T3)	6	110
Normal (N)	7	6
Total Data		**305**

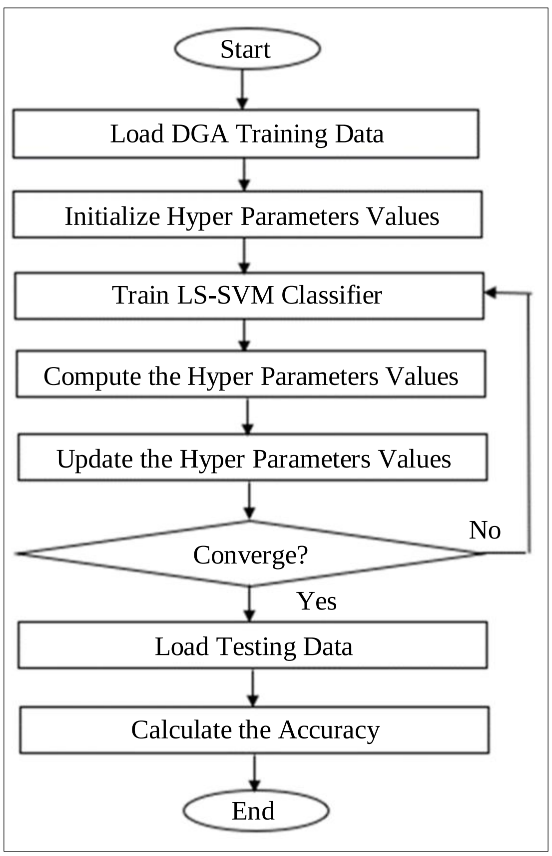

Fig. 1 Flowchart of LS-SVM for classification of fault types

3. Results and Discussion

This study presents actual data on flammable gasses found in transformer oil from utility companies. The dataset comprises measurements of dissolved gas levels and the status of power transformers. The conditions are partial discharge, low energy discharge, high energy discharge, thermal fault below 300°c, thermal fault between 300°c and 700°c, thermal fault over 700°c, and normal condition.

The simulation utilized a total of 305 samples. Multiple simulations were conducted for various case studies. In scenario 1, 90% of the data is allocated for training purposes, while the remaining 10% is reserved for testing. In Case 2, Case 3, Case 4, Case 5, Case 6, and Case 7, the training data consists of 85%, 80%, 75%, 70%, 65%, and 60% of the total data, respectively. The training outcomes for Cases 1 through 7 are illustrated in Figures 2 to 8 respectively.

