Comparative Study on the Accuracy of the Conventional DGA Techniques and Artificial Neural Network in Classifying Faults Inside Oil Filled Power Transformers

Gomotsegang Millicent Mokgosi

A research report submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science in Engineering

Johannesburg, 30 July 2024

Declaration

I declare that this dissertation is my own, unaided work, except where otherwise acknowledged. It is being submitted for the degree of Master of Science in Engineering in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Signed this ____ day of __________ 20___

____

Gomotsegang Millicent Mokgosi

Abstract

Power transformers are expensive yet crucial for power system reliability. As the installed base ages and failure rates rise, there is growing interest in advanced methods for monitoring and diagnosing faults to mitigate risks. Power transformer failures are often due to insulation breakdown from harsh conditions like overloading, that leads to prolonged outages, economic losses and safety hazards. Dissolved Gases Analysis (DGA) is a common diagnostic tool for detecting faults in oil-filled power transformers. However, it heavily relies on expert interpretation and can yield conflicting results, complicating decision-making. Researchers have explored Artificial Intelligence (AI) to address these challenges and improve diagnostic accuracy. This study investigates using Machine Learning (ML) techniques to enhance DGA for diagnosing power transformers. It employs an Artificial Neural Network (ANN) with Feed Forward Back Propagation, a Bayesian Regularizer for predictions, Principal Component Analysis (PCA) for feature selection and Adaptive Synthesizer (ADASYN) for data balancing. While traditional DGA methods are known for their accuracy and non-intrusiveness, they have limitations, particularly with undefined diagnostic areas. This research focuses on these limitations, to demonstrate that ANN provides more accurate predictions compared to conventional methods, with an average accuracy of 76.8% versus lower accuracies of 55% for Dornenburg, 40% for Duval, 38.4% for Roger and 31.8% for IEC (International Electrotechnical Commission) Methods. The study findings prove that ANN can effectively operate independently to improve diagnostic performance.

Acknowledgements

I would like to thank God for giving me the strength to do this research and for providing positive people around me to encourage me throughout my research.

I would like to thank my supervisor Prof Ken Nixon and co-supervisor Prof Cuthbert Nyamupangedengu for their expertise, guidance and encouragement throughout this journey.

A special thank you to my former Eskom Managers: Mr Andre Bekker for always showing interest in my research and being understanding when I needed time to go on campus. Not forgetting Mrs Amelia Mtshali for always believing in me and always encouraging me to take opportunities to further my studies whenever they were presented by Eskom.

I would like to thank my family for their support and always urging me to never give up no matter how hard it got. Special thanks to my partner Mosibudi Machaba for the late-night sleeps, proof reading my report and always believing in me even when I didn't at times.

To my sister, Pontsho Mokgosi thank you for allowing me to take time away when you took care of the kids so that I could focus on this research, I am forever grateful for you. To my son Tshiamo Mokgosi, thank you for understanding when our daily study time had to be shortened as I needed to focus on my research and for being independent in your own schoolwork while mommy was doing research. This is dedicated to you my boy, I love you.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

ADASYN Adaptive Synthesizer

AI Artificial Intelligence

ANFIS Adaptive Neuro Fuzzy Inference System

ANN Artificial Neural Network

BP Back Propagation

BR Bayesian Regularizer

D Discharge

D1 Discharge of Low Energy

D2 Discharge of High Energy

DGA Dissolved Gases Analysis

DP Degree of Polymerisation

FL Fuzzy Logic

FRA Frequency Response Analysis

HLPC High Performance Liquid Chromatography

KNN K-Nearest Neighbour

IOT Internet of Things

ML Machine Learning

MSE Mean Square Error

NB Naïve Beyers

PCA Principal Component Analysis

PD Partial Discharge

PPM Parts Per Million

RBFN Radial Basis Function network

RCM Reliability Centred Maintenance

RLC Resistance, Inductance and Capacitance

ROC Receiver Operating Curve

SFDA Smart Fault Diagnostic Approach

SVM Support Vector Modulation

TDCG Total Dissolved Combustible Gases

UHF Ultra High Frequency

1. Introduction

1.1. Background

A power transformer is one of the power systems' most important and costly assets. Its primary role is to connect two electrical circuits of different voltage levels through electromagnetic induction and to transmit electric power between the circuits. The major components (excluding the tap changer) are static while in operation, but their failure can be detrimental to the power system.

A power transformer's availability and reliability are critical in ensuring continuity of supply to customers. To reach its intended maximum lifespan and to ensure its reliability, it should have design safety margins, operate within its design limits (operating technical specification), have a proper maintenance plan and be protected from unexpected external faults in the network. The emerging technology advancements are making it possible to extend the life of the assets with improved monitoring and maintenance regime.

The power transformer is exposed to various stresses during operation. It is subjected to mechanical stress from short circuit currents, electrical stress from switch and lightning surges, thermal stresses from overloading and other stresses. While the transformer is designed to handle the stresses to a certain extent, over time, the components fail due to multiple exposures and fatigue. Due to regular operation, the insulating medium also degrades over time. When insulation has deteriorated to extreme levels, the current path gets destroyed then leakage current flows causing short circuit faults. This deterioration in insulation produces losses and further degradation of the insulation which then leads to electron discharges. When the gap between two conducting materials is bridged, the ultimate result is a flashover and arcing. Power transformer faults vary with energy dissipated during the fault. Energy dissipation is the highest in arcing, medium in overheating and less in corona discharges. At different temperature levels, the faults cause a gradual breakdown of insulation which releases gaseous decomposition products known as Hydrogen (H_2), Methane (CH_4), Ethane (C_2H_6), Ethylene (C_2H_4), Acetylene (C_2H_2), Carbon monoxide (CO) or Carbon dioxide (CO_2).

It has been reported that 75% of power transformer faults are due to insulation system problems [1]. In addition, the majority of these faults develop over a period of time. The biggest threats to power transformer insulation are heat, moisture and Oxygen. When these are combined, the transformer ageing process is accelerated immensely. Studies have shown that increasing the moisture level from 0.5% to 1% will reduce the transformer life by half if other parameters (heat and oxygen) are kept the same. Taking that into consideration, Cigre's transformer reliability survey of 2015 [2] has reported an average annual failure rate of 0,53% for substation power transformers. This translates to 5 power transformer failures per annum for a fleet of 1000 power transformers. To minimise the risk of failure of transformers and maximise on their lifecycle, proper asset care strategies should be prescribed. Condition monitoring tools are useful in monitoring the transformer's health while in operation. The condition analysis will lead to the correct diagnosis of potential faults. DGA (Dissolved Gases Analysis) is one of the condition monitoring techniques that is mostly preferred due to its effectiveness, economical and non-intrusive nature in diagnosis. DGA detects gases released in insulating oil when a transformer is subjected to different electrical stress. These gas concentrations are then tested, analysed and correlated to the nature and severity of the faults.

IEC (International Electrotechnical Commission) [3], IEEE (Institute of Electrical and Electronics Engineers) [4] and other international bodies have developed guidelines for interpretating DGA using computational and graphical methods. The techniques do not involve any mathematical formulation, they are solely based on expert knowledge and experience. Research in the subject of DGA has been effort to improve their accuracy.

1.2. Problem Statement

The challenges of DGA diagnostic methods have been covered vastly in the literature. Some of these include DGA not being able to predict multiple faults [5][6], no results for negligible gas quantities [5], conflicting diagnosis for the same sample, undefined areas and overlap on the IEC method for high and low energy discharge faults [3]. The limitations presented by the conventional DGA methods can potentially cause delays in decision making or wrong decisions being made. The time and accuracy factor can lead to transformer failures if the decision is not made timeously or a wrong decision is made. Unnecessary, costly and time-consuming repair maintenance will be undertaken if a wrong decision is made.

The large volume of maintenance data that plant specialists have to analyse daily can put even more pressure on decision making and result in even more errors in the corrective action.

There is an ageing installed asset base worldwide and the existence of global trends to decrease operational cost, digitalise and decentralise. Therefore, the adoption of intelligent techniques and sensor technology is a necessity to monitor and analyse the condition of the power transformer and to extend its lifecycle.

1.3. Research Question

This study will answer the question: To what extent can an ANN (Artificial Neural Network) algorithm be applied to the dissolved gases to improve the accuracy of fault classification of developing faults inside an oil-filled power transformer?

1.4. The Study Approach (Methodology)

1.4.1. Data Collection

To answer the research question, a dataset of 145 power transformer oil samples with their actual inspected conditions is sourced from the IEC TC 10 (IEC Technical Committee) [7] database and other credible literature sources. The IEC TC 10 is a database of faulty equipment inspected in service by experienced engineers and maintenance experts. The input dataset is randomly split into 70% for training, 20% for validation and 10% for testing. The dataset is made up of 5 features represented in the form of each transformer's dissolved gas concentrations (H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2) in parts per million, 145 observations and 5 class outputs corresponding to main fault conditions. This study follows a similar classification of main faults according to IEC 60599 publication where:

N – Normal degradation.

T_L- Temperature fault <300℃ evidenced by the paper turning brownish.

T_H - Temperature fault>300℃ evidenced by carbonised paper and oil, metal colourisation or fusion.

PD - Partial Discharge of low and high energy refers to PD (Partial Discharge) of the cold plasma (corona) type with possible X-Wax formation and of the sparking type inducing small, carbonised punctures in the paper.

D – Arcing including discharge of low and high energy evidenced by larger punctures in paper, tracking or carbon particles in oil. This includes discharges of high energy with power follow-through evidenced by extensive carbonisation and metal fusion.

1.4.2. Machine Learning Algorithm Selection and Application

The study explores the supervised machine learning ANN algorithm in Matlab™ R2000b simulation toolbox. The ANN classifier is selected for this study due to its robustness, ability to handle nonlinear dataset and ability to deal with inconsistencies in the dataset.

To deal with the imbalance in a dataset with minority class labels, a technique called Adaptive Synthesizer (ADASYN) is used. ADASYN is an oversampling method that generates synthetic samples using K-nearest neighbour for minority classes, balancing the dataset and improving classification accuracy. This technique increases the original dataset from 145 observations to 273 observations. Furthermore, the feature dimension reduction technique Principal Component Analysis (PCA), is used to reduce redundant features that are correlated from five features to three features by only keeping features with valuable information to determine clustering in the dataset. The dimensionality reduction is done to optimise the model and increase the generalisation speed.

The different conventional methods used to determine the transformer's condition which are: IEC, Dornenburg, Rogers and Duval Triangle are evaluated and compared. The results of the DGA methods were compared with the ANN classifier classification results and the actual condition of the transformer. A conclusion was made based on the comparison of whether ANN was efficient in making correct diagnosis. ANN performance was evaluated on accuracy, precision, recall and specificity.

1.4.3. Artificial Neural Network (ANN) Application

ANN consists of parallel connections with neurons resembling those of a human brain. This study uses the Multilayer Feed-Forward Back Propagation Neural Network which has a network architecture of three layers. A typical Multi-Layer Neural Network layout is depicted in Figure 1.

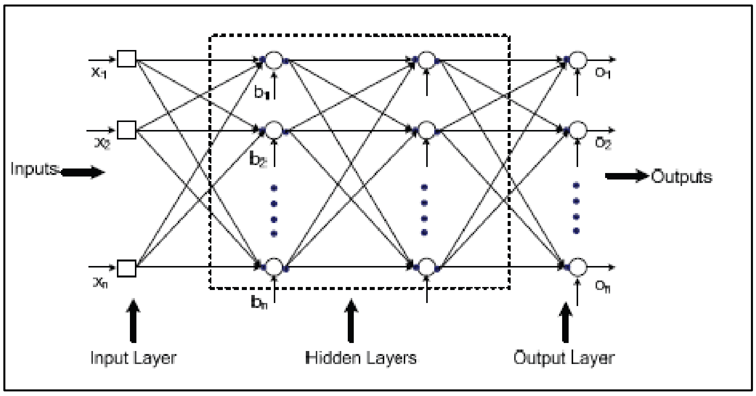

Figure 1: Multi-Layer Perceptron Neural Network [8]

a. **The input layer:** Consists of inputs with connecting weights, where input data with three transformed features will be supplied. The input signal is multiplied by connecting weights and added to the bias function before passing through a sigmoid function.
b. **The hidden layer:** With five hidden layers, the hidden layers were iteratively computed from 1 to 100 to obtain the number of hidden layers with the minimum error percentage. Each hidden layer consists of a sigmoid activation function where the computation takes place, a bias function to propagate the error between the computation output and the target output back.
c. **The output layer:** With five neurons where the model's output is shown.

The ANN classifier is initially trained with 70% of the dataset to learn and recognise the patterns, validated with 20% to strengthen its generalisation capability and tested with 10% to evaluate the performance. The error target is set to zero percent and the classifier is set to make 1000 iterations for training to find the optimal position where the model is not overfitted or underfitted. The model generalisation capability is evaluated for accuracy, recall, precision and specificity. Finally, a comparison is made between IEC, Dornenburg, Rogers, Duval Triangle and ANN with the actual condition of the transformer to determine the best performing technique.

1.5. Conclusion

This chapter outlined a justification for the current study. It described the challenges of conventional DGA methods for fault detection and classification and proposed a method to improve the accuracy. The research question and the approach that will be followed to answer the research question have been described. The following sections of this report will present literature and previous research work related to this topic.

2. Current Status of Power Transformer Condition Monitoring

This chapter covers the status and trends of tools and techniques used for power transformer condition monitoring. The DGA method detects most power transformer internal faults. However, other tests are necessary to detect faults in other components of the transformer such as the windings, tap changer, bushings, core and joints. These other techniques are used together with DGA or independently to provide a full spectrum of the health condition of a transformer. This chapter is structured by first giving an overview of the status of condition monitoring techniques in section 2.1. Section 2.2 covers the major components of the power transformer. Section 2.3 covers the failures experienced by the power transformer. Section 2.4 discusses conventional and modern diagnostic techniques and section 2.5 is the conclusion.

2.1. The Overview

Condition monitoring techniques for power transformers have advanced significantly over the years, with increasingly sophisticated methods being developed for continuous monitoring [9]. The demand for reliable and non-intrusive monitoring has driven the creation of modern diagnostic techniques that complement traditional methods such as insulation resistance, power frequency dissipation factor, and polarization index [9]. Both conventional and contemporary online detection techniques are guided by various international standards, including those from IEEE, IEC, and Cigre. The advent of smart grids has further transformed the energy sector, introducing more dependable online monitoring techniques beneficial for power transformer diagnostics.

The concept of Reliability Centered Maintenance (RCM), introduced by Nowlan and Heap [10] in the airline industry, revealed that 98% of aircraft failures were not time-dependent, further supporting the importance of condition monitoring. Their study emphasized the need to identify and prioritize failure modes, selecting appropriate and effective tasks to minimize the risk of failure.

