

ARTICLE

Transformer Fault Diagnosis: A Shallow Learning Approach for DGA-Based Incipient Fault Detection

Deepika Bhalla ¹ , **Avnesh Verma** ^{2*} 

¹ Department of Electrical Engineering, IKG Punjab Technical University, Kapurthala 144603, India

² Department of Instrumentation (Formerly USIC), Kurukshetra University, Kurukshetra 136119, India

ABSTRACT

Power transformers are exposed to electrical, thermal, and mechanical stresses during operation, leading to the degradation of insulation and the generation of dissolved gases. Utilities use IEEE and IEC standards use dissolved gas analysis (DGA) to detect incipient faults in oil-filled in-service transformers. Traditional gas ratio-based DGA methods, at times inconclusive diagnoses, limiting their effectiveness in scheduling preventive maintenance. This study presents the application of a shallow learning Backpropagation Neural Network (BP-NN) for assessing the condition of normal ageing and classification of incipient faults in oil-immersed power transformers. The model is trained using the concentrations (ppm) of five key gases—H₂, CH₄, C₂H₂, C₂H₄, and C₂H₆—as input features. The classified condition of a transformer is normal ageing and five fault type, namely partial discharge, low-energy and high-energy discharges, and thermal faults across two varying temperature ranges. The data set used for the classification of incipient faults within transformers is that where the fault type is confirmed by physical inspection. The 256 samples used in this work are from published sources, including the IEC TC10 database. The results achieved by the BP-NN demonstrate its capability to accurately classify normal ageing and diagnose five types of faults. For evaluating the performance of the trained NN, the IEEE/IEC method of classification, the benchmark used is the actual fault type. The shallow network of pattern recognition successfully identified the presence of normal ageing and the five fault types. The performance of the test set is 94.73%. The results highlight the potential of BP-NNs for enhanced transformer condition monitoring and early fault detection. As more high-quality

***CORRESPONDING AUTHOR:**

Avnesh Verma, Department of Instrumentation (Formerly USIC), Kurukshetra University, Kurukshetra 136119, India;
Email: verma.avnesh@rediffmail.com

ARTICLE INFO

Received: 3 November 2025 | Revised: 25 December 2025 | Accepted: 2 January 2026 | Published Online: 7 January 2025
DOI: <https://doi.org/10.30564/jeis.v8i1.11972>

CITATION

Bhalla, D., Verma, A., 2026. Transformer Fault Diagnosis: A Shallow Learning Approach for DGA-Based Incipient Fault Detection. *Journal of Electronic & Information Systems*. 8(1): 1–14. DOI: <https://doi.org/10.30564/jeis.v8i1.11972>

COPYRIGHT

Copyright © 2026 by the author(s). Published by Bilingual Publishing Group. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License (<https://creativecommons.org/licenses/by-nc/4.0/>).

labelled data become available, the diagnostic accuracy and robustness of the model are expected to improve further.

Keywords: Artificial Neural Networks; Backpropagation; Dissolved Gas Analysis; Power Transformer Diagnostics; Incipient Fault Detection

1. Introduction

In power transformers, incipient faults typically develop as a result of the gradual degradation of both solid and liquid dielectric materials. These faults arise due to the process of pyrolysis, hydrolysis, and oxidation of the insulation system. Incipient faults are generally categorized into thermal and electrical types. The degradation of dielectric materials leads to the generation of both combustible and non-combustible gases. These gases dissolve in the transformer's mineral oil and/or accumulate in the space above the oil level. If such conditions are not addressed promptly, they may lead to catastrophic transformer failure and interruption of the power supply. Common causes of gas generation include overheating due to overloading, failure of the cooling system, and electrical discharges such as arcing, corona, or low-energy sparking. The application of Dissolved Gas Analysis (DGA) for detecting incipient faults in oil-filled power transformers is a widely adopted diagnostic practice among power utilities; for which the IEC and IEEE standards are used irrespective of the transformer rating.

However, DGA techniques often yield conflicting or inconclusive diagnostic results, which may confuse operators or lead to an inability to determine the fault type. In such scenarios, maintenance decisions are typically based on expert judgment. To address these limitations, the application of Artificial Neural Networks (ANNs) for assisting fault diagnosis has been investigated.

While the IEEE/IEC method provide the highest diagnostic accuracy under standard conditions^[1] the use of Fuzzy logic^[2-5], Neural networks^[6-9], Support vector machines (SVM)^[10-12], Particle swarm optimisation^[13], Deep learning (DL)^[14,15] and (Machine learning)^[16,17] has demonstrated potential in overcoming the limitations posed by incomplete or overlapping gas ratio combinations.

Hybrid models using various Artificial Intelligence (AI) techniques have been extensively explored for the interpretation of DGA data; these include Genetic Algorithms^[18,19], SVM^[20], and Machine Learning^[21]. Studies have addressed

not only offline fault diagnosis but also online condition monitoring^[22,23]. In many cases, AI-based methods have successfully diagnosed faults where conventional methods failed to provide accurate classification. ML is being applied to transformer fault detection, exploring hybrid methods that combine conventional DGA methods with advanced models, and improved fault identification accuracy is achieved. Recently, DL algorithms have been used for fault classification. However, review of deep learning algorithms does not have capable capacity for DGA analysis as compared to shallow learning (SL) algorithms. The performance of deep learning (DL) algorithms was not significantly improved when compared to the results of shallow learning (SL) algorithms on the same DGA datasets^[24].

Table 1 analyzes some publications to decide upon the scope of work. There are fewer publications that consider normal ageing/no fault data. Most publications consider five gases. The classification is in a maximum of six conditions. In none of these did the data set used have all samples where the actual condition/fault type was known through physical inspection.

In this work, an SL Back propagation neural network (BP-NN) is employed due to its proven ability to recognize complex patterns for a small dataset where the classification pattern problems. The NN model is trained using the concentration (in ppm) of five gases: Hydrogen (H₂), Methane (CH₄), Acetylene (C₂H₂), Ethylene (C₂H₄), Ethane (C₂H₆). Based on these inputs, the model classifies the condition of a power transformer into six classes/categories, namely normal ageing/no fault, partial discharge, low energy discharge, high energy discharge, thermal fault of temperature < 700 °C, and thermal fault of temperature > 700 °C.

The dataset used consists of 256 fault samples sourced from the International Electro-technical Commission (IEC) TC10 database^[25], and published literature^[9,26-28] with actual fault known by physical inspection. The diagnostic performance of the IEEE/IEC method and the developed BP-NN is benchmarked against the actual fault. The IEEE/IEC

method is the combination of the respective standards considering three gas ratios: CH_4/H_2 , $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$, $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$ ^[29,30]. It is observed that the conventional IEEE/IEC method, being a ratio method, has its own limitations.

Table 1. Analysis of gases considered and condition accessed.