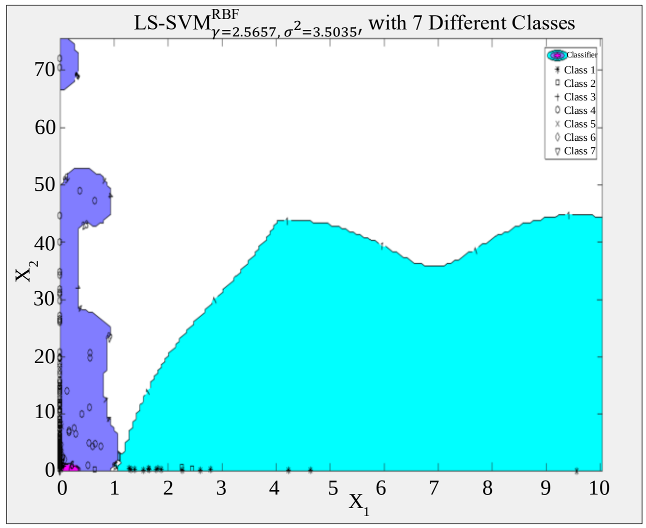

Fig. 2 Classification results of LS-SVM for case 1

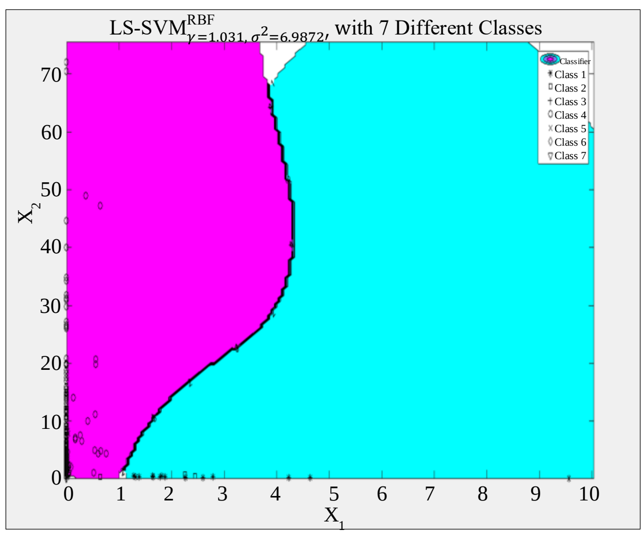

Fig. 3 Classification results of LS-SVM for case 2

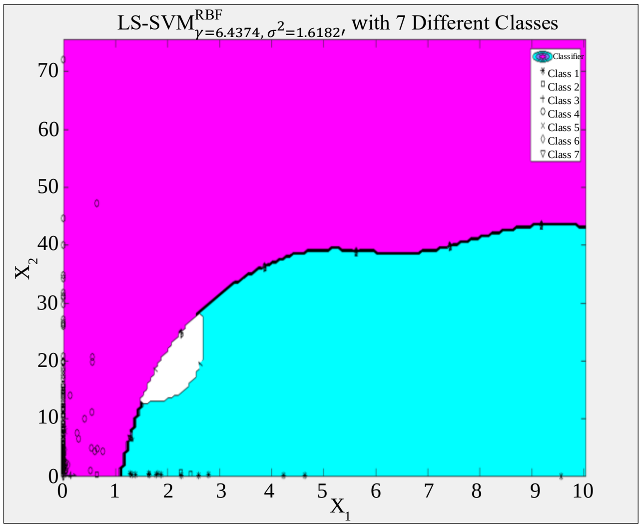

Fig. 4 Classification results of LS-SVM for case 3

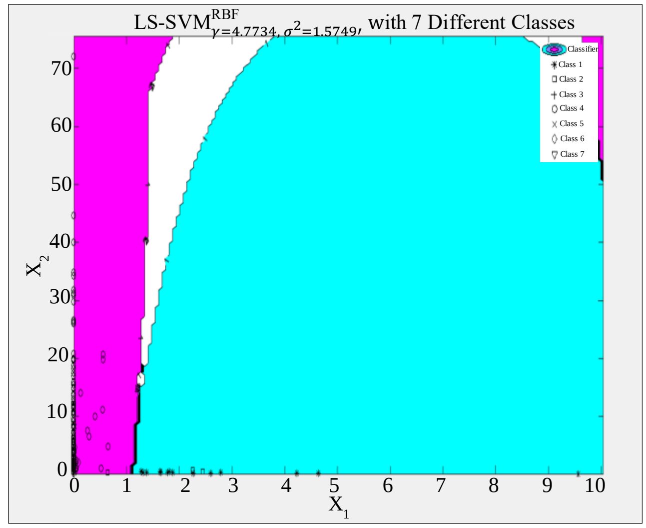

Fig. 7 Classification results of LS-SVM for case 6

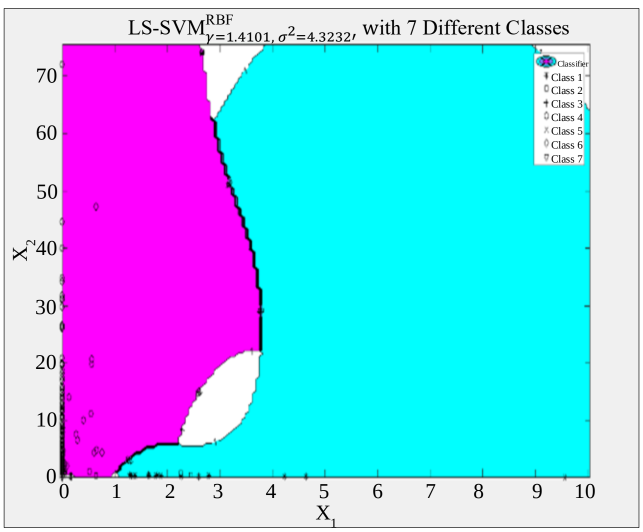

Fig. 5 Classification results of LS-SVM for case 4

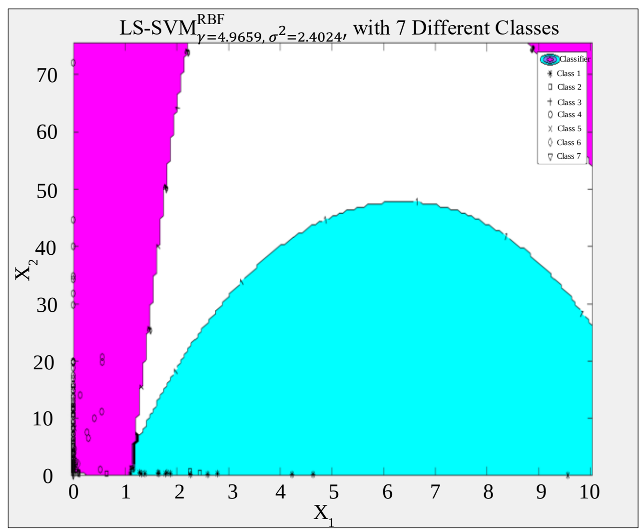

Fig. 8 Classification results of LS-SVM for case 7

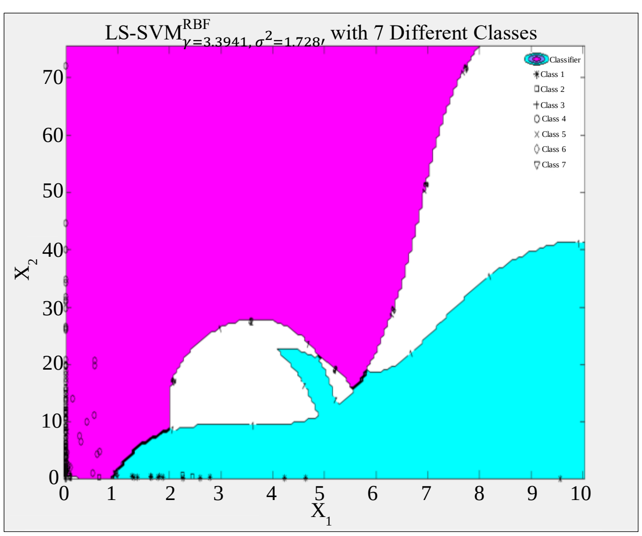

Fig. 6 Classification results of LS-SVM for case 5

The categorization accuracy exhibited variability across several cases, specifically Case 1 to Case 7, as depicted in Figures 2 to 8. The observed variation can be attributed to disparities in the quantity of training and testing data. Therefore, the selection of the amount of training and testing data is vital for attaining accurate classification models. Furthermore, the optimal selection of RBF parameters varies for each specific case study. The outcomes of the LS-SVM and DGA data-based classification precision evaluations for power transformer fault categories are displayed in Table 4.