Figure 2 illustrates the evolution of maintenance practices and their corresponding maturity levels, alongside their impact on the organisation. As maintenance regimes mature, they yield more significant business benefits, including enhanced safety, increased productivity, and reduced maintenance costs.

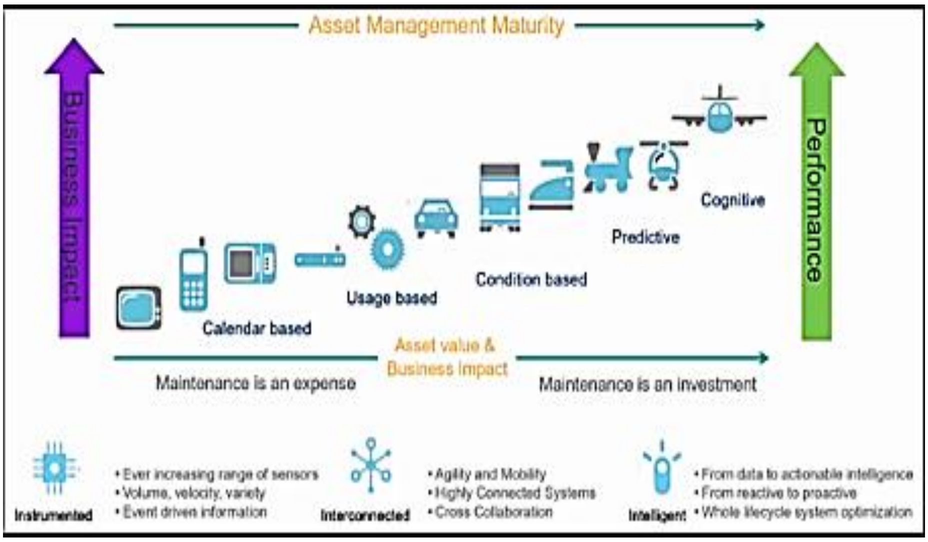

Figure 2: The Evolution of Maintenance [11]

2.2. Main Components of a Power Transformer

A power transformer primarily consists of its mechanical structure, insulation medium (oil and cellulose paper), and electrical and magnetic circuits. The mechanical components include the tank and radiators. The electrical components are the windings. The core represents the magnetic circuit. The insulation system comprises of the oil and paper, and the auxiliary components include the tap changer and Buchholz relay. Generally, the materials in a power transformer do not age significantly, except for the insulation system, which is therefore considered the most vulnerable part of the power transformer.

Figure 3 shows an example of a typical sub-transmission power transformer. To maximise the residual life and reliability of a power transformer, it is essential to understand the ageing processes and failure modes of its major components. The following section explores the key components of the power transformer and their susceptibility to failure.

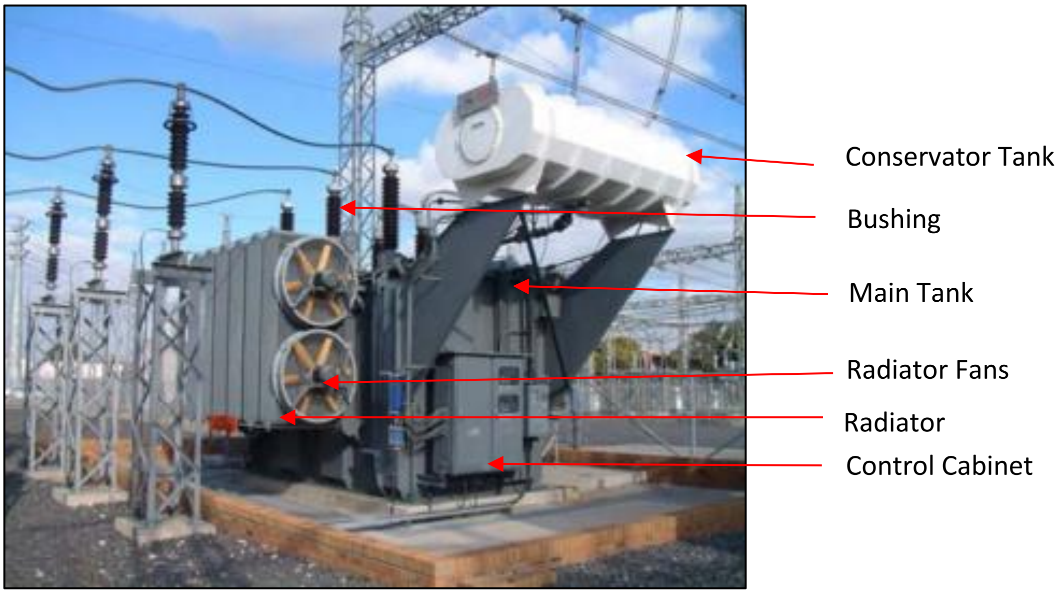

Figure 3: Power Transformer and its Major Components [12]

2.3. Ageing and Failure Analysis of the Major Components

2.3.1. Insulation Degradation

The power transformer consists of two types of insulating medium: solid (paper and pressboard) and liquid (insulating oil) insulation. In a properly designed power transformer, the insulation medium can perform reliably under normal voltage conditions and abnormal overvoltage conditions [12]. The leading cause of a power transformer failure is the overstressing of the insulation, where an electric field is applied to the material that is higher than the material's capability.

The lifespan of a power transformer is heavily weighted on the condition of the solid insulation (paper). The fundamental cause of paper ageing is heat; the cellulose Kraft paper has a degree of polymerization (DP) of around 1200. When the DP gets to about 200 to 250, it indicates that the paper has little remaining strength and the transformer is approaching the end of life [13].

Other factors that degrade insulation are moisture in the oil and the insulating paper and oxygen inside a transformer. When the transformer's paper insulation has deteriorated, the transformer is said to have reached the end of its economic life and needs to be replaced. While the liquid insulation can be replaced, filtered, or regenerated when degraded. The combination of moisture, heat and oxygen accelerates the deterioration of the transformer. These can get to a point where the insulation degrades itself over time. The insulating paper can act as a catalyst in producing acids in the oil which reacts with the paper, causing even further degradation and ultimately failure. DGA is a method used to evaluate the oil condition of the transformer through dissolved gases in transformer oil while Furan analysis is used for the paper insulation for power transformers. Sections 2.3.2 and 2.3.2 focus on the faults that the transformer components are vulnerable to that threaten the life of a transformer.

2.3.2. Effects of Through Faults

The mechanical structure of the power transformer can be affected by short circuits in the network due to forces associated with high current faults. These forces can lead to the deformation of the structure, resulting in additional complications. Network overvoltages, primarily caused by network switching, transient faults, and temporary overvoltage conditions, can contribute to power transformer insulation degradation.

2.3.3. Component Faults

Tap changers and bushings are critical components that can significantly contribute to transformer faults. Bushing failures are particularly severe, as they can lead to explosions, fires, and the destruction of the transformer unit, resulting in substantial losses to both the transformer and surrounding equipment. Additionally, bushing damage poses a risk to personnel, as exploding porcelain can eject its fragments into the air. Similarly with the auxiliary components of the power transformer, DGA is also used to evaluate their condition through oil analysis. The next section focuses on the different condition monitoring techniques used to assess the condition of different power transformer components.

2.4. Condition Monitoring Techniques

Most power transformer monitoring techniques primarily focus on the insulation system and can be categorised into online and offline methods. Online techniques are often favoured for monitoring transformers without interrupting production, while offline tests are typically considered more reliable [14]. The following subsections will explore various condition monitoring techniques used in power transformers. Although each technique is discussed separately, they are usually employed in combination to assess the overall condition of a transformer. Similarly, while DGA can detect the presence of faults, supplementary techniques are often required to pinpoint the exact location of the fault.

2.4.1. Traditional Fault Detection Techniques

Most traditional fault detection techniques are offline electrically based techniques and are intrusive in nature. They assess the electrical signals from electrical and mechanical changes in the resistivity, inductance, capacitance, moisture contamination or aging by-products in the insulation [9]. The disadvantage of these offline methods is that they do not account for the electrical and thermal insulation aspects present during the transformer's operating phase [14]. Due to the electrically based diagnostic technique's lack of moisture analysis in oil, it is recommended that they be combined with physiochemical techniques such as DGA for accurate oil diagnosis [9].

2.4.2. Modern Fault Detection Techniques

Modern diagnostic techniques are electrically based but non intrusive [15]. They are mainly sensor based techniques. They provide more detailed information about the condition of the transformer while in operation. This information is useful in diagnosing developing faults proactively and for trending purposes. Figure 4 below shows some of the modern and traditional techniques which will be discussed in the next subsections.

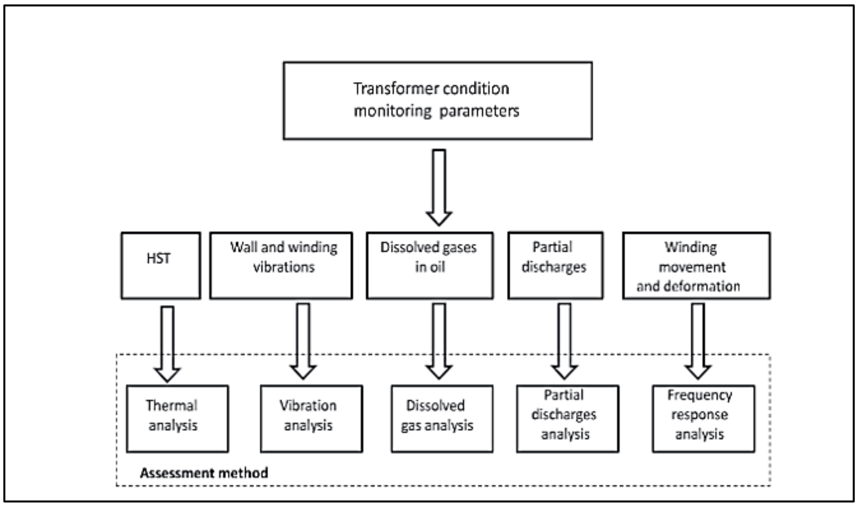

Figure 4: Condition Monitoring Method for Power Transformer [16]

2.4.2.1.Transformer Condition Assessment by Frequency Response Analysis (FRA)

Frequency Response Analysis (FRA) detects shifts and deformations in transformer mechanical structures and windings caused by mechanical stresses from short circuit fault currents or vibrations during transportation. Movement can damage the insulation, resulting in partial discharge or increased gas production due to insulation degradation. Winding movement can also affect the transformer's thermal performance by affecting cooling efficiency.

Mechanical faults, which often lead to prolonged outages, are common in older transformers with deteriorated mechanical structures [17]. This offline technique involves injecting a signal into one terminal of the transformer and measuring the response at another terminal, with results compared to previous measurements, data from similar transformers, or different phases of a three-phase transformer. FRA is valuable for assessing winding integrity, especially after transportation, and is guided by international standards [18-19]. It is praised for its sensitivity to winding faults and minimal dependence on historical data, though it requires a systematic interpretation approach. New advancements aim to automate FRA interpretation, and while Frequency Domain Spectroscopy is preferred for assessing insulation moisture levels [9], distinguishing between moisture and aging effects remains challenging.

2.4.2.2.Transformer Condition Assessment by Thermal Analysis

The most common thermal condition in a transformer is overloading and hot spot temperature or localised overheating. The infrared thermograph test is a popular electrical test for assessing the thermal condition of a power transformer. The test is carried out by a camera that captures non-visual infrared energy and converts the information into a pictorial image that is easier to interpret [20]. This method helps with fast hot spot analysis, confined heating, loose or bad networks, winding circulating currents, congested cooling, tap changer glitches, bushing and lightning arrester-related issues [21]. Temperature sensors are also used to measure the oil and winding temperature.

The initial indicator of transformer overloading is the generation of temperature-related gases such as CH_4, C_2H_6, and C_2H_4, with the specific gas indicating the severity of the rise in temperature. Dissolved Gas Analysis (DGA) can be performed to assess the thermal condition of the transformer oil. Additionally, infrared thermography can be used to examine the condition of the transformer, particularly at the bushing joints. Excessive heat can lead to insulation degradation and accelerated ageing of transformer components. Elevated temperatures can also affect vibration levels and increase the risk of partial discharge.

Research in this area are efforts to develop thermal models with improved accuracy for power transformers from Artificial Neural Network [22], Neuro-fuzzy [23] and Fuzzy logic [24]. These intelligent models [22][23] take load current, top oil temperature and ambient temperature as inputs to the Radial Basis Function Network (RBFN) and neuro-fuzzy, then output the maximum winding hot spot temperature. This fault diagnosis statistical method enables early detection of failures.

2.4.2.3.Transformer Condition Assessment by Vibration Analysis

Vibration analysis is a key method for evaluating the condition of load tap changers [15], winding and other internal components in power transformers. Accelerometer sensors are mounted to the wall of the transformer or tap changer to monitor vibrations and assess the integrity of the mechanical structures. This technique helps detect core and winding movement, loose or misaligned parts, the condition of tap changers and cooling fans, mechanical resonance issues, and overall structural integrity. Vibration analysis remains essential for diagnosing mechanical issues and ensuring reliable transformer operation. Vibration analysis and thermal analysis are complementary techniques used for monitoring power transformers. Mechanical issues identified through vibration analysis can lead to cooling inefficiencies, impacting thermal performance. Vibration can also cause physical movement or displacement of transformer windings affecting insulation and resulting in increased risk of partial discharge. Using both methods together provide a comprehensive view of the transformer's condition.

2.4.2.4.Transformer Condition Assessment by Partial Discharge (PD) Analysis

Partial Discharge (PD) is an electrical phenomenon caused by partial insulation breakdown or an uneven electric field. It is the most common failure mode in electrical insulation due to degradation over time [25]. For PD detection, chemical methods like Dissolved Gas Analysis (DGA) and High-Performance Liquid Chromatography (HPLC) are used. However, to locate faults, electrical and acoustic detection methods are employed. Ultra High Frequency (UHF) monitoring is particularly advantageous because it is less affected by external noise [26] around transformers or substations. PD is prevalent in power transformers [27] and can also occur in insulation bubbles. PD can be influenced by thermal conditions, winding movement, and mechanical stresses from vibration. Overheating and physical stress can accelerate insulation degradation, leading to increased PD activity. Monitoring PD helps identify insulation problems early.