Publication	Gases Considered	Condition Assessed	Performance	AI Technique
Jin et al. (2024) ^[15]	Five gases	Partial discharge Discharge of 2 type Thermal fault of 3 type	accuracy of fault diagnostics 87%.	Convolution Neural Network (type of DL)
Saroja et al. (2023) ^[16]	Five gases	No fault Discharge of 2 type Thermal fault of 2 type	accuracy of fault diagnostics 81.12%	Quantitative descriptive analysis (QDA) (type of ML)
Wang et al. (2024) ^[11]	Seven gases	Partial discharge Discharge of 2 type Thermal fault of 3 type	mean absolute percentage error of 1.81% and the root mean square error of 0.707 $\mu\text{L/L}$	decomposition-cuckoo search-support vector regression (type of SVM)
Vidal et al. (2023) ^[19]	Five gases	Partial discharge Discharge of 2 type Thermal fault of 2 type	accuracy of fault diagnostics 95.18%	Genetic programming (GP) technique

The implementation of the neural network model has been carried out using MATLAB R2025b.

This paper is organized as follows:

- Section 2 outlines the prevalent DGA methodology applied in this study.
- Section 3 presents the method adopted in this work
- Section 4 presents the architecture of the three-layer BP-NN, including the training process and the activation functions used in the hidden layers.
- Section 5 describes the fault classification process using the trained BP-NN.
- Sections 6 and 7 include the results and discussions, and conclusions, respectively.

2. Dissolved Gas Analysis

Utilities employ various Dissolved Gas Analysis (DGA) techniques to assess the operational condition of power transformers. DGA is a critical diagnostic tool used to detect incipient faults based on the identification and quantification of gases dissolved in transformer oil, which are produced due to electrical and thermal stress.

The two most widely accepted DGA standards are:

- IEEE Standard C57.104:2019^[29]
- IEC Standard 60599:2022^[30]

The IEEE standard incorporates three main methods:

- The Key Gas Method
- Gas Ratio Methods (including Dornenburg's and Roger's methods)

- Graphical Methods, such as Duval's Triangle and Duval's Pentagon

The IEC standard also applies gas ratio methods, using the following ratios: CH_4/H_2 , $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$, and $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$.

The graphical methods, particularly Duval's Triangle and Pentagon, are visual tools used to classify fault types based on gas concentrations. However, each of these diagnostic techniques has certain limitations. These methods were primarily developed based on expert experience and empirical rules, which can lead to diagnostic inconsistencies. The limitations of the gas ratio methods include:

- Incomplete coverage of all possible ratio combinations
- Overlapping ratio ranges leading to diagnostic ambiguity

The other issues are variability in gas data due to sampling, measurement errors, and storage conditions. As a result, reliance solely on conventional DGA standards may result in misdiagnosis or inconclusive results, particularly in cases involving multiple or evolving fault types.

2.1. IEEE/IEC Method

In the IEEE/IEC method, incipient fault diagnosis is performed using predefined codes based on the ratios of specific hydrocarbon gases produced during developing thermal or electrical conditions within the transformer.

The three primary gas ratios defined in the IEEE/IEC methodology are:

- $R1 = \text{C}_2\text{H}_2/\text{C}_2\text{H}_4$

- $R2 = \text{CH}_4/\text{H}_2$
- $R3 = \text{C}_2\text{H}_4/\text{C}_2\text{H}_6$

These ratios are interpreted using diagnostic codes, which are matched against standard reference tables to determine the fault type.

The range of values for these ratios, along with the corresponding diagnostic codes, are provided in **Table 2** and the diagnostic interpretation based on these codes is outlined in **Table 3**. In this work, BP-NN is applied to overcome the incompleteness of the ratio combination and invalid ratios.

Table 2. Range of ratio and ratio codes for IEEE/IEC method^[6].

Range of Ratio	R1	R2	R3
<0.1	0	1	0
0.1–1.0	1	0	0
1.0–3.0	1	2	1
>3.0	2	2	2

Table 3. Diagnosis of fault type by IEEE/IEC method^[6].

R1	R2	R3	Fault Type	Fault Code
0	0	0	No Fault	F_0
0*	1	0	Low energy partial discharge	F_1
1	1	0	High energy partial discharge	F_2
1–2	0	1–2	Low energy discharge	F_3
1	0	2	High energy discharge	F_4
0	0	1	Low thermal fault < 150 °C	F_5
0	2	0–1	Medium thermal fault 150–700 °C	
0	2	2	High thermal fault > 700 °C	

*: The range is negligibly small.

3. Methodology

To assess the condition of a transformer based on DGA, data used is from published in literature. These samples are in the form of ppm level of the gas dissolved in the oil. The IEC TC10 database, which has 126 samples has been used for the past two decades as the standard data base for DGA based diagnosis. In this work the shallow BP-NN is trained using these 256 samples. The three layered neural network has five and six neurons in the input and output layer respectively. To decision of the number of neurons in the hidden layer is based on the general performance. The results of the for the NN are observed for the maximum output value across the three neurons and the average of which is taken as threshold (onset) value. The performance of the BP-NN results is compared with that of IEEE/IEC ratio method using the same data base of five input gases. The benchmark of the evaluation of the diagnosis is the actual fault type. **Figure 1** gives the flowchart on which the architecture of this work is based upon.

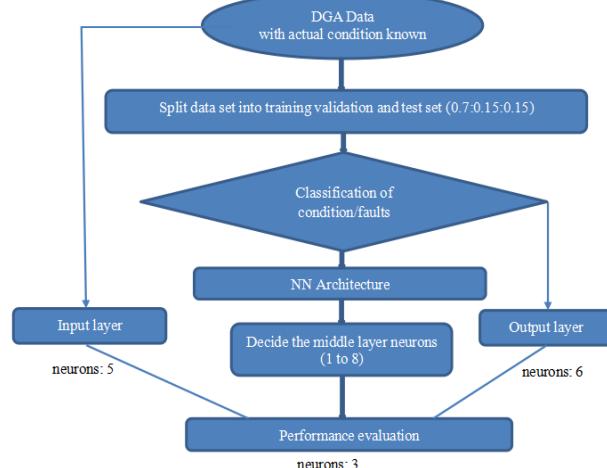


Figure 1. Flowchart for the NN architecture.

4. Network Architecture: Multilayer Backpropagation Neural Network

The multilayer backpropagation neural network (BP-NN) used in this study consists of three layers: an input layer,

a hidden layer, and an output layer. For the input layer, a linear transfer function is used, and for the hidden and output layers employ a sigmoidal transfer function is used. The sigmoidal activation function is mathematically defined as:

$$O = \frac{1}{1 + e^{-\lambda I}} \quad (1)$$

where I is the input to the neuron, λ is the sigmoidal gain, and O is the output of the neuron. The nonlinear mapping implemented by the three-layer structure is representing the input, hidden, and output layers respectively. The neurons in each layer are denoted by subscripts: I : input layer neurons, H : hidden layer neurons, and O : output layer neurons.