The findings are reported for different ratios of training and testing data. The highest level of accuracy, reaching 97.37%, was attained by employing a blend of 229 training datasets and 76 testing datasets. These datasets were divided into a ratio of 75% for training and 25% for testing, derived from a total of 305 samples. The recommended values for gamma and sigma are 1.4101 and 4.3232, respectively. Table

4 unequivocally illustrates that the precision of fault detection for power transformers in each of the seven categories is above 90%. By allocating 75% of the data for training and 25% for testing, a high level of accuracy, precisely 97.37%, can be achieved. Figure 9 provides a summary of the findings for each case, including the total amount of training data utilized, as well as the number of TRUE and FALSE data points. The term “TRUE” refers to data that has been accurately predicted, whereas “FALSE” denotes data that has been inaccurately forecasted. Case 4 exhibits the lowest number of inaccurate data, with a total of 229 data points and 223 correctly identified data points. There were just six instances where the data points were categorized improperly. The most significant error is evident in Case 2, where 23 data points were erroneously classified. Comparable observations can be drawn in Case 1, where the imprecision is very significant. This may be attributed to insufficient testing data. Case 1 and Case 2 included 31 and 46 testing data points, respectively. The given results suggest that, apart from the parameter settings of RBF, the amount of training and testing data can also impact the accuracy of classification. Consequently, it necessitates careful and suitable selection.

Table 4. Performance of each classifier trained in LS-SVM

No.	Case Study	Percentage of Training Data (%)	Percentage of Testing Data (%)	Gamma	Sigma	Accuracy (%)	Error (%)
1.	Case 1	90	10	2.5657	3.5035	93.33	6.67
2.	Case 2	85	15	1.0310	6.9872	91.30	8.70
3.	Case 3	80	20	6.4374	1.6182	95.08	4.92
4.	Case 4	75	25	1.4101	4.3232	97.37	2.63
5.	Case 5	70	30	3.3941	1.7280	96.70	3.30
6.	Case 6	65	35	4.7734	1.5749	96.26	3.74
7.	Case 7	60	40	4.9659	2.4024	95.90	4.10

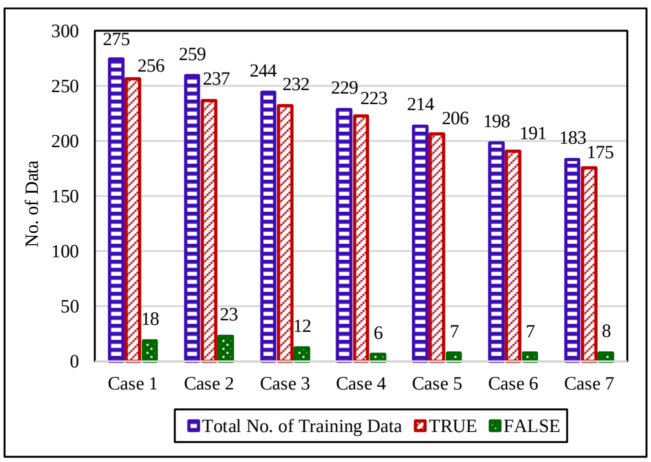

Fig. 9 Summarized results of LS-SVM for all cases

4. Conclusion

This work introduced the categorization of power transformer faults in a power system using the Least-Square Support Vector Machine (LS-SVM) technique. A total of 305 data samples were collected by Dissolved Gas Analysis (DGA). The DGA approach utilized IEC key-gas ratios to detect and pinpoint the first faults in transformers effectively. The dataset has seven distinct categories: partial discharge, discharge of low energy, discharge of high energy, thermal fault at temperatures less than 300°c, thermal fault with temperatures ranging from 300°c to 700°c, thermal fault at a temperature more than 700°c and normal condition. The efficacy of the LS-SVM model was evaluated through multiple case studies. The test results indicate that the LS-SVM classification model provided can attain a diagnosis accuracy of 97.37%, which is considered high. Utilizing it would be advantageous for the utilities as it would help prevent breakdowns and eliminate superfluous maintenance of power transformers.

Funding Statement

The funding for the research is provided by the Solar Research Institute (SRI) and the College of Engineering at University of Technology MARA Malaysia.