2.4.2.5.Transformer Condition Assessment using Dissolved Gases Analysis

The oil sampling of electrical apparatus is conducted in accordance with ASTM D923 [28] and a chromatographic analyser is normally used to measure the concentrations of gases in PPM (Parts Per Million). The chromatography technique performs a physical-chemical analysis through a separation of chemical compounds [30]. DGA detects the dielectric breakdown of the oil molecules or cellulose molecules of the insulation due to developing faults. The technique's success is highly reliant on the data quality and reliability of DGA results. Gases released at different temperature levels are CO, CO_2, H_2, C_2H_6, C_2H_4, C_2H_2 and CH_4 [21]. Faults responsible for the released gases are: overheating of oil, partial discharge, or arcing inside a transformer. Classification of these faults is facilitated by using the IEC method, Key Gas Method, Rogers Ratio Method, CIGRE, Dornenburg ratio method, Duval Triangle [28,30] and Pentagon methods [21]. Figure 5 lists a few DGA methods. The abovementioned DGA techniques only provide information about the fault's existence and the type of fault. However, the location of the fault is not provided. The other monitoring techniques such as PD and thermal Imaging that are discussed in this chapter provide information about the location of the fault.

Intelligent machine learning techniques are an area that researchers have explored to improve the DGA fault classification. Researchers have explored ways to improve the accuracy and overcome the limitations introduced by DGA techniques. Fuzzy logic [24,30, 31, 32, 33,34], Artificial neural network [16, 23, 28, 35-37] and Support Vector Machine [38 - 40] are some of the popular algorithms researchers have used to improve DGA accuracy. The AI techniques have been used on DGA techniques, in combinations (hybrid approach) or individually to improve classification accuracy. These approaches are discussed in detail in chapter 4. Online sensors, spectroscopes and chromatography are key components in DGA based analysis [25].

Past studies have linked DGA in power transformer oil to developing faults [20]. Research interest has surged in methods for diagnosing potential faults hidden within power transformers [16]. DGA has offered the possibility of evaluating the condition of an asset and detecting developing faults earlier.

The DGA technique remains the most preferred as it does not require switching the power transformer out of service and furthermore, it is not affected by external factors such as the electric and magnetic fields. Elevated temperatures and partial discharge generate gases in the transformer oil. DGA can indicate problems like overheating, insulation degradation, or partial discharge. The presence of certain gases can correlate with specific issues such as thermal stress or winding movement in a power transformer.

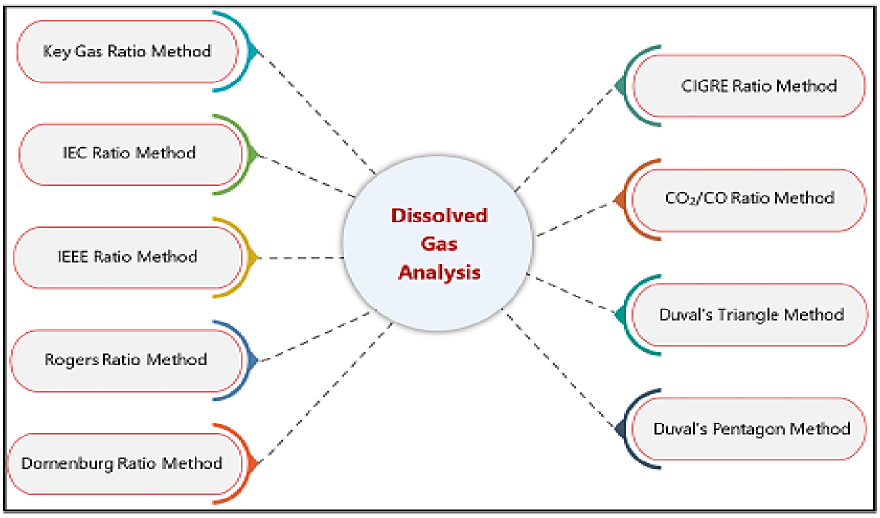

Figure 5: Different DGA Ratio and Graphical Methods [21]

2.4.2.6. Intelligent Monitoring Techniques

In recent years, the rapid development of sensor technologies, computer technologies, information technologies and other new technologies have brought new advances in condition monitoring and fault diagnosis of power equipment [5]. Intelligent monitoring techniques for power transformers leverage advanced technologies to enhance diagnostics, improve reliability, and optimize maintenance. These techniques have been introduced to deal with the deficiencies which include the reliability of signal collection from sensors, the accuracy of data treatment and analysis, the anti-interference performance of test equipment and the appropriate models used for condition evaluation. Traditional testing methods are being upgraded with sensor technology for real-time monitoring and independent analysis, utilizing advanced tools driven by AI and machine learning. Combined with machine learning algorithms they can predict and diagnose faults and be useful for trend analysis.

2.5. Conclusion

This chapter presented the status of the current methodologies in power transformer condition monitoring. Each diagnostic technique is unique in that it provides valuable information about transformer components. There is a strong link between the different techniques as they provide insight into potential issues affecting the transformer's health. There is a notable shift towards sensor technology and intelligent systems in condition monitoring involving combining the conventional and computational techniques to improve the reliability of analysis. Condition monitoring is a leading factor in fault prediction to minimise the risk of failure. With the correct diagnosis the asset owner can execute corrective action to realise a return on investment. The upcoming chapter focuses on Dissolved Gas Analysis (DGA) and its use in evaluating transformer oil. Various DGA techniques are explored to understand their strengths and limitations.

3. DGA as a Power Transformer Diagnostic Tool

This chapter explores the various DGA diagnostic techniques and their applications in power transformer condition diagnosis.

3.1. The Overview

Since the inception of DGA methodologies in the 1960s, various diagnostic techniques have been developed. The development of the DGA methods was initially based on observing the relationship between the actual condition of transformer insulation through dissolved gasses and the faults generated inside the transformer. Its success rate has been reported to depend heavily on the combination of expert knowledge and other relevant field information. This introduces subjectivity in the interpretation and is prone to errors as it can be interpreted according to a person's experience, knowledge and feeling of the machine.

The DGA testing process involves drawing an oil sample from the power transformer sampling point and testing it in the lab using chromatography. Chromatography is the most commonly used method of testing transformer oil samples. It measures and correlates the concentration of dissolved gases to the nature of a fault present. The seamless technique is not only reliable but also reduces the cost of maintenance by reducing the dependence on exhaustive statistical procedures [25] and is done while the transformer is in service.

Currently, the Rogers, Dornenburg, IEC, Duval and Key gases methods are the most widely used. Some of the other DGA conventional methods are illustrated in Figure 5 in section 2.4.2.5. Researchers who studied DGA ratio techniques from gas concentration have developed more comprehensive diagnostic techniques. Later, the graphical methods (Duval, Pentagon, etc.) were introduced as an improved version of the ratio methods to address the limitations of the ratio methods.

Figure 6 shows the relationship between different gas decomposition in oil insulation at various temperature levels. In Figure 6, H_2 is a gas that is present at a temperature of 150°C and above. There is another concept called stray gassing [41] which is present in temperatures below 150°C. This type of gassing creates a false impression of a developing fault, which can sometimes be confused with a real fault. Other gases are released as the temperature rises; C_2H_2 is associated with arcing inside a transformer.

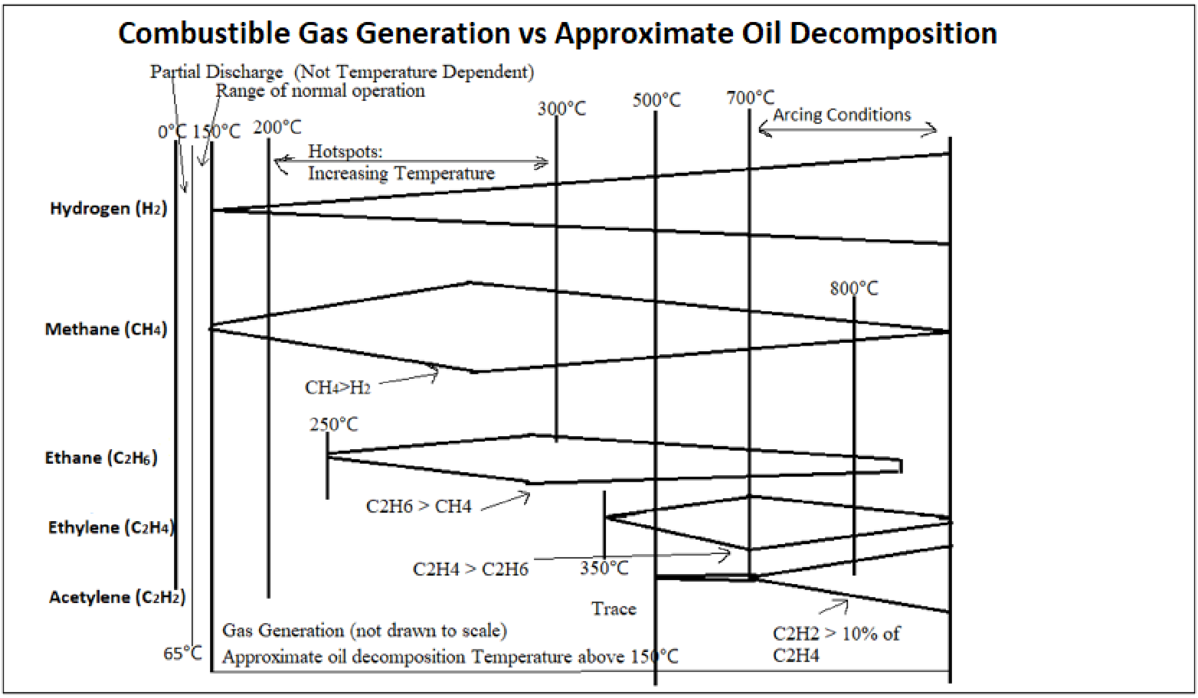

Figure 6: Oil Decomposition Based on Temperature [32]

3.2. The Duval Triangle Method of DGA Interpretation

The Duval Triangle and Pentagon methods are the most commonly used graphical methods. IEEE C57-104 [25] reports on the comprehensive applications and benefits of various graphical methods (Duval Triangle 1, 4 and 5 as well as Duval Pentagon 1 & 2). The Duval Triangle detects six different types of fault gases, which are categorised as electrical, thermal and a combination of both. Among the conventional DGA methods, the Duval Triangle has proven to be the most accurate [25][39]. Its application on samples with low gas concentration is discouraged due to its nature of always identifying a fault. The Duval Triangle method was developed to overcome the undefined ranges of ratio methods [31]. It has been reported to perform better in classifying faults with a high percentage of correct predictions with a high success rate [7].

Its failure to correctly detect and classify PD and thermal faults has been linked to its disregard for H2 and C2H6 [42,43]. Some interference between thermal and electrical faults has been reported in the area of the electrical and thermal fault, on the DT zone [42], see Figure 8. A case study in a Cigre study of 2019 [31] revealed that the Duval and Pentagon methods outperformed other conventional methods in fault classification. Figure 7 shows the Duval Triangle 1 method [25] that will also be applied in the present study.

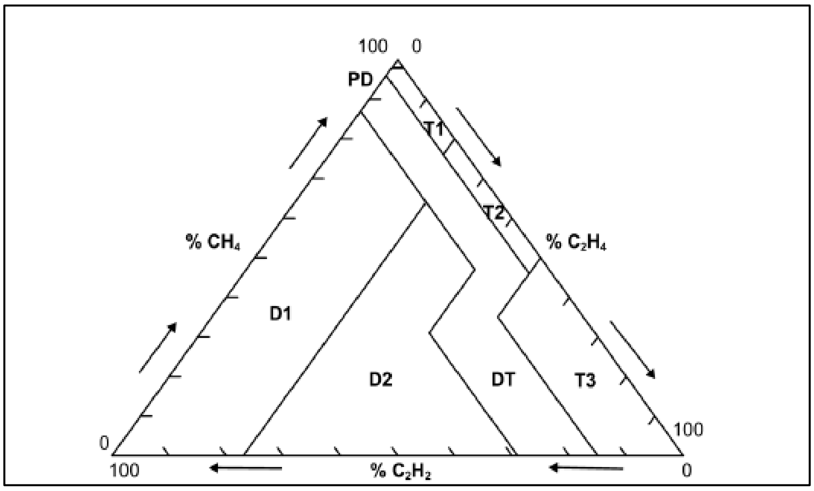

Figure 7: Duval Triangle 1 Method [4]

3.3. Ratio Method of DGA Interpretation

The ratio methods are unique in that they can account for the volume of oil in the transformer by analysing the ratio of gas pairs rather than absolute values [31]. They are, however, limited in their ability to diagnose cellulose insulation and do not handle a transition [44] of gas level in boundary cases very well. The most significant limitation of all ratio methods is their inability to identify faults for some ranges, even when a fault is obvious. A comprehensive list of the various ratio methods is shown in Figure 5 in section 2.4.2.5.

3.3.1. Dornenburg Ratio Method of DGA Interpretation

The ratio techniques were first proposed by Dornenburg [21] and modified by Rogers before they were revised in IEC standard 60599 [43]. Dornenburg uses four ratios namely: CH_4/H_2, C_2H_2/C_2H_4, C_2H_2/CH_4 and C_2H_6/C_2H_2 to detect six faults as per the Table 1 below. The Dornenburg ratio principle is considered valid if the gas concentrations of H_2, CH_2, C_2H_2 and C_2H_4 exceed twice the fixed limit of each gas and C_2H_6 exceeds three times the fixed limit [45].

Table 1: Dornenburg Ratio Limits [25]

Gas Ratio	Value	Code
$R_1 = CH_4/H_2$	$R_1<0.1$	0
	$0.1 \le R_1 \le 1$	1
	$R_1>1$	2
$R_2 = C_2H_2/C_2H_4$	$R_2<0.75$	0
	$R_2>0.75$	1
$R_3 = C_2H_2/CH_4$	$R_3<0.3$	0
	$R_3>0.3$	1
$R_4 = C_2H_6/C_2H_2$	$R_4<0.4$	0
	$R_4>0.4$	1
Normal Operation		3
The approach could not identify the fault		4
R_2 is not involved		5

Table 2: Dornenburg Fault Code [28]

No	Fault Type	Ratio Code				Output Fault Code
		R1	R2	R3	R4	
1	NF	3	3	3	3	[1 0 0 0 0]T
2	UD	4	4	4	4	[0 1 0 0 0]T
3	TH	2	0	0	1	[0 0 1 0 0]T
4	PD	0	5	0	1	[0 0 0 1 0]T
5	AR	1	1	1	0	[0 0 0 0 1]T

3.3.2. Rogers Ratio Method of DGA Interpretation

Roger's ratio is the most common ratio method [28] of DGA interpretation. It employs four ratios—C2H2/C2H4, CH4/H2, C2H6/CH4, and C2H4/C2H6—to identify four types of fault conditions: Partial Discharge (with or without tracking), as well as electrical and thermal faults of varying severity levels. The Rogers method, like the Dornenburg and IEC ratio methods suffers from the out of range problem.