The input layer's output; considering linear activation function is:

$$\{O\}_I = \{I\}_I \quad (2)$$

$$l \times 1 \quad l \times 1$$

There is a synapse that connects hidden and input node. The weight of the synapse of the i^{th} input nodule with j^{th} hidden nodule is v_{ij} . The input of hidden nodule is the weighed addition of the output of the input neurons, the input to the p^{th} hidden nodule I_{Hp} is:

$$I_{Hp} = v_{1p}O_{I1} + v_{2p}O_{I2} + \dots + v_{Ip}O_{Il} \quad (3)$$

$$(p = 1, 2, 3, \dots, s)$$

The weight matrix between input and hidden neurons is:

$$[V] \quad (4)$$

$$l \times s$$

The input to the hidden nodule is:

$$\{I\}_H = [V]^T \quad \{O\}_I \quad (5)$$

$$s \times 1 \quad s \times l \quad l \times 1$$

The value of O_{Hp} , output of the p^{th} hidden nodule, O_{Hp} is found using sigmoidal activation function, it is:

$$O_{Hp} = \frac{1}{(1 + e^{-\lambda(I_{Hp} - \theta_{Hp})})} \quad (6)$$

where θ_{Hp} is the onset of the p^{th} hidden nodule.

The hidden nodule output is:

$$\{O\}_H = \left\{ \frac{1}{(1 + e^{-\lambda(I_{Hp} - \theta_{Hp})})} \right\} \quad (7)$$

$$\begin{array}{c} - \\ - \\ 1 \\ - \\ - \end{array}$$

If each component of the input of the hidden nodule is treated separately, then the output of the hidden nodule as given by Equation (7).

I_{Oq} is the input to the q^{th} output nodule, and it is:

$$[I_{Oq}] = w_{1q}O_{H1} + w_{2q}O_{H2} + \dots + w_{sq}O_{Hs} \quad (8)$$

$$(q = 1, 2, 3, \dots, n)$$

If the weighed matrix between the hidden and the output neurons is $[W]$, then output nodule's input is:

$$\{I\}_O = [W]^T \quad \{O\}_H \quad (9)$$

$$n \times 1 \quad n \times s \quad s \times 1$$

Then, O_{Oq} is the output of the q^{th} output nodule, and it is:

$$O_{Oq} = \frac{1}{(1 + e^{-\lambda(I_{Oq} - \theta_{Oq})})} \quad (10)$$

where, θ_{Oq} is the onset of the q^{th} nodule. This onset is handled by taking an extra O^0 nodule of hidden layer with -1 as output.

The output neurons are:

$$\{O\}_O = \left\{ \frac{1}{(1 + e^{-\lambda(I_{Oq} - \theta_{Oq})})} \right\} \quad (11)$$

$$\begin{array}{c} - \\ - \\ 1 \\ - \\ - \end{array}$$

The learning algorithms for a neural network is supervised.

Network Training

The samples need to distribute the available data into subsets, namely: training set, cross-validation set, and testing set. When the data set (I_p, y_p) , $p = 1, 2, 3, \dots, K$ has p patterns; the input $I_p \in R^N$, and target $y_p \in R$. The training set has s hidden neurons. The weight of the synapse v_{ij} is from j (input nodule) to i (hidden nodule). The weight of the synapse to the output node from the hidden node i is w_i . The hidden node activation function $f(x)$ is sigmoidal function.

For input I_p activation value A_{ip} The hidden neurons and their projected value are calculated by function \tilde{y}_p as

follows:

$$A_{ip} = f \left(\sum_{j=i}^N v_{ij} I_{jp} \right) = f(g_{ip}) \quad (12)$$

$$\tilde{y}_p = \sum_{i=1}^l w_i A_{ip} \quad (13)$$

Where I_{jp} is the value for the pattern p with input j .

5. Training of BP-NN for Incipient Fault Diagnosis

A highly non-linear activation function is employed in this work to enhance the classification capability of the neural network. Specifically, the sigmoidal function is used in combination with the gradient descent learning algorithm, which has proven effective for training backpropagation neural networks. The learning approach adopted is supervised learning. To avoid issues such as shallow convergence and overfitting, the architecture is deliberately kept shallow, consisting of three layers.

5.1. Fault Classification Scheme

The fault types are derived from the IEEE/IEC methods. However, due to a limited number of samples, certain categories are merged for condition of transformer classification by BP-NN:

- The low and high energy partial discharge faults are combined into a single class.
- The thermal faults at temperatures below 700 °C are grouped together.

This results in six final output attributes used for training the BP-NN for assessing the condition of transformer. The mapping of the no fault and five classes of faults by

the IEEE/IEC method and the target/output attributes for the BP-NN are given in **Table 4**.

Table 4. Mapping fault codes to BP-NN classification.

Actual Fault (Code)	Fault Type for BP-NN
F ₀	No Fault normal ageing (NF)
F ₁	Partial discharge (PD)
F ₃	Low energy discharge (LED)
F ₄	High energy discharge (HED)
F ₅	Medium temperature fault < 150 °C (MTH)
F ₆	High temperature > 700 °C (HTH)

5.2. Dataset and Pre-Processing

The dataset consists of 256 samples with known fault classifications, sourced from published literature^[6,23–25]. The classification of condition of transformer is initially done using the IEEE/IEC method and merging similar types of faults were sufficient samples were not available, and then mapped to the BP-NN fault categories. The dataset is divided into training set, validation set and test set manually; while ensuring the splitting of the data in the proportion of (0.7:0.15:0.15 and also that there are sufficient samples of each type of fault in each set, which cannot be achieved by random splitting of data. 70% data in training set large enough to capture complexity but still leaves enough data for evaluation. 15% data in validation set is sufficient to provide feedback without taking too much away from training an also ensures that the model adjustment are based on representative sample. And 15% data in test set statistically is a meaningful sample size and prevents misleading results that could occur if the set was smaller^[25–28].

This ensured network is trained and validated effectively while maintaining generalization capability. **Table 5** provides the detailed mapping of IEEE/IEC fault codes to BP-NN fault classifications, as well as the distribution of samples across the training, validation, and test sets.

Table 5. Mapping of fault codes and dataset distribution.

Fault Code (BP-NN)	Actual Fault (Code)	Total Number of Samples	Training Set	Validation Set	Test Set
NF	F ₀	22	12	5	5
PD	F ₁	18	12	3	3
LED	F ₃	51	37	7	7
HED	F ₄	71	53	9	9
LTH	F ₅	48	36	6	6
MTH	F ₆	46	30	8	8
Total		256	200	38	38

One hidden layer multilayer feedforward network is trained using five gases are the distinct input attributes. Due

to the number of inputs being limited pruning is not required. Training is done with the concentration of five gases CH₄,

C_2H_6 , C_2H_4 , C_2H_2 and H_2 , as the input attributes and six conditions as the output attributes.