References

[1] Khmais Bacha, Seifeddine Souahlia, and Moncef Gossa, “Power Transformer Fault Diagnosis Based on Dissolved Gas Analysis by Support Vector Machine,” *Electric Power Systems Research*, vol. 83, no. 1, pp. 73-79, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[2] Mohamad Aiman Syahmi Rosli et al., "Power Transformer Incipient Fault Diagnosis Using Support Vector Machine," *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, IPOH, Malaysia, pp. 182-185, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mohd Syukri Ali et al., "Conventional Methods of Dissolved Gas Analysis Using Oil-Immersed Power Transformer for Fault Diagnosis: A Review," *Electric Power Systems Research*, vol. 216, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Rahman Azis Prasojo et al., "A Fuzzy Logic Model for Power Transformer Faults' Severity Determination Based on Gas Level, Gas Rate, and Dissolved Gas Analysis Interpretation," *Energies*, vol. 13, no. 4, pp. 1-20, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Syahiduddin Misbahulmunir, Vigna K. Ramachandaramurthy, and Yasmin H. MD. Thayoob, "Improved Self-Organizing Map Clustering of Power Transformer Dissolved Gas Analysis Using Inputs Pre-Processing," *IEEE Access*, vol. 8, pp. 71798-71811, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Secil Genc, and Serap Karagol, "Fuzzy Logic Application in DGA Methods to Classify Fault Type in Power Transformer," *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, pp. 1-4, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mohammad Ali Akhtar Siddique, and Shabana Mehfuz, "Artificial Neural Networks Based Incipient Fault Diagnosis for Power Transformers," *2015 Annual IEEE India Conference (INDICON)*, New Delhi, India, pp. 1-6, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[8] Nitchamon Poonnoy, Cattareeya Suwanasri, and Thanapong Suwanasri, "Neural Network Approach to Dissolved Gas Analysis for Fault Analysis in Power Transformers," *2022 International Electrical Engineering Congress (iEECON*), Khon Kaen, Thailand, pp. 1-4, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Kalinda D. Patekar, and Bhoopesh Chaudhry, "DGA Analysis of Transformer Using Artificial Neutral Network to Improve Reliability in Power Transformers," *2019 IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON)*, Chennai, India, pp. 1-5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] D. Saravanan et al., "Fault Prediction of Transformer Using Machine Learning and DGA," *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, Greater Noida, India, pp. 1-5, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Hazlee Azil Illias et al., "Fault Identification in Power Transformers Using Dissolve Gas Analysis and Support Vector Machine," *2021 IEEE International Conference on the Properties and Applications of Dielectric Materials (ICPADM)*, Johor Bahru, Malaysia, pp. 33-36, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Fernanda Rabelo Souza, and Bhuvana Ramachandran, "Dissolved Gas Analysis to Identify Faults and Improve Reliability in Transformers Using Support Vector Machines," *2016 Clemson University Power Systems Conference (PSC)*, Clemson, USA, pp. 1-4, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[13] Nutthaphan Boonsaner et al., "The Study and Analysis of Oil-Immersed Power Transformer by Using Artificial Neural Network for Designing Program Apply in the Industry of Testing Oil-immersed Transformers," *2020 8th International Conference on Condition Monitoring and Diagnosis (CMD)*, Phuket, Thailand, pp. 274-277, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Arian Dhini et al., "Intelligent Fault Diagnosis for Power Transformer Based on DGA Data Using Support Vector Machine (SVM)," *2018 3rd International Conference on System Reliability and Safety (ICSRS)*, Barcelona, Spain, pp. 294-298, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] Teguh Handjojo Dwiputranto, Noor Akhmad Setiawan, and Teguh Bharata Adji, "DGA-Based Early Transformer Fault Detection Using GA-Optimized ANN," *2021 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*, Jakarta, Indonesia, pp. 342-347, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Jagdeep Singh et al., "Condition Assessment of Power Transformer Using SVM Based on DGA," *2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), AIC-MITCSA 2016*, Baghdad, Iraq, pp. 1-5, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[17] Lov Kumar et al., "Effective Fault Prediction Model Developed Using Least Square Support Vector Machine (LSSVM)," *Journal of Systems and Software*, vol. 137, pp. 686-712, 2018. [CrossRef] [Google Scholar] [Publisher Link]