Table 3: Rogers Ratio Limits [25]

Gas Ratio	**Value**	**Code**
$R_1 = CH_4/H_2$	$R_1<0.1$	2
	$0.1\leq R_1\leq 1$	0
	$R_1>1$	1
$R_2 = C_2H_2/C_2H_4$	$R_2<0.1$	0
	$0.1\leq R_2<1$	1
	$1\leq R_2\leq 3$	2
$R_3 = C_2H_4/C2H_6$	$R3<1$	0
	$1\leq R3\leq 3$	1
	$R3>3$	2

Table 4: Roger's Fault Code [28]

[R1 R2 R3 R4]	**Fault**
[0000]	Normal deterioration
[5000]	Partial Discharge
[1000], [2000]	Slight Overheating, T < 150°C
[1100], [2100]	Overheating, 150°C < T < 200°C
[0100]	Overheating, 200°C < T < 300°C
[0010]	General Conductor Overheating
[1010]	Winding Circulating Currents
[1020]	Core and Tank Circulating Currents, Overheated Joints
[0001]	Flashover without Power follow Though
[0011], [0012], [0021]	Arc with Power follow Though
[0022]	Continuous Sparking to Floating Potential
[5001], [5002]	Partial Discharge with Tracking

3.3.3. IEC Ratio of DGA Interpretation

The IEC ratio method uses only three ratios C_2H_2 / C_2H_4, CH_4 / H_2 and C_2H_4 / C_2H_6 to identify six main faults [5]. Like the other ratio techniques, the IEC method also suffers from the outside the range limits diagnosis. This is often attributed to a mixture of faults or a new fault that combines with a high background gas level [5]. IEC 60599 – 2019 the revised version makes provision for the out of range diagnosis by relating a diagnosis to the closest case through visualisation in Figure 8. The IEC method further provides a simplified interpretation for those cases with no fault diagnosis. There also exists an overlap between D1 and D2 which IEC attributes to the conflict in diagnosing D1/D2 faults.

Table 5: IEC Ratio Limits [6, 28]

Gas Ratio	**Value**	**Code**
$R_1 = CH_4/H_2$	$R_1<0.1$	1
	$0.1\le R_1\le 1$	0
	$R_1>1$	2
$R_2 = C_2H_2/C_2H_4$	$R_2<0.1$	0
	$0.1\le R_2<3$	1
	$R_2>3$	2
$R_3 = C_2H_4/C_2H_6$	$R_3<1$	0
	$1\le R_3\le 3$	1
	$R_3>3$	2

Table 6: IEC Fault Code [28]

[R1 R3 R4]	**Fault**
[000]	Normal deterioration
[100]	PD with low energy density
[101]	PD with high energy density
[011], [012], [022]	Discharge (arc) with low energy
[021]	Discharge (arc) with high energy
[010]	Thermal fault of temperatures < 150℃
[200]	Thermal fault of temperatures 150℃ < T < 300℃
[210]	Thermal fault of temperatures 300℃ < T < 700℃
[220]	Thermal fault of temperatures T > 700℃

Table 7 below summarizes the fault types for each DGA technique. The corona effect on the structure of paper produces H_2 gas. CH_4 and C_2H_6 are linked to low-temperature faults. Due to medium-temperature faults inside the oil-paper insulation, H_2 and C_2H_4 are produced. C_2H_2 is associated with high-temperature faults and heavily discharged arcs.

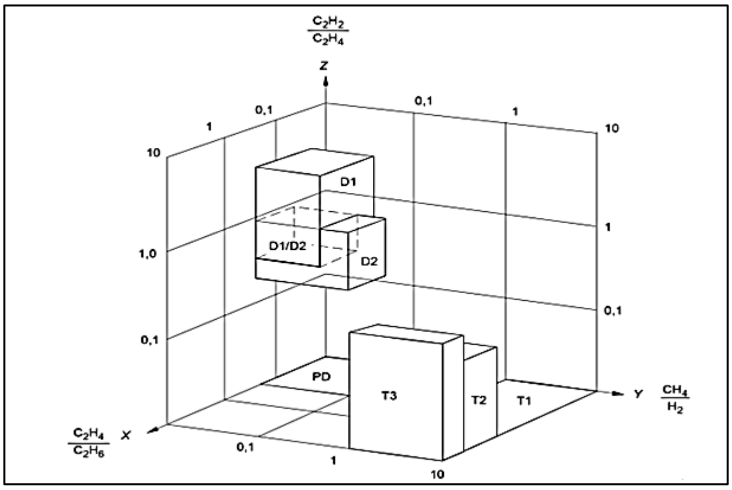

Figure 8: IEC Ratio Boundary Limits [3]

Table 7: DGA Ratio Methods with Fault Types [21]

Method	F1 Thermal fault (cellulose)	F2 Thermal fault (oil)	F3 Electrical fault (Corona)	F4 Electrical fault (Arcing)
Rogers ratio	Slight overheating <150°C Overheating 150-200°C Overheating 200-300°C	Conductor overheating Winding circulating current Core/tank	Low energy electrical discharge PDs PDs with tracking	High energy electrical discharge Flashover Arcing Continous sparking
IEC ratio	Thermal fault <150°C Thermal fault 150-300°C	Thermal fault 300-700°C Thermal fault >700°C	Low energy electrical discharge PD of low energy density	High energy electrical discharge Discharge of low energy
Dornenburg ratio	Thermal decomposition with	Thermal	Corona	Arcing
Duval triangle	Thermal fault <300°C	Thermal fault 300-700°C	PDs Mix thermal and electrical faults	Lower energy discharge High energy

3.4. Conclusion

There is a trend towards the adoption of computational techniques. DGA must first be assessed for normal conditions before the diagnostic techniques are used. This is to avoid misclassifications for instances of smaller gas concentrations that are normal operating conditions. While the different DGA methods discussed in this section have their own merits, none can classify a full spectrum of faults in a transformer for a wide range of DGA data. With the key gas method, a history is needed to interpret the fault consistently. The ratio techniques have limitations with undefined ranges. Duval Triangle is unable to classify cellulose and PD in faults.

Secondly, the ratio method's undefined areas create a challenge for diagnosing certain faults that fall outside the boundary. Researchers have often associated this with multiple faults. The transitioning of the boundary code is another challenge specifically being reported with the IEC method where an undefined value is returned for a transition between different faults.

Lastly, the conventional DGA methods do not use all the key gases that are generated in oil according to the relationship between oil temperature and gas concentration as in Figure 8, which will always raise doubts about fault classification accuracy as some gases which are linked to certain faults are not represented by a ratio combination.

4. Review of Machine Learning Techniques on DGA

This section reviews artificial intelligence techniques for use in DGA for incipient fault diagnosis. Although there are numerous machine learning algorithms covered in literature, this study only focuses on the ones that stand out for DGA applications to provide a rationale for the current choice for this study.

4.1. The Background

The high number of datasets available has motivated the use of AI to automate decision making and reduce human intervention. AI allows machines to think like humans and make decisions on certain tasks that human decisions would require. The research conducted broadly related to DGA- has been efforts by researchers to mainly address the challenges with the DGA classification accuracy, with the focus being on the unresolved cases and incorrect diagnosis [31].

Various ML algorithms, including Artificial Neural Network [28,35], Fuzzy Logic [30,31,32,46], Support Vector Machine (SVM) [39], KNN (K Nearest Neighbour) and others, that have been developed have outperformed the conventional DGA techniques. Depending on the nature of the problem, researchers have found ways to improve the models by using an individual, a combination of different learning algorithms or a mixture of DGA combined with algorithms in their investigations.

4.2. Machine Learning Tools

4.2.1. Artificial Neural Network

ANN is one of the non-parametric classifiers reported to perform best with new dataset compared to linear classifiers [38]. In [25], five different classifiers are tested for DGA fault classification on the Duval Pentagon 1 method: Auto Weka, ML, NB (Naive Bayes), LibSVM (Library for Support Vector Machines) and KNN. What stands out with all these models is that while their performance is good, no one model is perfect for all situations. Some models will be sensitive to certain faults and classify them correctly while they perform badly on others. The author [28] used a Smart Fault Diagnostic Approach (SFDA) to improve the accuracy of DGA classification. This was done by taking the most accurate output of the 5 DGA methods and feeding them into a decision-making multilayer feed-forward ANN with the backpropagation ANN technique. The overall model improved the fault classification by overcoming the shortcomings of each conventional DGA method used in the study, therefore yielding the best performance.

ANFIS (adaptive neuro-fuzzy inference system) is an ANN-Fuzzy logic hybrid technique introduced to supplement ANN with Fuzzy Logic. It makes decisions using a combination of ANN and FIS (Fuzzy Inference System). It applies the rule base used in Fuzzy systems and the superior learning capability of ANN. The authors in [30,32] investigated the unique intelligence of ANN with the combination of FIS's if-then rules to improve the model's accuracy and to solve multiple fault diagnosis that the most conventional technique is unable to diagnose.

In [16], the author dealt with the ANN's large dataset and reduction in training duration by introducing a rough set fuzzy wavelet ANN. The rough set fuzzy wavelet ANN model was used to simplify the dimension of the input data and improve generalisation capability and speed. The work further suggested ensemble learning methods for better fault prediction. The study highlighted the importance of data preprocessing and data balancing methods on DGA data. Although the study achieved good performance, the author highlighted the lack of individual class label performance in the low thermal class. This is attributable to the closeness between the high and low thermal classes.

Neural Network has been reported by most researchers to require larger dataset for better generalisation ability. In [36], the author trained and tested a CNN (Convolutional Neural Network) on a sample of 589 data samples. Despite claims in the literature that a larger training dataset of thousands or more is required to make an accurate generalization, this study produced good prediction accuracy. The present study takes a similar approach with an initial number of samples of 145. To leverage the diagnostic ability of ANN with the minimum dataset, an Adaptive Synthesiser is introduced to increase the dataset to 273 in order to handle the imbalance in the dataset and increase the dataset for ANN to perform efficiently.

Multiple faults in the transformer and how it affects DGA is another area that has greatly been researched. Researchers explored ANN and the combination of both to diagnose these faults that most DGA method fail to diagnose. Duval Triangle does however make provision for classifying multiple faults in the DT zone of the triangle. The other conventional methods only detect a dominant fault as the classified fault.

The majority of the research related to DGA fault diagnosis is focused on undetermined cases and incorrect diagnosis by applying soft computing techniques [31]. The present study further aims to build upon existing knowledge around enhancing the accuracy of the ANN classifier in detecting undefined areas that the IEC ratio method is unable to detect. This will be achieved by training, optimising and testing with the new dataset to determine the accuracy and to answer the research question. The output of the optimised ANN will be compared with the actual inspected condition of the machine as well as the results of the literature.

4.2.2. Fuzzy Logic Application on DGA

Fuzzy logic (FL) is one of the powerful classifiers that has been explored and has shown good performance in classifying faults. It uses the membership function to replace a single value output with a range to enable the degree of belongingness to a particular fault. FL has been applied to resolve the uncertainty at the boundaries where multiple faults are present. It establishes the dominant fault and the less dominant [30,32,44] as well as the severity of each fault [25]. It has been used satisfactorily in cases where values fall on the threshold or the boundary for ratio methods, in cases of multiple faults where different faults mislead resulting in ratio confusion and in cases where the DGA code is undefined. Fuzzy logic (FL) uses a simplified linguistic approach to make decisions, making them more transparent, explainable and desirable [33]. The model's inherent nature of operation is based on the degree of belongingness to a fuzzy set rather than a binary approach, which makes it more adaptable to dealing with emerging faults.

In the literature, researchers used FL [31] to Dornenburg, key gas [30,34], IEC, Rogers and Duval [30] to successfully diagnose incipient faults. A fuzzy inference system [30,32,44] has been used extensively to deal with the ratio method's unclassified ranges, which were linked to multiple fault diagnoses. FIS was used with the IEC ratio (Fuzzy - IEC) in [32] to diagnose incipient faults in a transformer and to further quantify the severity of each fault. FIS has been used more often with the ratio methods to deal with the unclassified codes that the ratios introduce. The authors in [30,32,47] linked the issue of undefined ranges in ratio methods to the existence of multiple faults. This was successfully implemented and later validated by the next DGA sample result. Fuzzy logic has also been used successfully to handle boundary faults. The boundary condition is said to tend to change the codes. Fuzzy logic handles this by incorporating an overlap in its membership function to ensure a smooth transition between faults.

The author in [47] accommodated a 10% tolerance on concentration measurement from the gas analyser with the boundary of the Fuzzy Inference System for each ratio overlapping between two consecutive codes.

The concept of extension methods was introduced to extend the ratio ranges where generalisation normally yields no output. This method employs a fuzzy approach to extend the original data into compatible and solvable domains through the use of a membership function. The extension method was used for the first time in power transformer diagnosis to extend the Fuzzy values from [0,1] to $(-\infty,\infty)$ [46]. The author employs a diagnosis method based on a matter element model and an extension set, in which the degree of extended relational functions directly identifies the incipient faults. This too, produced promising fault classification results [46, 48]. Another author [48] used Fuzzy's extended methods to solve contradictions and incompatibility problems in ratio methods.

The main disadvantage of fuzzy has been reported to be its difficulty in fixing the membership function and fuzzy parameters and its construction requires many rule guides, expert experience [49]. The construction of the membership function is said to be on a trial-and-error basis and it is too time consuming and ineffective.

4.2.3. Support Vector Machine (SVM) and Other ML Classifiers Applications on DGA

According to [39], Support Vector Machines (SVM) demonstrate exceptional classification capability and accuracy, particularly with smaller sample sizes. In that study, power transformer fault classification was successfully carried out using a hybrid feature selection method combined with a Genetic Algorithm-SVM classifier (GA-SVM). Additionally, the Adaptive Synthetic (ADASYN) technique and the arctangent transformation method were employed to enhance the statistical properties of the training data. This study adopts a similar approach with ADASYN technique to tackle dataset imbalance in Dissolved Gas Analysis (DGA). SVM has shown impressive generalization performance, proving its ability to accurately classify and predict new samples. It excels not only with smaller sample sizes but also in handling high-dimensional and linear problems.

Ensemble learning is a concept that arose from meta machine learning, where researchers used multiple learners to solve a problem and combine the decisions. It combines the strengths and weaknesses of multiple classifiers to create the best classifier possible. The models are frequently used for decision-making in a variety of disciplines, including finance, biomedical engineering and power engineering [50]. By avoiding the selection of a single weak classifier, ensemble learning increases the likelihood of selecting more accurate classifiers [52].

Wavelet networks are also used in fault diagnosis. This algorithm is similar to ANN, except that activation functions play a larger role in determining diagnosis accuracy than neural networks, where network structure (number of neurons and hidden layers) determines model accuracy [30].