5.3. Network Architecture Inputs and Output Attributes

The input layer receives five attributes, representing the normalised value of the gas concentrations (in ppm) in oil. The following dissolved gases are used as input attributes: Hydrogen (H_2), Methane (CH_4), Ethylene (C_2H_4), Ethane (C_2H_6), Acetylene (C_2H_2). The hidden layer is decided using sigmoidal activation functions to introduce non-linearity. The output layer contains six neurons, corresponding to the normal condition and five types of distinct faults. **Table 6**

maps output attributes are mapped to the condition of the transformer.

5.4. BP-NN Applied to DGA Database and Its Analysis

Code for DGA with a neural network suitable for a pattern recognition problem is used is available in **Appendix A**. Since the problem is of pattern recognition with low data supervised learning so the training function used is Levenberg–Marquardt backpropagation. The neural network used was tested for one to eight neurons in the middle layer. **Table 7** gives the percentage error, number of epochs, and performance.

Table 6. Mapping condition/fault codes and target/output attributes.

Actual Condition/Fault (Code)	Condition/Fault Type for BP-NN	Output Attributes
F_0	NF	1 0 0 0 0 0
F_1	PD	0 1 0 0 0 0
F_3	LED	0 0 1 0 0 0
F_4	HED	0 0 0 1 0 0
F_5	MTF	0 0 0 0 1 0
F_6	HTF	0 0 0 0 0 1

Table 7. Error percentage and performance of the middle layer for a three-layer.

Neurons in the Middle Layer	1	2	3	4	5	6	7	8
% Error	0.4375	0.3242	0.2188	0.3359	0.2148	0.3086	0.2266	0.2578
Performance	0.0943	0.0748	0.0585	0.0814	0.0539	0.0906	0.0544	0.0620
Number of epochs	30	43	37	21	51	29	73	139
Best epoch	24	37	31	15	48	23	67	133
Best training performance value	0.0942	0.0748	0.0542	0.0785	0.0443	0.0917	0.0468	0.0593
Best validation performance	0.0904	0.0582	0.0617	0.0865	0.0650	0.0886	0.0545	0.0331
Test performance value	0.0982	0.09111	0.0758	0.0905	0.0884	0.0874	0.0903	0.0835

Evaluation of **Table 7**: 1–2 neurons: Too few → high error, weak generalization, 3 neurons: Excellent balance—low error, strong training, and best test performance, 5 neurons: Very good training fit, but validation/test slightly weaker than 3 neurons, 6 neurons: Underperforming across metrics, 7 neurons: Good training, but test performance drops (possible overfitting), 8 neurons: Outstanding validation performance, but test performance not as strong as 3 neurons.

Analysing the above for decision on middle layer: The best overall generalization (lowest test error, strong training/validation) is achieved by 3 neurons. The best validation performance, but risk of overfitting (test weaker) is with 8 neurons. The strong training fit, but slightly weaker generalization than 3 neurons is with 5 neurons. However, 3 hidden

neurons balances accuracy, performance, and generalization most effectively.

5.5. Hyper-Parameters for Supervised Learning (Pattern Recognition)

1. Architecture:

Number of layers → 3

Number of neurons → Input: 5, middle: 6, output: 6

Activation function → sigmoidal

Training function → Levenberg–Marquardt algorithm

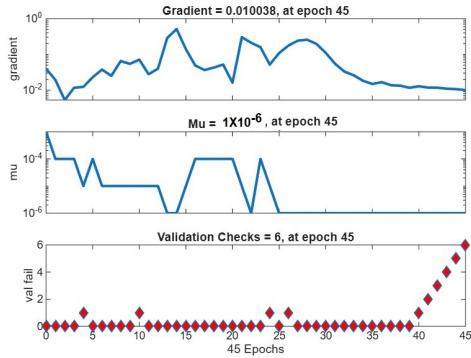
2. Training Parameters:

epochs → 1000

goal → performance goal (default 0).

training time → Inf

`min_grade` → 1.0000×10^{-7}
`max_fail` → 6
`mu` → 1.0000×10^{-3}
`mu_dec` → 0.1000
`mu_inc` → 10
`mu_max` → 1.0000×10^7

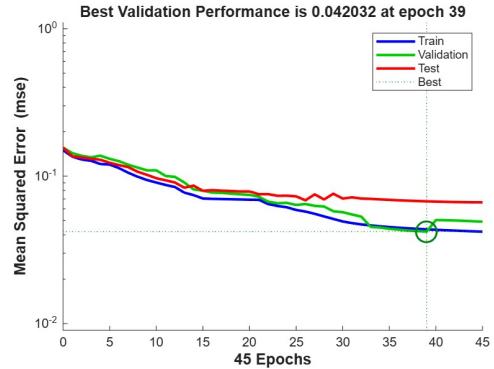


(a) Gradient and the validation checks.

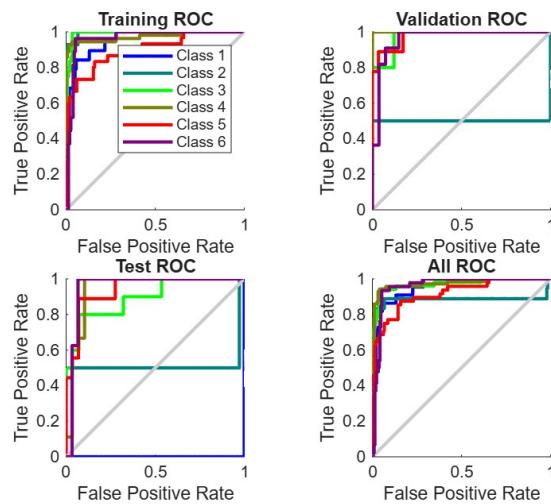
3. Performance:

`perform_Fnc` → mse
`derivFnc` → default (Jacobian matrices)

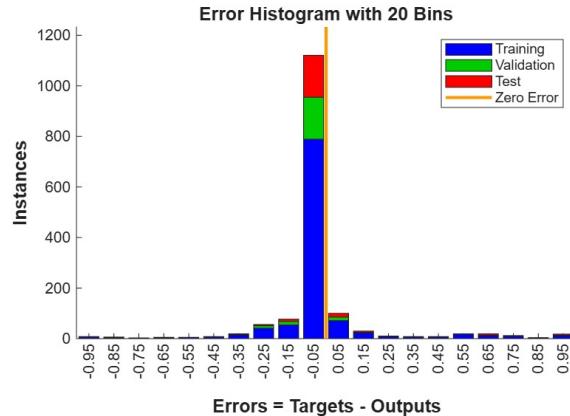
The graph's computational efficiency for the three-layer BP-NN with three neurons in the middle layer is shown in **Figure 2**.



(b) Performance Graph.



(c) Receiving Operating Characteristics.



(d) Error Histogram.

Figure 2. The performance of the three-layer BP-NN.

6. Results

6.1. IEEE/IEC Method

The result of the IEEE/IEC ratio method for all the 256 samples used in this work is appended in **Table 8**. In the absence (ppm being 0) of any one or more of three gases (H_2 , C_2H_4 and C_2H_6) the value of the ratio is invalid, in this dataset there are 28 such samples. Further for 58 samples the

ratio code does not exist and 43 samples were wrongly diagnosed. The accuracy of the IEEE > IEC method is 51.56% for the complete dataset.

6.2. BP-NN Method

The normalised confusion matrix of all data is given in **Figure 3**.

Table 8. Results of IEEE/IEC method.

Total Number of Samples	Correct Diagnosis	Ratio Undefined	Ratio Range not Classified	Incorrect Diagnosis	% Result
256	132	23	58	43	51.56%

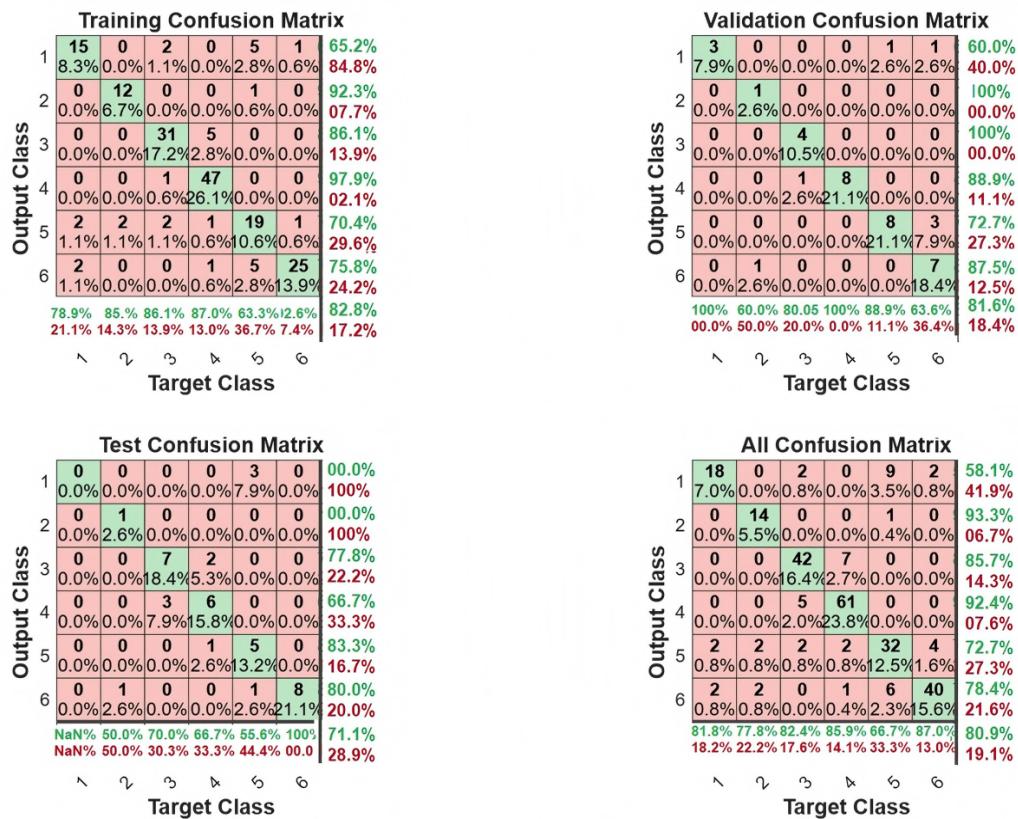


Figure 3. The confusion matrix of the dataset.

It is observed that:

- Of the 22 samples of no fault condition there is incorrect prediction of 2 samples each to be medium and high thermal fault.
- Of the 18 samples of partial discharge, there is incorrect prediction of 2 samples each to be medium and high thermal fault.

- Of the 51 samples of low energy discharge, there is incorrect prediction of 5 samples to be high energy discharge, and 2 samples each of partial discharge and high thermal fault.
- Of the 71 samples of high energy discharge, there is incorrect prediction of 7 samples to be low energy discharge, and 2 samples to be medium thermal fault and

1 sample to be high thermal fault.

- Of the 48 samples of medium thermal fault there is incorrect prediction of 9 samples to be normal ageing, and 6 samples to be high thermal fault and 1 sample to be partial discharge.
- Of the 46 samples of high thermal fault there is incorrect prediction of 2 samples to be normal ageing, and 4 samples to be medium thermal fault.

The overall prediction of the dataset is 80.9%.

6.3. Analysis of Test Set

The ppm concentrations of dissolved gases in all samples of the test set used in Backpropagation Neural Network (BP-NN) for assessing normal condition and five type of incipient fault classification are available in **Appendix B**. The value from the BP-NN corresponding to the output neurons are given in **Table 9**. A threshold (onset) value of 0.8137 has been defined as the minimum output required to indicate the presence of a fault type.

Table 9. Output of BP-NN for test set.