4.3. Comparison Across Various ML Techniques as Applied to DGA

In general, non linear classifier was shown to be superior to linear classifiers [25]. ANN's knowledge is embedded in the network and is difficult to interpret because it is encoded in a mathematical format. There are some considerations when applying ANN algorithms: it requires historical data to train a classifier before making optimal decisions. In addition, its accuracy is highly dependent on the availability and authenticity of DGA data [25]. ANN's training time is relatively longer or slower depending on the large amount and high dimensionality of data supplied.

The robustness and fault tolerance of neural networks, the network's flexibility in automatically adjusting to a new environment, its ability to deal with a variety of data situations such as noisy and inconsistent data [51], probabilistic and fuzzy data and its ability to perform collective computation are some appealing features that distinguish it from other machine learning algorithms. This is also the primary reason for the choice of ANN in the current study.

In comparison, the fuzzy inference process is more transparent and explainable. However, the associated membership functions and diagnostic rules are given based on experiences or trial-and-error tests, which can be a time-consuming activity. Fuzzy logic has been reported to have a high dependency on the accuracy of the data involved [25]. It is also found to lack the ability to adapt to new fault cases, which means a classifier will require an update whenever there's a new fault condition that it does not recognise. On the other hand, ANN provides approximately correct output for unforeseen data for which the model is not trained for [25].

4.4. Conclusion

In this chapter, the capabilities of some machine learning methods have been discussed along with their advantages and disadvantages. While these intelligent techniques have been applied and achieved better results than the conventional DGA methods, each has its own limitations. It is evident that some algorithms can perform better in some situations and be bad in other situations. It is also a reason there is currently no one model in the industry that is fitted for all purposes. Further research is needed in this area to explore whether a standardised model can be developed that can perform better in all instances of classification with good accuracy. In the following chapter, the proposed ANN classifier will be simulated using DGA samples.

5. Experimentation with ANN on Improving DGA Accuracy: A Case Study.

The previous chapter discussed the advantages and disadvantages of popular machine learning algorithms that are commonly applied to DGA. This chapter evaluates the proposed ANN on the transformer oil samples and discusses the outcomes of the model. In any machine learning study, the authenticity of the data used is critical. In the present study effort was directed at obtaining and preconditioning the DGA data as presented in the next sections.

5.1. Data Collection

This study uses DGA data collected from the IEC TC 10 database and other data from credible sources in the literature. A total of 145 samples from transformers with varying voltage ratios and MVA (Mega Volt Amperes) are first evaluated using four DGA methods which are: Rogers, Duval, Dornenburg and IEC ratio. The dataset consists of five gas concentrations comprising of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 that will be used as input to the ANN algorithm. The gas concentrations are represented in parts per million (ppm). The raw data is presented in the format shown in Table 8. Appendix A contains the rest of the 145 samples of the dataset. In the next section, this dataset will first be evaluated on each of the four conventional DGA methods to determine the results of each individual method.

Table 8: Dataset of Dissolved Gasses Obtained from IEC TC 10 and the Literature.

Sample Name	**H_2**	**CH_4**	**C_2H_6**	**C_2H_4**	**C_2H_2**
Sample 1	230	16	16	2	2
Sample 2	112	19	104	6	4
Sample 3	124	14	4	0	13
Sample 4	102	6	6	7	10
Sample 5	99	170	20	200	190
Sample 6	310	230	54	610	760
Sample 7	24	109	69	0	0
Sample 8	10	26	147	6	0
Sample 9	89	461	95	185	18

5.2. Evaluating the Conventional DGA Criteria

In this section, a total of 27 out of 145 randomly selected DGA samples are evaluated using Dornenburg, Rogers, IEC and Duval Triangle methods. The results of the conventional DGA methods are compared against the actual condition of each sample in Table 9. The green highlight indicates where the conventional method was successful in making a correct prediction, while the red indicates an incorrect or no prediction was made. The Duval Triangle method shows better accuracy compared to the other conventional methods. However, as noted in the literature, it always provides a fault prediction for samples in normal conditions or for samples with low gas concentrations. Among the conventional methods, IEC performed poorly with most of its output falling out of range.

Table 9: Comparison of Four DGA Methods with the Actual Sample Condition

Sample No	H2	CH4	C2H6	C2H4	C2H3	IEC faultcode	Dorn Faultcode	Duval Faultcode	Rogers Faultcode	Actual Fault
Sample 15	227	759	156	3133	191	NA	TH	TH	TH	TH
Sample 23	170	320	53	520	3	NA	TH	TH	TH	TH
Sample 32	650	53	34	20	0	PD	PD	TH	PD	PD
Sample 34	300	490	180	360	95	NA	TH	TH	NA	TH
Sample 69	117	167	48	481	7	NA	TH	TH	TH	TL
Sample 76	307	22	2	33	109	NA	NA	D	NA	D
Sample 77	60	144	67	449	9	NA	TH	TH	TH	TL
Sample 80	208	1774	408	1431	1	NA	TH	TH	TH	TL
Sample 87	75	15	7	14	26	D	D	D	NA	D
Sample 90	1230	163	27	233	692	D	D	D	D	D
Sample 98	755	229	32	404	460	D	D	D	D	D
Sample 12	105	125	71	166	10	NA	TH	TH	TH	N
Sample 13	80	0	200	100	4	PD	NA	TH	PD	N
Sample 13	125	100	100	150	20	D	NA	TH	NA	N
Sample 13	50	30	0	0	5	D	NA	D	NA	N
Sample 14	144	13	36	3	2	NA	PD	D	NA	PD
Sample 15	503	40	25	14	0	PD	PD	TH	PD	PD
Sample 20	246	596	322	872	3	NA	TH	TH	TH	TL
Sample 20	1245	3865	721	2449	12	NA	TH	TH	TH	TL
Sample 20	35	78	29	210	5	NA	TH	TH	TH	TL
Sample 22	546	1367	228	1196	14	NA	TH	TH	TH	TH
Sample 22	334	872	221	1481	37	NA	TH	TH	TH	TH
Sample 23	70	235	96	369	7	NA	TH	TH	TH	TH
Sample 23	730	1588	431	3972	74	NA	TH	TH	TH	TH
Sample 23	360	2933	23292	21341	94	TL	TH	TH	NA	TH
Sample 24	52	31	3	2	5	TL	NA	D	NA	N
Sample 27	107	91	79	157	36	D	NA	TH	NA	N

Table 10 shows the overall performance of the four DGA methods per fault type on the 27 samples that were randomly selected for testing. Out of the four methods, Dornenburg shows superiority over the other three methods with an accuracy of 55%. The IEC method is the inferior method amongst all the four methods with an accuracy of 32% for the given test dataset. This is attributed to the smaller area of classification with a larger area being undefined thereby returning an undefined value for most of its predictions. This is noticeable in Figure 8 in section 3.3.3. As already mentioned in the literature, the successful implementation of the conventional method will require other information related to the power transformer. Meaning that the maintenance or performance history will need to be assessed in conjunction with the conventional DGA method.

Table 10: Accuracy Measure for all Criteria

Fault	number of samples	IEC	Dorn	Duval	Rogers
PD	3	67%	100%	0%	67%
D	4	75%	75%	100%	50%
TL	6	17%	0%	0%	0%
TH	8	0%	100%	100%	75%
N	6	0%	0%	0%	0%
Average		32%	55%	40%	38%

5.3. Data Pre-Processing

Data pre-procession is done to clean up the raw data. The process involves but is not limited to removing missing and errors on the data as well as data normalising, scaling, or transforming. The next sections discuss the data processing methods that are relevant to the dataset used for this study.

5.3.1. Rescaling

The log transform is used to rescale the highly skewed distribution of samples. The skewness in the distribution is introduced by a mixture of very low gas concentrations of a few PPM and samples with very high gas concentrations. The skewness in the dataset normally distorts the visualisation of the data points. The log transform for transforming the data is computed by Equation 1 below:

$$X_{log} = log_{10}\ (1+X) \qquad (1)$$

Where X is each individual gas concentration in ppm.

An assumption is made that DGA criteria will not discriminate between a zero gas content and a negligible quantity of one ppm and thus all zero ppm and below concentration are kept at one ppm to avoid dealing with negative log values and errors. For example, a gas concentration of H_2 of zero ppm will be transformed as $X_{H2}= \log_{10} (1 + (0\text{ ppm})) = 0$

5.3.2. Normalising the Dataset

Normalising is done to rescale the original data range to achieve a significant separation margin between big and small values in the data. It is applied to the dataset to avoid attributes in a larger numeric range dominating those in smaller numbers. The scaling process used in this study is the min – max normalisation shown with the Equation 2 below.

$$Xnorm = \frac{X - Xmin}{Xmax - Xmin} \tag{2}$$

Where ***Xnorm*** is the normalised value,

X is the original value,

Xmin is the minimum value of the feature,

Xmax is the maximum value of the feature,

5.3.3. Data Balancing

Data balance is a technique applied to deal with imbalances in a dataset caused by the unequal number of observations per class label or fault type. Generally, a power transformer failure is a rare event in a power system. As a result, there are few data points acquired during failure compared to data points acquired in normal operating mode. This causes an imbalance in the dataset. A dataset with only 5% of faulty transformers, will result in a classifier with an accuracy of 95% if it ignores its inputs and classifies all data points as normal. Some of the common techniques used for data balancing are Bootstraps, Synthetic minority oversampling technique (SMOTE) and adaptive synthetic sampling algorithm (ADASYN) [33]. While the data balancing technique presents a good-performing model, it is reported to present biases and non-realistic accuracy on the classifier as the model is not built on the actual machine condition. For the purpose of this study, data balancing is used to balance the minority class outputs as well as to increase the dataset for training the classier at a later stage. The techniques to handle imbalanced learning problems are categorised below:

Manual entry: This method involves manually duplicating the minority class to equate to the majority class. This is called oversampling while undersampling involve deleting the dataset to maintain the number of majority to the minority class. The disadvantage of this method is that it’s tedious and for undersampling, some important samples and features for classifier accuracy might be missed. This method is disregarded in this study due to the compromised accuracy as well as the time factor.

SMOTE: This next method generates an arbitrary number of synthetic minority examples to shift the classifier learning bias towards the minority class. The technique works better with a linear dataset, but for a non-linear dataset it introduces a lot of errors. Due to the non linearity of the dataset in this present study, this method has been discarded.

Adaptive Synthesizer (ADASYN): This method aims to overcome the imbalance in the original dataset by artificially generating data samples for the minority data samples. This is achieved by creating data points in the vicinity of the boundary between two classes using the majority vote from the K-Nearest Neighbour approach. The input of the technique includes specifying the number of K for density estimation. ADASYN is the preferred technique for the present study. The outcome of ADASYN is the increased number of observations or samples from 145 to 273 samples. The next section covers the feature selection which is a data pre processing technique applied to reduce the dimensionality in the dataset. This is achieved by selecting the features that carry more information about the dataset.

5.3.4. Feature Selection

In this section, feature selection is used to identify a smaller set of highly relevant variables from the original dataset. This process involves eliminating irrelevant, redundant and noisy variables to achieve a simpler, more effective and accurate generalization. Research has shown that combining multiple feature selection methods can yield more robust and reliable results [34]. This study employs Principal Component Analysis (PCA) to identify suitable features for the ANN classifier. PCA is a statistical technique that transforms a set of correlated variables into a set of uncorrelated variables through orthogonal transformation. The advantages of using PCA here are that it prevents overfitting, reduces computation time and improves model accuracy.

The original five features, represented as gas concentrations, are reduced to four principal components as illustrated in Figure 9. These first three new features, derived from PCA, capture the most variance in the dataset. The first component explains approximately 59% of the variance, indicating it contains 59% of the dataset's information. The first two components together account for about 79% of the total variance. In total, the first three components cover 90% of the dataset's variance. The fourth component, which explains only 5% of the variance, will be discarded to enhance the overall performance of the classifier.

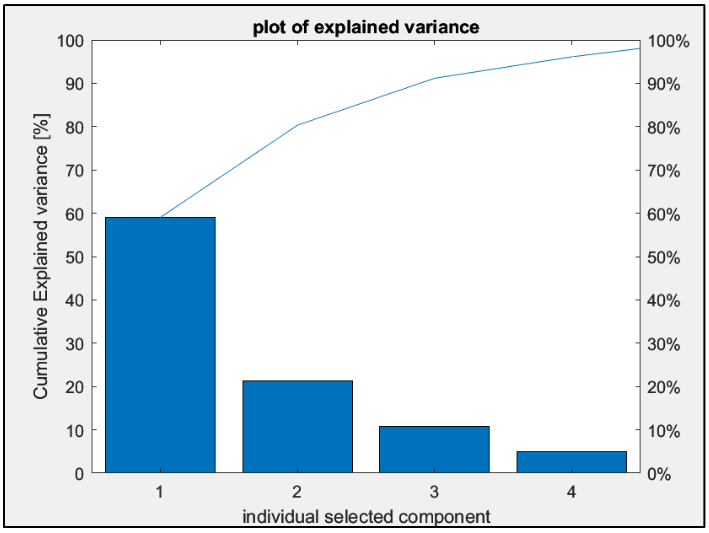

Figure 9: Plot Showing Variance in Data

Figure 10 below is an outcome of PCA transformed dataset. In Figure 10, PCA is used for data visualisation. The reduced number of variables from five to three dimensions makes it simpler to visualise and interpret the dataset. In addition, there is a noticeable separation in the class outputs. In the next section, ANN classifier is used to process the transformed dataset with 273 observations and three transformed features.

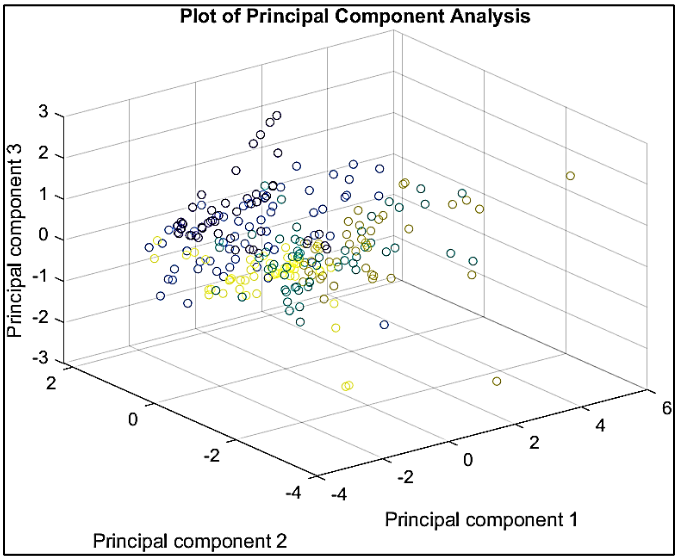

Figure 10: Principal Component Analysis Applied for Data Visualisation.