Sample No.	Actual Fault	Output Neuron 1	Output Neuron 2	Output Neuron 3	Output Neuron 4	Output Neuron 5	Output Neuron 6	Fault Type by BP-NN	Remarks
1.	F0	4.01×10^{-1}	3.22×10^{-2}	1.98×10^{-1}	4.53×10^{-2}	2.74×10^{-1}	4.99×10^{-2}	NF	Correct
2.	F0	6.44×10^{-1}	5.27×10^{-2}	5.1×10^{-2}	4.36×10^{-4}	2.49×10^{-1}	2.55×10^{-3}	NF	Correct
3.	F0	9.4×10^{-1}	1.44×10^{-2}	9.02×10^{-3}	5.26×10^{-6}	3.63×10^{-2}	9.6×10^{-5}	NF	Correct
4.	F0	4.20×10^{-1}	4.08×10^{-2}	1.91×10^{-1}	2.36×10^{-2}	2.93×10^{-1}	3.18×10^{-2}	NF	Correct
5.	F0	9.24×10^{-1}	1.66×10^{-2}	9.15×10^{-3}	7.91×10^{-6}	5.01×10^{-2}	1.43×10^{-4}	NF	Correct
6.	F1	2.11×10^{-8}	7.78×10^{-1}	6.40×10^{-6}	1.43×10^{-12}	2.22×10^{-1}	2.11×10^{-10}	PD	Correct
7.	F1	0.00×10^0	9.88×10^{-1}	1.17×10^{-2}	0.00×10^0	0.00×10^0	0.00×10^0	PD	Correct
8.	F1	0.00×10^0	1.00×10^0	2×10^{-15}	0.00×10^0	6.5×10^{-5}	0.00×10^0	PD	Correct
9.	F2	1.08×10^{-2}	6.84×10^{-3}	8.40×10^{-1}	1.22×10^{-1}	1.05×10^{-2}	9.15×10^{-3}	LED	Correct
10.	F2	0.00×10^0	0.00×10^0	1.00×10^0	1.34×10^{-5}	0.00×10^0	0.00×10^0	LED	Correct
11.	F2	4.80×10^{-14}	4.77×10^{-11}	9.13×10^{-1}	8.65×10^{-2}	1.00×10^{-14}	1.26×10^{-9}	LED	Correct
12.	F2	0.00×10^0	0.00×10^0	1.00×10^0	1.98×10^{-7}	0.00×10^0	0.00×10^0	LED	Correct
13.	F2	2.62×10^{-2}	6.52×10^{-3}	8.52×10^{-1}	9.54×10^{-2}	1.02×10^{-2}	9.41×10^{-3}	LED	Correct
14.	F2	0.00×10^0	0.00×10^0	1.00×10^0	3.28×10^{-5}	0.00×10^0	0.00×10^0	LED	Correct
15.	F2	7.02×10^{-5}	2.34×10^{-4}	9.71×10^{-1}	2.81×10^{-2}	3.51×10^{-5}	1.28×10^{-4}	LED	Correct
16.	F3	9.83×10^{-8}	1.90×10^{-6}	7.57×10^{-1}	2.43×10^{-1}	9.08×10^{-8}	1.99×10^{-5}	LED, HED	Correct
17.	F3	0.00×10^0	0.00×10^0	3.01×10^{-9}	1.00×10^0	0.00×10^0	0.00×10^0	HED	Correct
18.	F3	6.00×10^{-14}	1.20×10^{-11}	1.87×10^{-2}	9.81×10^{-1}	1.12×10^{-13}	6.55×10^{-8}	HED	Correct
19.	F3	0.00×10^0	0.00×10^0	4.54×10^{-9}	1.00×10^0	3×10^{-15}	4.36×10^{-8}	HED	Correct
20.	F3	0.00×10^0	0.00×10^0	9.65×10^{-11}	1.00×10^0	0.00×10^0	0.00×10^0	HED	Correct
21.	F3	0.00×10^0	0.00×10^0	0.00×10^0	1.00×10^0	0.00×10^0	2.32×10^{-13}	HED	Correct
22.	F3	6.47×10^{-6}	2.91×10^{-6}	1.02×10^{-1}	8.97×10^{-1}	2.56×10^{-6}	8.83×10^{-4}	HED	Correct
23.	F3	0.00×10^0	0.00×10^0	9.61×10^{-4}	9.99×10^{-1}	0.00×10^0	0.00×10^0	HED	Correct
24.	F3	0.00×10^0	0.00×10^0	1.99×10^{-6}	1.00×10^0	0.00×10^0	0.00×10^0	HED	Correct
25.	F4	0.00×10^0	2.10×10^{-2}	0.00×10^0	0.00×10^0	9.79×10^{-1}	0.00×10^0	MTH	Correct
26.	F4	2.65×10^{-1}	5.45×10^{-2}	2.14×10^{-1}	3.14×10^{-2}	3.98×10^{-1}	3.72×10^{-2}	MTH,NF	Correct
27.	F4	3.00×10^{-1}	5.34×10^{-2}	2.13×10^{-1}	2.68×10^{-2}	3.73×10^{-1}	3.35×10^{-2}	MTH,NF	Correct
28.	F4	2.51×10^{-3}	1.83×10^{-4}	4.77×10^{-6}	1.76×10^{-2}	6.82×10^{-1}	2.98×10^{-1}	MTH,HTH	Correct
29.	F4	1.55×10^{-13}	0.00×10^0	0.00×10^0	2.87×10^{-6}	7.36×10^{-5}	1.00×10^0	HTH	Correct
30.	F4	8.62×10^{-1}	2.88×10^{-2}	6.95×10^{-3}	5.1×10^{-6}	1.02×10^{-1}	1.39×10^{-4}	NF	Incorrect
31.	F5	3.15×10^{-3}	6.28×10^{-6}	8.26×10^{-7}	7.35×10^{-2}	7.9×10^{-2}	8.44×10^{-1}	HTH	Correct
32.	F5	5.69×10^{-3}	2.78×10^{-5}	3.08×10^{-6}	6.45×10^{-2}	1.81×10^{-1}	7.49×10^{-1}	HTH	Correct
33.	F5	7.26×10^{-2}	4.44×10^{-3}	9.93×10^{-3}	2.23×10^{-1}	3.72×10^{-1}	3.18×10^{-1}	MTH,HTH	Incorrect
34.	F5	1.00×10^{-15}	0.00×10^0	0.00×10^0	1.68×10^{-7}	7.78×10^{-2}	9.22×10^{-1}	HTH	Correct
35.	F5	0.00×10^0	0.00×10^0	0.00×10^0	9.75×10^{-2}	0.00×10^0	9.02×10^{-1}	HTH	Correct
36.	F5	1.00×10^{-15}	0.00×10^0	0.00×10^0	1.68×10^{-7}	7.78×10^{-2}	9.22×10^{-1}	HTH	Correct
37.	F5	0.00×10^0	0.00×10^0	0.00×10^0	7.4×10^{-3}	0.00×10^0	9.93×10^{-1}	HTH	Correct
38.	F5	0.00×10^0	0.00×10^0	0.00×10^0	9.75×10^{-2}	0.00×10^0	9.02×10^{-1}	HTH	Correct

Note:

The results in **bold** indicate that the output coincides with the actual condition/fault.

The results in *italics* indicate that there are two outputs where the onset value is high; however, the higher value coincides with the actual condition/fault.

The results in underlined do not coincide with the actual fault.

In the test set:

- 28 samples showed onset values exceeding the threshold.
- For 8 samples of the remaining 10 samples the higher value coincided with the actual fault.
- Samples no. 30 the actual fault is medium thermal fault,

while the BP-NN outputs exceeded the threshold for no fault/normal ageing.

- Samples no. 33 the actual fault is high thermal fault, while the BP-NN outputs exceeded the threshold for none, however the order of fault indicated is medium thermal and high thermal in that order.

The diagnostic accuracy of the BP-NN on the test set is calculated to be 94.73%. These discrepancies suggest that the performance of the BP-NN could be further improved with the inclusion of more samples for normal ageing, and partial discharge and thermal fault in the dataset.

6.4. Discussion

The interpretation of the classification errors: it is found that for NF condition had the weakest performance (58.1%). There is heavy misclassification into HTF (9 cases) and some leakage into HED & MTF. Suggests NF features overlap with MTF and HTF. This feature that overlaps is absence or traces of C_2H_2 during normal ageing and thermal faults. The accuracy of PD very good (>90%), the model distinguishes these fault types clearly. The accuracy of LED very good (>90%), this can be attributed that there are more number of samples of HED in the dataset. The accuracy of HED is good (85.7%) there is misclassification with LED, this is on account of the gasses being released during electrical discharge being same. The thermal faults MTF and HTF have moderate accuracy (~73–78%). These are thermal faults the gases associated with MTF are CH_4 , C_2H_6 and C_2H_4 so classification will be an issue as there is fuzziness in the boundary. The HTF is associated with C_2H_2 .

The limitation of this method is that the normal aging and five type of faults are attempted simultaneously. Even the international standards (IEEE and IEC) DGA methods other than Rogers ratio method consider normal ageing along with fault diagnostics of power transformer. **Table 9** illustrates the Output of BP-NN for test results.

7. Conclusion

The developed Backpropagation Neural Network (BP-NN) demonstrates promising performance in the classification of incipient faults in power transformers using DGA data. Unlike conventional IEEE/IEC ratio methods, the BP-NN-based diagnostic model exhibits the capability to overcome

the limitations of the ratio being undefined due to no trace of C_2H_4 and/or H_2 and/or C_2H_6 . Also, the issue of the range of ratios is eliminated. Offer improved fault identification where conventional methods may be ambiguous or inconclusive. Furthermore, the BP-NN is trained with data representing normal operating conditions, hence it can be extended for use in online condition monitoring of transformers. As more high-quality labelled data—especially for partial discharge and thermal faults—becomes available, the diagnostic accuracy and robustness of the ANN model are expected to improve further. Also, if data of multiple faults is available, then it could be used in scenarios involving overlapping fault signatures or complex fault conditions.

Author Contributions

Conceptualization, D.B. and A.V.; methodology, D.B.; software, D.B.; validation, D.B. and A.V.; formal analysis, D.B. and A.V.; investigation, D.B.; resources, D.B.; data curation, D.B.; writing—original draft preparation, D.B. and A.V.; writing—review and editing, A.V.; visualization, A.V.; supervision, D.B. and A.V.; project administration, D.B.; Both authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding. This research did not receive any external funding. The work presented in the paper is entirely self-sustained and self-supported, with all funding contributed by the authors themselves.

Institutional Review Board Statement

The research was conducted in accordance with the approved research protocol for software-based analysis. As the study is entirely based on soft computing methods and does not involve human or animal subjects, Institutional Review Board approval was not applicable. All applicable ethical guidelines were strictly followed.

Informed Consent Statement

Not applicable. The study is based solely on software analysis and soft computing techniques and does not involve

human participants or animal subjects.

Data Availability Statement

The data used in this study were generated through software analysis and soft computing techniques and are included within the manuscript, as presented in **Tables 1–9**.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix A. Code for BP-NN

% Solve a Dissolved gas analysis problem with Neural Pattern Recognition

```

x = inp;
t = out;
% Training Function
trainFcn = 'trainlm';
% Create Network with one hidden layer
hiddenLayerSize = 6;
net = patternnet(hiddenLayerSize, trainFcn);

```

% Setup Division of Data for Training, Validation, Testing

```

N = 256;
trainInd = 1: round(0.7 * N);
valInd = round(0.7 * N) + 1: round(0.85 * N);
testInd = round(0.85 * N) + 1: N;
% Apply to your data
X_train = x(:, trainInd);
T_train = t(:, trainInd);
X_val = x(:, valInd);
T_val = t(:, valInd);
X_test = x(:, testInd);
T_test = t(:, testInd);
% Train the Network
[net, tr] = train(net, x, t);
% Test the Network
y = net(x);
e = gsubtract(t, y);
performance = perform(net, t, y);
tind = vec2ind(t);
yind = vec2ind(y);
percentErrors = sum(tind ~= yind) / numel(tind);

```

Appendix B

Table A1. DGA data of samples used in test set.

Sample No.	Dissolved Gas Data (ppm)					Actual Fault (Code)	Sample No.	Dissolved Gas Data (ppm)					Actual Fault (Code)
	H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆			H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆	
1	0	6	0	4	3	F0	20	2770	660	763	712	54	F3
2	30	32	0	3	63	F0	21	3090	5020	2540	3800	323	F3
3	11.82	3.12	1.98	0.67	120.8	F0	22	120	31	94	66	0	F3
4	14.7	11.275	0.2	2.7	10.5	F0	23	5900	1500	2300	1200	68	F3
5	21.54	3.8	0	0.98	113	F0	24	13,500	6110	4040	4510	212	F3
6	2240	168	0	0	25	F1	25	0	18900	330	540	410	F4
7	92,600	10,200	0	0	0	F1	26	55	22	0	2.6	0.5	F4
8	36,036	619	0	233	58	F1	27	86	8	0	2.5	2.5	F4
9	131	32	38.7	18.8	7.3	F2	28	130	140	0	120	24	F4
10	4230	690	1180	196	5	F2	29	111	559	0	707	243	F4
11	595	80	244	89	9	F2	30	14	44	1	7	124	F4
12	41	112	4536	254	0	F2	31	15.9	55.98	0.21	137.25	22.33	F5
13	60	5	29	6	1	F2	32	21	75	0	126	24	F5
14	890	110	700	84	3	F2	33	150	22	11	60	9	F5
15	253	21.5	72.4	16.1	5.9	F2	34	290	1260	8	820	231	F5
16	390	62.6	133.6	65.3	12.8	F3	35	740	2227	42	4258	567	F5
17	1570	1110	1830	1780	175	F3	36	290	1260	8	820	231	F5
18	545	130	239	153	16	F3	37	860	1670	40	2050	30	F5
19	1500	395	323	395	28	F3	38	740	2227	42	4258	567	F5

References

[1] Farag, A.S., Mohandes, M., Al-Shaikh, A., 2001. Diagnosing Failed Distribution Transformers Using Neural Networks. *IEEE Transactions on Power Delivery*. 16(4), 631–636. DOI: <https://doi.org/10.1109/61.956749>

[2] Verma, A., Vishnu, A.T., 2013. Optimum Fuzzy Based Approach to Improve the Instrument's Performance Affected by Environmental Conditions. *Serbian Journal of Electrical Engineering*. 10(2), 309–318. DOI: <https://doi.org/10.2298/SJEE121219006V>

[3] Genc, S., Karagol, S., 2020. Fuzzy Logic Application in DGA Methods to Classify Fault Type in Power Transformer. In Proceedings of the International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Ankara, Turkey, 26–28 June 2020; pp. 1–4. DOI: <https://doi.org/10.1109/HORA49412.2020.9152896>

[4] Malik, I.M., Sharma, A., Naayagi, R.T., 2023. A Comprehensive and Practical Method for Transformer Fault Analysis With Historical Data Trend Using Fuzzy Logic. *IEEE Transactions on Dielectrics and Electrical Insulation*. 30(5), 2277–2284. DOI: <https://doi.org/10.1109/TDEI.2023.3286795>

[5] Verma, A., Tiwari, V.A., 2019. A Fuzzy Logic Controller to Minimize the Effect of Input Temperature Fluctuations on the Deviation of Relative Humidity to Avoid Instrument's Deterioration. *International Journal of Research in Advent Technology*. 7(6), 1–6. DOI: <https://doi.org/10.32622/ijrat.76201948>

[6] Thang, K.F., Aggarwal, R.K., Esp, D.G., et al., 2000. Statistical and Neural Analysis of Dissolved Gases in Power Transformers. In Proceedings of the Eighth International Conference on Dielectric Materials, Measurements and Applications, Edinburgh, UK, 17–21 September 2000; pp. 324–329.