5.4. Network Architecture

The ANN used in this study is trained by a Back Propagation (BP) – Bayesian Regularizer (BR) algorithm with a sigmoid function. The network architecture is a three-layer system comprising of the input layer, the hidden layer and the output layer. The back propagation algorithm calculates the error iteratively and updates the weights backward through the connections to the input layer until the desired target output is achieved. The whole process involves adjusting the synaptic connections that exist between the neurons until the desired outcome is achieved. Figure 11 shows the constructed ANN that is used for this study. It consists of three input neurons at the input layer, eight hidden layers and five output neurons at the output layer. Appendix B contains the Matlab™ code for simulating the ANN model.

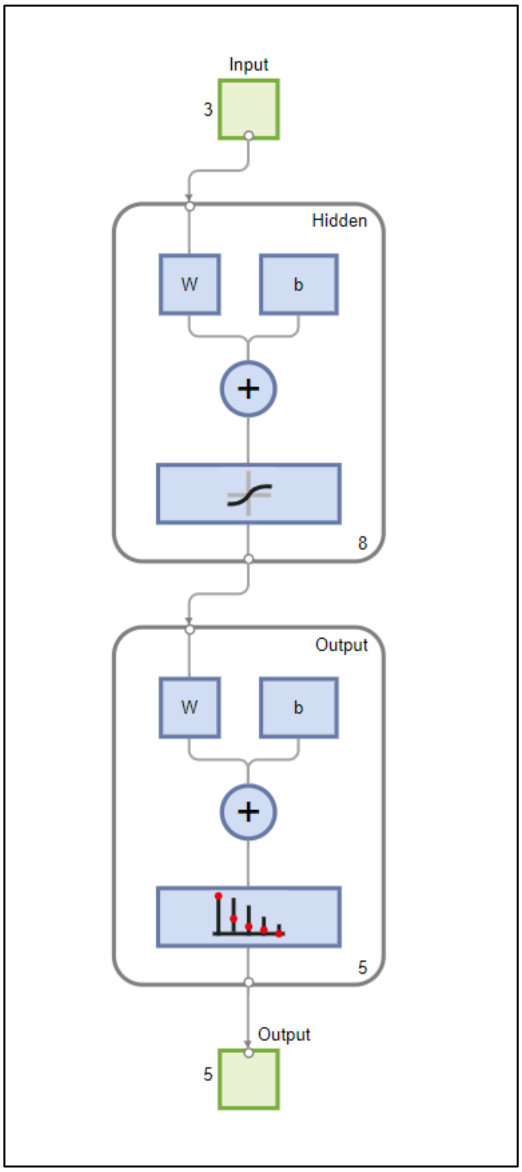

Figure 11: Matlab™ Model of ANN Network Architecture

5.5. Evaluating the Neural Network with DGA Dataset

The selected network architecture of an ANN consists of three input layers, five hidden layers and five output layers. The sigmoid function is the activation function selected for this simulation represented mathematically in Equation 3. The sigmoid activation function is where the computation of the model takes place. The three PCA transformed inputs are processed at the input neurons, and evaluated by the sigmoid function which processes the inputs and makes calculations in the hidden layers.

$$F(x) = \frac{1}{1+ e^{-x}} \tag{3}$$

Where ***x*** is the input value where the samples will be received at the input layer and **e** is the mathematical constant equivalent to **2.718**. The output of the model is shown on the output layer. Table 11 shows the typical output in a vector format as it appears at the output layer and its fault description.

Table 11: Description of Fault Code on ANN Output Layer

No	Fault code	Condition description
1	$[0\ 0\ 0\ 0\ 1]^T$	Normal degradation (N)
2	$[0\ 0\ 0\ 1\ 0]^T$	TL- Temperature fault <300
3	$[0\ 0\ 1\ 0\ 0]^T$	TH - Temperature fault>300
4	$[0\ 1\ 0\ 0\ 0]^T$	PD - Partial discharge of low and high energy
5	$[1\ 0\ 0\ 0\ 0]^T$	D – Arcing including discharge of low and high energy

5.6. Training and Validation

The dataset is randomly divided into training, validation and testing with a proportion of 70%, 20% and,10% respectively. The Bayesian Regularizer is the selected algorithm with 1000 iterations for training with the error target of zero. The 20% of the dataset is validated to strengthen the generalisation capability of the classifier.

5.7. Testing the ANN Classifier

Once the classifier has been trained with the raw data, the new dataset that has not been used before is fed into the classifier to test its generalisation capability. The 27 test samples which represent 10% of the total samples are selected randomly to test the accuracy of the classifier. Table 12 shows the results of the test for the 27 samples. It shows the comparison of the four conventional DGA methods and ANN method to the actual condition of the machine. From Table 12, the IEC, Dornenburg and Rogers ratio out of the range challenge is noticeable with 16, 6 and 9, respectively of test samples falling outside the classification range. This supports the literature on its limitation to provide a decision for samples that fall as followed by the outside its range. The ANN model has managed to make predictions for all of the test samples supplied. Seven of its predictions were misclassified.

Table 12: Comparison of Results with Actual Condition

Sample No	H2	CH4	C2H6	C2H4	C2H3	IEC faultcode	Dorn Faultcode	Duval Faultcode	Rogers Faultcode	Actual Fault	Predicted Fault
Sample 15	227	759	156	3133	191	NA	TH	TH	TH	TH	TH
Sample 23	170	320	53	520	3	NA	TH	TH	TH	TH	TH
Sample 32	650	53	34	20	0	PD	PD	TH	PD	PD	PD
Sample 34	300	490	180	360	95	NA	TH	TH	NA	TH	D
Sample 69	117	167	48	481	7	NA	TH	TH	TH	TL	TH
Sample 76	307	22	2	33	109	NA	NA	D	NA	D	PD
Sample 77	60	144	67	449	9	NA	TH	TH	TH	TL	TL
Sample 80	208	1774	408	1431	1	NA	TH	TH	TH	TL	TL
Sample 87	75	15	7	14	26	D	D	D	NA	D	D
Sample 90	1230	163	27	233	692	D	D	D	D	D	D
Sample 98	755	229	32	404	460	D	D	D	D	D	D
Sample 12	105	125	71	166	10	NA	TH	TH	TH	N	PD
Sample 13	80	0	200	100	4	PD	NA	TH	PD	N	N
Sample 13	125	100	100	150	20	D	NA	TH	NA	N	TH
Sample 13	50	30	0	0	5	D	NA	D	NA	N	N
Sample 14	144	13	36	3	2	NA	PD	D	NA	PD	PD
Sample 15	503	40	25	14	0	PD	PD	TH	PD	PD	PD
Sample 20	246	596	322	872	3	NA	TH	TH	TH	TL	TL
Sample 20	1245	3865	721	2449	12	NA	TH	TH	TH	TL	TL
Sample 20	35	78	29	210	5	NA	TH	TH	TH	TL	N
Sample 22	546	1367	228	1196	14	NA	TH	TH	TH	TH	TL
Sample 22	334	872	221	1481	37	NA	TH	TH	TH	TH	TH
Sample 23	70	235	96	369	7	NA	TH	TH	TH	TH	TH
Sample 23	730	1588	431	3972	74	NA	TH	TH	TH	TH	TH
Sample 23	360	2933	23292	21341	94	TL	TH	TH	NA	TH	TH
Sample 24	52	31	3	2	5	TL	NA	D	NA	N	N
Sample 27	107	91	79	157	36	D	NA	TH	NA	N	N

Figure 12 is a graphical representation of Table 12. The ANN, Dornenburg, Rogers, Duval and IEC prediction accuracy are compared to the actual condition of the machine for each fault type. Overall, ANN outperforms individual conventional methods in terms of accuracy. This image supports the literature's contention that no single model is ideal for all situations [3].

Figure 12 shows evidence of the challenges discussed in the literature regarding the ratio methods returning an undefined value, with IEC, Dornenburg and Rogers unable to classify 16, 6 and 9 samples, respectively. This was followed by the Duval Triangle's tendency to detect a fault even in normal conditions. According to the literature, Duval Triangle should only be used when a fault is present, as it always predicts a fault even for a normal sample [5,28].

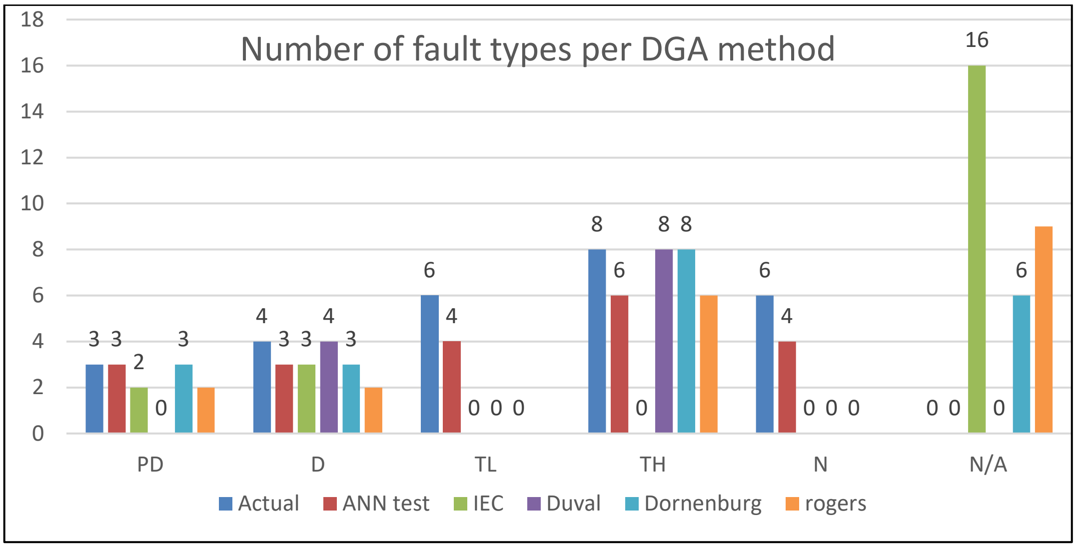

Figure 12: Fault Types per DGA Method

6. Results and Discussion

This chapter discusses and compares the overall performance of the Machine learning technique to the conventional DGA methods as applied in this study. The popularly used performance measures are applied to the ANN classifier to evaluate its overall performance. The performance measures will be assessed to evaluate the ANN model's ability to consistently make correct predictions.

6.1. Performance Metrices for ANN Model

The performance measures calculate the accuracy, specificity, precision, recall and the Mean Square Error (MSE). These performance metrics examine the model's overall performance to ensure no bias in performance evaluation. It is to be noted that where classes 1 to 5 are denoted in the diagram, they represent the class outputs as follows:

Class 1 – Partial Discharge (PD).

Class 2 – Discharge of low and high energy (D)

Class 3 – Thermal Temperature below 300 degrees Celsius (TH)

Class 4 – Thermal Temperature above 300 degrees Celsius (TH)

Class 5 – Normal/ No Fault (N)

The performance metrics are discussed in detail in the next sections.

6.1.1. The Confusion Matrix for Accuracy in Prediction

A confusion matrix is a visual representation of the ground truth of the model prediction. Accuracy metrics such as a confusion matrix are often not enough to determine the model's performance. They need to be supplemented by other metrics that are discussed in the next sections to get an overall performance of the model.

Figures 13 and 14 below show the model's performance for testing and training, respectively. The horizontal axis indicates the targeted output, while the vertical axis is the model's predicted output. The fault codes are represented on the axis by numbers 1 to 5 and represent the class output as discussed in section 6.1 with a 1 denoting the class 1 fault of Partial discharge. In Figure 13, all three partial discharge (class 1) fault conditions were correctly predicted by the model, meaning the model predicted Partial discharge with an accuracy of 100%. It can also be seen in Figure 14, that the model accurately predicted all PD faults.

The model reaches an overall accuracy of 92% on training the algorithm to recognise patterns in the dataset, however once the model is fed with new testing data, it can make 74% of correct predictions. Compared to conventional DGA methods, this is a good performance.

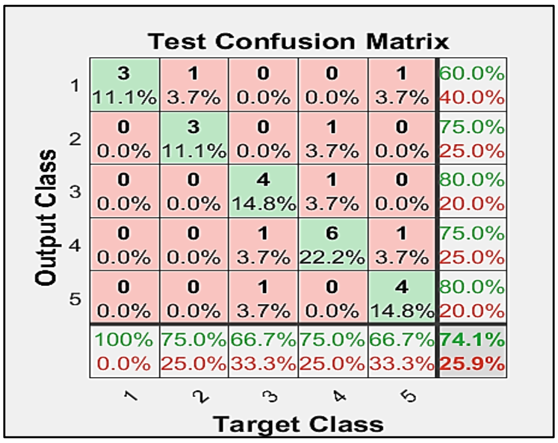

Figure 13: Confusion Matrix for Testing Data

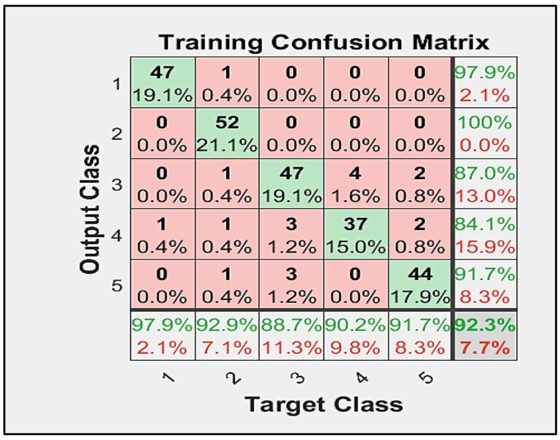

Figure 14: Confusion Matrix for Training Data

6.1.2. Mean Square Error

The Mean Square Error (MSE) is calculated as the average of the squared differences between the actual and predicted outputs. It measures the average squared deviation of the predictions from the true values. It reflects the disparity between the actual and expected values. During training, the model undergoes 1000 iterative epochs. Figure 15 illustrates that as the classifier begins to process training data, the error decreases for both training and testing datasets. However, around 190 iterations, the model begins to overfit; while the training error continues to decline, the testing error starts to rise. The optimal performance of the classifier is observed at approximately 190 epochs, as indicated in the diagram. A lower error value signifies better prediction accuracy.

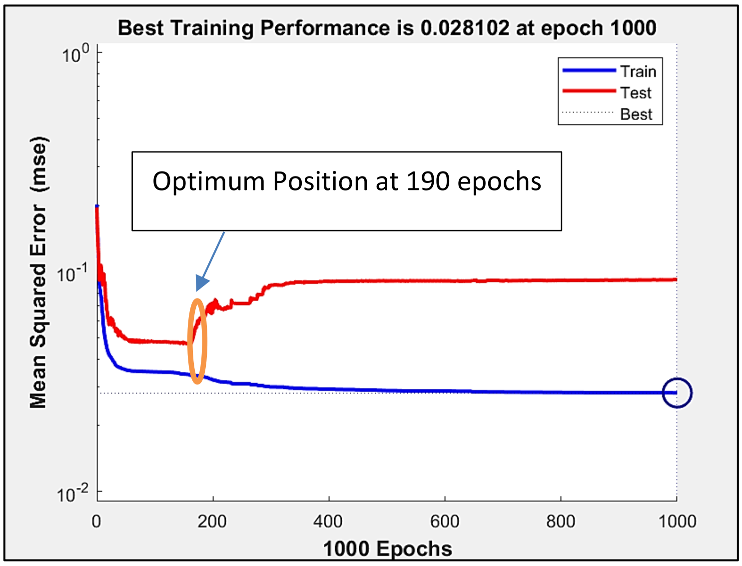

Figure 15: Mean Square Error

6.1.3. Receiver Operating Curve

The receiver operating curve in Figure 16 shows the true positive rate for the different fault classes. The diagonal line represents the area where the model randomly guesses the positive class half the time. The Y-axis represents the rate at which the model makes correct classification, the X-axis represents the rate at which the model makes errors in classification. The closer the graph is to the Y -axis, the better the generalisation and accuracy of the model. In this case while there is a variance in the rate at which the model makes positive prediction, the model performance is performing well with the rate closer to one.

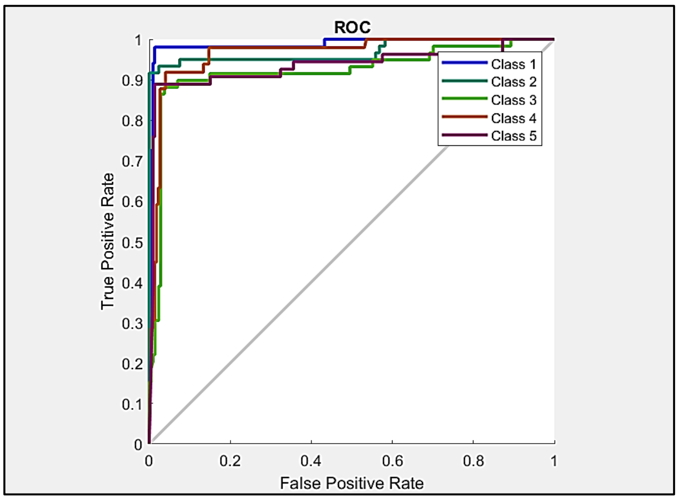

Figure 16: Receiver Operating Curve

6.1.4. Accuracy

Accuracy represents the proportion of correct predictions that a classifier makes relative to the overall predictions made. Equation 4 below represents accuracy.

$$Accuracy = \frac{Tp+Tn}{Tp+Tn+Fp+Fn}\ x\ 100 \tag{4}$$

Tp represents True Positive, ***Tn*** denotes True Negative, ***Fp*** denotes false positive and ***Fn*** represents False negative predictions.

Accuracy is a good measure of prediction accuracy however, used independently it doesn't give the true picture of the overall performance of the model. For an example, a binary model with an imbalanced dataset that makes predictions with 95% of the dataset having a certain class output will tend to favour the class output with the majority. Such a model can make predictions with an accuracy of 95%, which is misleading.

6.1.5. Precision

Precision shows the comparison between the true positives with relative to the total positives outcomes. It is a metric that shows the proportion of the positive predictions that was predicted correctly. Precision is represented by Equation 5 below.

$$\text{Precision} = \frac{\text{Tp}}{Tp+Fp}\ \text{x}\ 100 \tag{5}$$

6.1.6. Recall

The sensitivity/Recall/ Hit rate of the model is the comparison between the true positives to all the positives in ground truth. Recall is represented by Equation 6 below.

$$\text{Recall} = \frac{\text{Tp}}{Tp+Fn}\ \text{x}\ 100 \tag{6}$$

6.1.7. Specificity

The specificity/selectivity measures the proportion of the actual negatives that were identified correctly. Specificity is represented by Equation 7 below.

$$\text{Specificity} = \frac{\text{Tn}}{Tn+Fp}\ \text{x}\ 100 \tag{7}$$

6.1.8. Overall Performance

The overall performance of the model is depicted by Table 13. From Table 13, the model performs better for predicting PD and D type of faults with an average performance of 87% and 85% respectively.

Table 13: Performance Metrics of the ANN Overall Performance

fault type	**Accuracy**	**Precision**	**Recall**	**Specificity**	**Average Performance**
PD	89%	60%	100%	100%	87%
D	93%	75%	75%	96%	85%
TL	85%	80%	67%	91%	81%
TH	85%	75%	75%	89%	81%
N	89%	80%	67%	91%	82%

The average performance for thermal faults of low energy and high energy is the lowest, with an average performance of 81%. This means the ANN model found it more challenging to make correct predictions for temperature faults (TL and TH). Relating this to the confusion matrix, seen in Figure 17, the model made four correct predictions for TL fault (Class 3) and two misclassifications i.e. Normal (N) and high temperature fault (TH). This is not too unusual as these are the adjacent conditions, which means that one misclassified condition could be at the early development at the boundary between normal and low temperature fault and the other at the advanced phase at the boundary between low and high temperature fault. Thereby the model mistaken them for a low temperature fault.

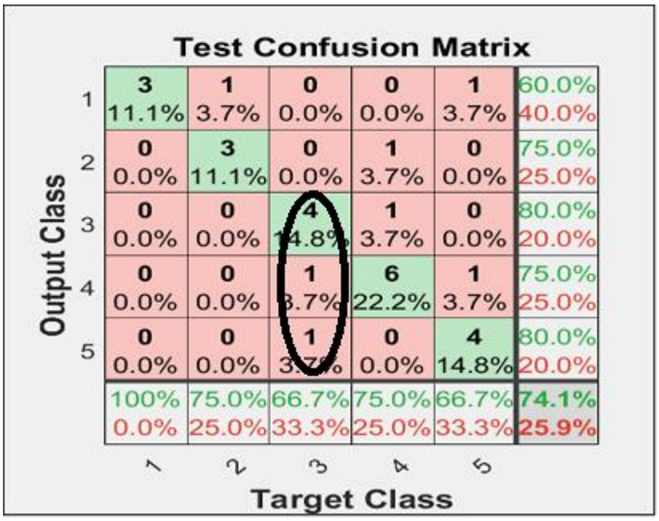

Figure 17: Confusion Matrix Showing Class 3 (TL) Fault
Misclassified as Class 5 (N) and Class 4 (TH)

6.2. Conclusion

The study results show that the ANN classifier achieves a probabilistic output for the fault diagnosis of oil-immersed transformers. It overcomes the deficiency of traditional DGA methods. The results show that ANN gave accurate results where the conventional method failed, in addition the learning process is faster and can be implemented easily on a software. The classifier must be tested with inconsistent data to test its robustness to the erroneous dataset. In addition, real data need to be acquired to model as the adaptive synthesizer generated data does not paint a realistic picture of the machine.

7. Conclusion

This study aimed to investigate the extent to which machine learning algorithms correctly classified DGA. This was achieved through a Matlab™ based machine learning classifier for the DGA dataset from different literature with most of the dataset acquired from IEC TC10. The classifier applied in this study is Artificial Neural Network. ANN model was constructed, evaluated and compared with the four popular conventional DGA methods to determine their accuracy in classifying faults correctly.

The buildup of this research was by first investigating the topic covering the challenges brought by the conventional DGA methods and the opportunities currently available to address them. Dissolved Gases Analysis tools remain the most preferred techniques to detect and classify power transformer faults before a catastrophic failure. The main challenge is the time and accuracy factor associated with it.

The proposed ANN model achieved a higher accuracy than the conventional DGA methods. This shows that ANN could be a powerful fault prediction technique especially for costly and critical assets in the power system. The conventional methods, while easy to implement, still require expert knowledge to apply successfully. They are also heavily reliant on an individual's state of mind. A large volume of maintenance data can be overwhelming for humans to interpret while, if appropriately trained, machines can execute timeously and accurately.

Other benefits of the ANN and machine learning tools in general are that they can be easily implemented on a computer software with less hassles and can be monitored by non-experience maintenance staff for decision making. The challenge with machine learning classifiers however is that they do not adapt to new data, therefore for new data it will constantly need to be retrained so that it learns the pattern of the new data to be able to make predictions.

This study yielded the intended outcome which was to evaluate the extent to which ANN can be used to make accurate predictions. The predictions were made for all class output with an average accuracy of 77% for ANN, 55% for Dornenburg, 40% for Duval, 38% for Roger and 32% for IEC, as in Table 10. While most of the research in the literature uses machine learning applied to conventional DGA methods for predictions, this study has shown that ANN can be used independently and still yield a good performance. The conventional methods have been used successfully over the years, the present study has proved that their accuracy is still a major concern. Condition monitoring has evolved and technology has presented intelligent ways for machines to make decisions independently with less human intervention.

Future work in this area can consider an increased number of samples with a better balanced dataset to avoid using the synthetic dataset generated through the ADASYN technique. While the adaptive synthesiser helps create pseudo data points for the purpose of this study, it is however not a true reflection of the actual condition of the machine.

References

[1] S. Li, J. Li, “Condition monitoring and diagnosis of power equipment: Review and prospective”, IET, High Volt, vol. 2, no. 2, pp.82–91, 2017.

[2] CIGRE WG A2.37, “Transformer Reliability Survey”, Brochure 642, 2015.

[3] IEC 60599, “Mineral oil-filled electrical equipment in service–Guidance on the interpretation of dissolved and free gases analysis”, IEC, 2022.

[4] IEEE C57.104-2019, “IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformer”, IEEE Transformer Committee

[5] S. Wani, A. Rana, S. Sohail, O. Rahman, S. Parveen and S. Khan, "Advances in DGA based condition monitoring of transformers: A review", Renewable and Sustainable Energy Reviews, vol. 149, p. 111347, 2021.

[6] CIGRE TB D1/A2, “Advance in DGA interpretation”, TB 771, 2019

[7] M. Duval, “Interpretation of Gas-In-Oil analysis Using New IEC Publication 60599 and IEC TC 10 Databases”, IEEE Electr. Insul. Mag., Vol. 17, No. 2, pp. 31-41, 2001.

[8] S. Ghoneim, I. Taha and N. Elkalashy, "Integrated ANN-based proactive fault diagnostic scheme for power transformers using dissolved gas analysis", IEEE Transactions on Dielectrics and Electrical Insulation, vol. 23, no. 3, pp. 1838-1845, 2016.

[9] I. Fofana and Y. Hadjadj, “Electrical-Based diagnostic techniques for assessing insulation condition in aged transformers” Energies, vol. 9, no. 9, p. 679, August 2016, doi: 10.3390/en9090679.

[10] F.Nowlan, H. Heap, “Reliability Centred Maintenance study”, 1978

[11] T.Mokwana, “ISO55000 Series: Maintenance Strategies”, 2020

[12] J. Holtzhausen, Vosloo. W, “High Voltage Engineering Practice and Theory”,

[13] 240-170000825, “2021 power transformer appraisal report for units installed in the distribution network”, Eskom Standard, 2021

[14] M. Hussain, S. Refaat and H. Abu-Rub, "Overview and Partial Discharge Analysis of Power Transformers: A Literature Review", IEEE Access, vol. 9, pp. 64587-64605, 2021.

[15] A. Abu-Elanien, M. Salama, “Asset management techniques for transformers”, Electric Power Systems Research, Vol 80, Issue 4, pp 456-464, 2010

[16] G. Rigatos and P. Siano, “Power transformers’ condition monitoring using neural modeling and the local statistical approach to fault diagnosis,” International Journal of Electrical Power & Energy Systems, vol. 80, pp. 150–159, 2016.

[17] A. Vosoughi, M. Samimi, “Transformer fault type discrimination based on window calculation method and frequency response analysis, Measurement 202, 2022
[18] IEEE Std C57.149, “IEEE Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers”, IEEE Standards Association, 2012
[19] Cigré WG A2/26, “Mechanical condition assessment of transformer windings using Frequency Response Analysis (FRA)”, CIGRE, 2008.
[20] R Liao, H Zheng, S Grzybowski, “Particle swarm optimization least-squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers”, Electr Power Syst Res, vol 81, pp 2074–2080.
[21] . Soni and B. Mehta, "Review on asset management of power transformer by diagnosing incipient faults and faults identification using various testing methodologies", Engineering Failure Analysis, vol. 128, p. 105634, 2021.
[22] G. Rigatos and P. Siano, “Power transformers’ condition monitoring using neural modeling and the local statistical approach to fault diagnosis,” International Journal of Electrical Power & Energy Systems, vol. 80, pp. 150–159, 2016.
[23] V. Galdi, L. Ippolito, A. Piccolo and A. Vaccaro, “Neural diagnostic system for transformer thermal overload protection,” IEE Proceedings - Electric Power Applications, vol. 147, no. 5, p. 415, 2000.
[24] L. Ippolito, P. Siano, “Identification of Tagaki–Sugeno–Kang fuzzy model for power transformers predictive overload system”, IEE Proc – Gener Transm Distrib, pp. 582-589, 2004
[25] S. Islam, T. Wu and G. Ledwich, "A novel fuzzy logic approach to transformer fault diagnosis”, IEEE Transactions on Dielectrics and Electrical Insulation, vol. 7, no. 2, pp. 177-186, 2000.
[26] S. Tenbohlen, S. Coenen, M. Djamali, “Diagnostic measurements for power transformers”, Energies, pp. 347–372, 2016
[27] A. Mohamed, “Comparative Study of Four Supervised Machine Learning Techniques for Classification”, International Journal of Applied Science and Technology, vol. 7, no.5, 2017
[28] ASTM D923, “Standard Practices for Sampling Electrical Insulating Liquids”, 2015
[29] H. de Faria, J. Costa and J. Olivas, "A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis", Renewable and Sustainable Energy Reviews, vol. 46, pp. 201-209, 2015.

[30] U. Rao, I. Fofana, K. Rajesh and P. Picher, "Identification and Application of Machine Learning Algorithms for Transformer Dissolved Gas Analysis", IEEE Transactions on Dielectrics and Electrical Insulation, vol. 28, no. 5, pp. 1828-1835, 2021.

[31] S. Wani, D. Gupta, M. Farooque and S. Khan, "Multiple incipient fault classification approach for enhancing the accuracy of dissolved gas analysis (DGA)", IET Science, Measurement & Technology, vol. 13, no. 7, pp. 959-967, 2019.

[32] Q. Su, C. Mi, L. Lai and P. Austin, "A fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer", *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 593-598, 2000.