[7] Guardado, J.L., Naredo, J.L., Moreno, P., et al., 2001. A Comparative Study of Neural Network Efficiency in Power Transformer Diagnosis Using Dissolved Gas Analysis. *IEEE Transactions on Power Delivery*. 16(4), 643–647.

[8] Souahlia, S., Bacha, K., Chaari, A., 2012. MLP Neural Network-Based Decision for Power Transformers Fault Diagnosis Using an Improved Combination of Rogers and Doernenburg Ratios DGA. *International Journal of Electrical Power and Energy Systems*. 43(1), 1346–1353. DOI: <https://doi.org/10.1016/j.ijepes.2012.05.067>

[9] Bhalla, D., Bansal, R.K., Gupta, H.O., 2012. Function Analysis Based Rule Extraction from Artificial Neural Networks for Transformer Incipient Fault Diagnosis. *International Journal of Electrical Power and Energy Systems*. 33, 1196–1203.

[10] Zou, D., Li, Z., Quan, H., et al., 2023. Transformer Fault Classification for Diagnosis Based on DGA and Deep Belief Network. *Energy Reports*. 9(Suppl. 12), 250–256. DOI: <https://doi.org/10.1016/j.egyr.2023.09.183>

[11] Wang, N., Li, W., Li, J., et al., 2024. Prediction of Dissolved Gas Content in Transformer Oil Using the Improved SVR Model. *IEEE Transactions on Applied Superconductivity*. 34(8), 1–4. DOI: <https://doi.org/10.1109/TASC.2024.3463256>

[12] Wu, Y., Sun, X., Zhang, Y., et al., 2022. A Power Transformer Fault Diagnosis Method Based on Hybrid Improved Seagull Optimization Algorithm and Support Vector Machine. *IEEE Access*. 10, 17268–17286.

[13] Taha, I.B.M., Hoballah, A., Ghoneim, S.S.M., 2020. Optimal Ratio Limits of Rogers' Four-Ratios and IEC 60599 Code Methods Using Particle Swarm Optimization Fuzzy-Logic Approach. *IEEE Transactions on Dielectrics and Electrical Insulation*. 27(1), 222–230.

[14] Moradi, E., Elsisi, M., Mahmoud, K., et al., 2025. Robust Deep Neural Network-Based Internet of Things for Power Transformer Fault Diagnosis Under Imbalanced Data and Uncertainties. *International Journal of Electrical Power and Energy Systems*. 168, 110731. DOI: <https://doi.org/10.1016/j.ijepes.2025.110731>

[15] Jin, L., Kim, D., Chan, K.Y., et al., 2024. Deep Machine Learning-Based Asset Management Approach for Oil-Immersed Power Transformers Using Dissolved Gas Analysis. *IEEE Access*. 12, 27794–27809.

[16] Saroja, S., Haseena, S., Madavan, R., 2023. Dissolved Gas Analysis of Transformer: An Approach Based on Machine Learning and Multi-Criteria Decision Making. *IEEE Transactions on Dielectrics and Electrical Insulation*. 30(5), 2429–2438. DOI: <https://doi.org/10.1109/TDEI.2023.3271609>

[17] Rao, U.M., Fofana, I., Rajesh, K.N.V.P.S., et al., 2021. Identification and Application of Machine Learning Algorithms for Transformer Dissolved Gas Analysis. *IEEE Transactions on Dielectrics and Electrical Insulation*. 28(5), 1828–1835. DOI: <https://doi.org/10.1109/TDEI.2021.009770>

[18] Nanfak, S., Eke, F., Meghnefi, F., et al., 2023. Hybrid DGA Method for Power Transformer Faults Diagnosis Based on Evolutionary k-Means Clustering and Dissolved Gas Subsets Analysis. *IEEE Transactions on Dielectrics and Electrical Insulation*. 30(5), 2421–2428. DOI: <https://doi.org/10.1109/TDEI.2023.3275119>

[19] Vidal, J.F., Castro, A.R.G., 2023. Diagnosing Faults in Power Transformers With Variational Autoencoder, Genetic Programming, and Neural Network. *IEEE Access*. 11, 30529–30545. DOI: <https://doi.org/10.1109/ACCESS.2023.3258544>

[20] Abdo, A., Liu, H., Mahmoud, Y., et al., 2024. Hybrid Model of Power Transformer Fault Classification Using C-Set and MFCM–MCSVM. *CSEE Journal of Power and Energy Systems*. 10(2), 672–685. DOI: <https://doi.org/10.17775/CSEEPES.2020.04010>

[21] Mashifane, L.D., Mendu, B., Monchusi, B.B., 2025. State-of-the-Art Fault Detection and Diagnosis in Power Transformers: A Review of Machine Learning and Hybrid Methods. *IEEE Access*. 13, 48156–48172. DOI: <https://doi.org/10.1109/ACCESS.2025.3546861>

[22] Butler-Purry, K.L., Bagriyanik, M., 2003. Identifying Transformer Incipient Events for Maintaining Distribution System Reliability. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6–9 January 2003; pp. 1–8.

[23] Izzularab, M.A., Aly, G.E.M., Mansour, D.A., 2004. On-line Diagnosis of Incipient Faults and Cellulose Degradation Based on Artificial Intelligence Methods. In Proceedings of the IEEE International Conference on Solid Dielectrics, Toulouse, France, 5–9 July 2004; pp. 767–770.

[24] Chen, H.C., Zhang, Y., 2025. Rethinking Shallow and Deep Learnings for Transformer Dissolved Gas Analysis: A Review. *IEEE Transactions on Dielectrics and Electrical Insulation*. 32(1), 3–10. DOI: <https://doi.org/10.1109/TDEI.2025.3526080>

[25] Duval, M., DePabla, A., 2001. Interpretation of Gas-in-Oil Analysis Using New IEC Publication 60599 and IEC TC 10 Databases. *IEEE Electrical Insulation Magazine*. 17(2), 31–41.

[26] Yang, H.T., Liou, C.C., Chow, J.H., 2001. Fuzzy Learning Vector Quantization Networks for Power Transformer Condition Assessment. *IEEE Transactions on Dielectrics and Electrical Insulation*. 8(1), 143–149.

[27] Lin, C.E., Ling, J.M., Huang, C.L., 1993. An Expert System for Transformers Fault Diagnosis Using Dissolved Gas Analysis. *IEEE Transactions on Power Delivery*. 8(1), 231–238.

[28] Tomsovic, K., Tapper, M., Ingvarsson, T., 1993. A Fuzzy Information Approach to Integrate Different Transformer Diagnostic Methods. *IEEE Transactions on Power Systems*. 8(3), 1638–1646.

[29] IEEE SA Standards Board, 2019. ANSI/IEEE Standard C57.104-2019: IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers. IEEE Standards Association: Piscataway, NJ, USA.

[30] IEC, 2022. IEC 60599-2022: Mineral Oil-Impregnated Electrical Equipment in Service: Guide to the Interpretation of Dissolved Free Gas Analysis. IEC: Geneva, Switzerland.