[33] H. Yang, C. Liao, "Adaptive fuzzy diagnosis system for dissolved gas analysis of power transformers,", IEEE Transactions on Power Delivery, vol. 14, no. 4, pp. 1342-1350, 1999

[34] A. Abu-Siada, S. Hmood and S. Islam, “A New Fuzzy Logic Approach for Consistent Interpretation of Dissolved Gas-in-Oil Analysis”, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 20, No. 6, 2013

[35] Y. Zhang, X. Ding, Y. Liu and P. Griffin, “Artificial Neural Network approach to Transformer Fault diagnosis”, IEEE Transactions on Power Delivery, Vol. 11, No. 4, 1996

[36] I. Taha, S. Ibrahim and D. Mansour, "Power Transformer Fault Diagnosis Based on DGA Using a Convolutional Neural Network with Noise in Measurements," *IEEE Access*, vol. 9, pp. 111162-111170, 2021.

[37] J. Guardado, J. Naredo, P. Moreno and C. Fuerte, “A Comparative Study of Neural Network Efficiency in Power Transformers Diagnosis Using Dissolved Gas Analysis”, IEEE Transactions on Power Delivery, vol. 16, no 4, Pages 643 – 647, 2001

[38] P. Mirowski and Y. LeCun, "Statistical Machine Learning and Dissolved Gas Analysis: A Review", IEEE Transactions on Power Delivery, vol. 27, no. 4, pp. 1791-1799, 2012.

[39] T. Kari et al, “Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm”, IET Gener Transm. Distrib, vol. 12, no. 21, pp. 5672–5680, 2018.

[40] C. Wei, W. Tang, Q. Wu, “Dissolved gas analysis method based on novel feature prioritisation and support vector machine”, IET Electr. Power Appl., vol8, no 8, pp. 320–328, 2014

[41] IEEE Std C57.149, “IEEE Guide for the Application and Interpretation of Frequency Response Analysis for Oil-Immersed Transformers”, IEEE Standards Association, 2012

[42] S. Ghoneim and I. Taha, "A new approach of DGA interpretation technique for transformer fault diagnosis", International Journal of Electrical Power & Energy Systems, vol. 81, pp. 265-274, 2016.

[43] O. Gouda, S. El-Hoshy, S. Ghoneim, “Enhancing the Diagnostic Accuracy of DGA Techniques Based on IEC-TC10 and Related Databases”, IEEE Access, vol 9,2021

[44] R.De Aquino, M. Lira, et La, “a fuzzy system for detection an incipient fault in power transformers based on gas-in-oil analysis”, IEEE international conference on fuzzy systems, 2020

[45] H Sun, Y Huang, C Huang, “a review of Dissolved gas analysis I Power Transformers”, Energy procedia, vol 14, p 1220 – 1225, 2012

[46] M. Wang, “A novel extension method for transformer fault diagnosis”, IEEE Trans Power Deliv, Vol 18, no 1, Pages 164 – 169, 2003.

[47] N. Dhote, J. Helonde, “Improvement in transformer diagnosis by DGA using fuzzy logic”, Journal of Electrical Engineering and Technology, vol. 9, No.2, pp 615 – 621, 2014

[48] W. Cai, “The extension set and incompatibility problem,” J. Scientific Exploration, vol. 1, pp. 81–93, 1983.

[49] Y. Hua et al, “A fault diagnostic method for oil-immersed transformer based on multiple probabilistic output algorithms and improved DS evidence theory”, Electrical Power and Energy Systems, vol 137,2022

[50] A. Peimankar et al, “Ensemble classifier selection using multi-objective PSO for fault diagnosis of power transformers”, IEEE Congress on Evolutionary Computation, pp 3622–3629, 2016.

[51] J. N’cho, I. Fofana, Y. Hadjadj and A. Beroual, “Review of Physicochemical-Based Diagnostic Techniques for Assessing Insulation Condition in Aged Transformers,” Energies, vol. 9, no. 5, p. 367, May 2016, doi: 10.3390/en9050367.

[52] R. Polikar, “Ensemble based systems in decision making”, Circuits and Systems Magazine, IEEE, vol. 6, no. 3, pp. 21–45, 2006. D. Sarma, G. Kalyani, “ANN approach for condition monitoring of power Transformers using DGA”, Warangal – 506004, pp 444 -447,2004,

[53] N. Pattanadec, et al, “The conformity of DGA interpretation techniques: Experience from transformer 132 units”, IEEE 20th International Conference on Dielectric Liquids (ICDL), 2019

[54] I.Taha, S. Ghoneim, A. Duaywah,"Refining DGA methods of IEC Code and Rogers four ratios for transformer fault diagnosis", IEEE Power & Energy Society, 2016

Bibliography

[1] S. Li, G. Wu, B. Gao, et al.: "Interpretation of DGA for transformer fault diagnosis with complementary SaE-ELM and arctangent transform", IEEE Trans. Dielectr. Electr. Insul., vol.23, no1, pp. 586–595, 2016,

[2] H. Malik, et al, "Probabilistic Neural Network Based incipient fault identification using DGA dataset", Second international Symposium on Computer Vision and the internet, pp 665- 672, 2015

[3] I. Taha, S. Desspuky, S. Ghoneim, "Transformer fault types and severity class prediction based on neural pattern-recognition techniques",Electric Power Systems Research 191, 2021.

[4] K. Patekar, B. 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, 2020

Appendix A: DGA Database

H2	CH4	C2H6	C2H4	C2H2	actual	Reference
230	15,9	16,3	1,9	2,4	PD	[53]
111,7	19,4	104,1	6,4	3,8	PD	[53]
124	14	4	0,0001	13	D1	[53]
102	6	6	7	10	D1	[53]
99	170	20	200	190	D2	[53]
310	230	54	610	760	D2	[53]
24	109	69	0,0001	0,0001	T1	[53]
10	26	147	6	0,0001	T1	[53]
88,7	460,8	94,9	184,6	18,3	T2	[53]
119	670	286	934	19	T2	[53]
290	1260	231	820	8	T3	[53]
1550	2740	816	5450	184	T3	[53]
2078	486	380	1508	700	D1	[41]
57	589	211	749	0	T3	[41]
227	759	156	3133	191	T3	[41]
14,7	3,7	10,5	2,7	0,2	N	[45]
345	112,3	27,5	51,5	58,8	D1	[45]
181	262	41	28	0	T2	[45]
173	334	172	812,5	37,7	T4	[45]
127	107	11	154	224	D2	[45]
60	40	6,9	110	70	D2	[45]

220	340	42	480	14	T4	[45]
170	320	53	520	3,2	T4	[45]
27	90	42	63	0,2	T3	[45]
565	53	34	47	0	D1	[45]
56	286	96	928	7	T4	[45]
200	48	14	117	131	D2	[45]
78	161	86	353	10	T4	[45]
32,4	5,5	1,4	12,6	13,2	D2	[45]
980	73	58	12	0	PD1	[45]
160	130	33	96	0	T1	[45]
650	53	34	20	0	PD1	[45]
95	110	160	50	0	T2	[45]
300	490	180	360	95	T3	[45]
200	700	250	740	1	T3/T4	[45]
625	130	47	2	0	T2	[45]
56	61	75	32	31	D2	[45]
117	17	1	3	1	PD	[41]
32,93	2397	157	0,001	0,001	PD	[41]
78	20	11	13	28	D1	[41]
1230	163	27	233	692	D1	[41]
8200	3790	250	4620	277	D2	[41]
13	3	1	3	6	D2	[41]
130	140	2	120	0,001	T1	[41]
78	66	283	2,6	0,001	T1	[41]

30,4	117	44,2	138	0,1	T2	[41]
27	90	42	63	0,2	T2	[41]
1100	1600	221	2010	26	T3	[41]
290	966	299	1810	57	T3	[41]
24	13	5	43	319	Arcing	[51]
266	584	328	862	1	overheating	[51]
160	10	3	1	1	Discharge	[51]
80	619	326	2480	0	severe overheating	[51]
231	3997	1726	5584	0	severe heating	[51]
127	24	0	32	81	Arcing	[51]
9474	4066	353	6552	12997	Arcing	[51]
507	1053	297	1440	17	severe local heating	[51]
416	697	74	867	0	heating and arcing	[51]
441	207	43	224	261	Arcing	[51]
65	61	16	143	3	overheating	[51]
16	87	75	395	30	Overheating	[51]
212	38	15	47	78	Arcing	[51]
800	1393	304	2817	3000	Arcing	[51]
199	770	217	1508	72	overheating	[51]
4906	8784	1404	9924	9671	Arcing	[51]
425	17424	7299	37043	158	overheating	[51]
1076	95	4	71	231	PD	[51]

244	754	172	1281	27	overheating	[51]
117	167	48	481	7	overheating	[51]
858	1324	208	2793	7672	Arcing	[51]
137	369	144	1242	16	overheating	[51]
274	27	5	33	97	Arcing	[51]
1249	370	56	606	1373	Arcing	[51]
240	20	5	28	96	pd	[51]
33	79	30	215	5	overheating	[51]
307	22	2	33	109	Arcing	[51]
60	144	67	449	9	overheating	[51]
2004	9739	2750	5113	0	overheating	[51]
127	107	11	154	224	Arcing	[51]
208	1774	408	1431	1	overheating	[52]
9	453	590	211	0	Arcing	[52]
66	60	2	7	0	T1	[7]
3420	7870	1500	6990	33	T2	[7]
30	200	114	308	8	T3	[7]
9340	995	60	6	7	PD	[7]
1330	10	20	66	182	D1	[7]
75	15	7	14	26	D2	[7]
78	20	11	13	28	D1	[7]
305	100	33	161	541	D1	[7]
1230	163	27	233	692	D1	[7]
645	86	13	110	317	D1	[7]

95	10	0	11	39	D1	[7]
595	80	9	89	244	D1	[7]
1330	10	20	66	182	D1	[7]
440	89	19	304	757	D2	[7]
545	130	16	153	239	D2	[7]
7150	1440	97	1210	1760	D2	[7]
755	229	32	404	460	D2	[7]
1570	1110	175	1780	1830	D2	[7]
3090	5020	323	3800	2540	D2	[7]
1820	405	35	365	634	D2	[7]
260	215	35	334	277	D2	[7]
75	15	7	14	26	D2	[7]
60	5	2	21	21	D2	[7]
1500	395	28	395	323	D2	[7]
20000	13000	1850	29000	57000	D2	[7]
3700	1690	128	2810	3270	D2	[7]
2770	660	54	712	763	D2	[7]
1170	225	18	312	325	D2	[7]
1270	3450	520	1390	8	T1/2	[7]
3420	7870	1500	6990	33	T1/2	[7]
48	610	29	10	0	T1/2	[7]
12	18	4	4	0	T1/2	[7]
66	60	2	7	0	T1/2	[7]
14	44	124	7	1	T1/2	[7]

8800	64064	72128	95650	0	T3	[7]
6709	10500	1400	17700	750	T3	[7]
290	966	299	1810	57	T3	[7]
2500	10500	4790	13500	6	T3	[7]
400	940	210	820	24	T3	[7]
6	2990	29990	26076	67	T3	[7]
290	1260	231	820	8	T3	[7]
107	143	34	222	2	T3	[7]
134	134	157	45	0	N	[7]
100	200	200	200	20	N	[7]
0	225	225	110	3	N	[7]
105	125	71	166	10	N	[7]
100	50	65	50	15	N	[7]
100	70	70	170	10	N	[7]
150	0	0	220	8	N	[7]
0	224	224	112	5	N	[7]
200	50	50	200	3	N	[7]
85	0	80	35	70	N	[7]
175	0	100	375	3	N	[7]
80	0	200	100	4	N	[7]
150	0	200	100	15	N	[7]
125	100	100	150	20	N	[7]
200	3	50	200	0	N	[7]
50	30	0	0	5	N	[7]

100	70	70	170	10	N	[7]
95	280	250	150	10	N	[7]
60	40	50	60	3	N	[7]
84	79	52	166	56	N	[7]
66	111	90	110	15	N	[7]

Appendix B: The ANN Classifier Code Used.

```
clear; clc; close all; warning off;

% import TrainDataR.*

Data = readtable('ProjectData.xlsx');

% keeping heading names in categorical format

feature_heading =["H2" "CH4" "C2H6" "C2H4" "C2H2"];

feature_heading=categorical(feature_heading);

%converting faultcode into ANN version

Data.faultgroup = grp2idx(Data.faultgroup);

Data.faultgroup = dummyvar(Data.faultgroup);

%defining features and outputs

All_Features = Data(:,1:5);

All_Labels = Data.faultgroup;

%assigning empty variable to store new features and output:ADASYN

new_feat = [];

new_Output = [];

%measuring size of the data collected

new_Output_size = [];

%initiate while loop

i = 1;

%assign output matrix for fault output

Out = [[1 0 0 0 0]; [0 1 0 0 0] ;[0 0 1 0 0]; [0 0 0 1 0]; [0 0 0 0 1]];

%finding new features to balance the dataset ADASYN

while i < 6

[out_featuresSyn, out_labelsSyn] = ADASYN([All_Features.H2 All_Features.CH4 All_Features.C2H6
All_Features.C2H4 All_Features.C2H2], All_Labels(:,i),0.3);
```

```
new_feat = [new_feat ;out_featuresSyn];

X= repelem(Out(i,1:5),sum(out_labelsSyn),1);

new_Output = [new_Output; X];

new_Output_size = [new_Output_size sum(out_labelsSyn)];

i = i + 1;

end

%converting from table to array

All_Features1 = table2array(All_Features);

%combining new feature and outputs with new ones

All_Features = [All_Features1; new_feat ];

All_Features_DGA = [All_Features1; new_feat ];

%All_Features = array2table(All_Features, "VariableNames",feature_heading);

All_Labels = [All_Labels;new_Output];

my_data= [All_Features All_Labels];

%counting the size of each class output : the balanced data

X_count_size = [sum(All_Labels(:,1)) sum(All_Labels(:,2)) sum(All_Labels(:,3)) sum(All_Labels(:,4))
sum(All_Labels(:,5))];

X = All_Features;

X1 = All_Features;

t = All_Labels;

%preserving the names of the class output

classOrigin=["PD" "D" "TL" "TH" "N"];

%converting class output to names

t= onehotdecode(t,classOrigin,2);

tx= t;

%changing features data to log format
```

```
Xlog=log10(1+X);

%taking the mean of X

Xlog(:,1)= (Xlog(:,1)-mean(Xlog(:,1)))/std(Xlog(:,1));

Xlog(:,2)= (Xlog(:,2)-mean(Xlog(:,2)))/std(Xlog(:,2));

Xlog(:,3)= (Xlog(:,3)-mean(Xlog(:,3)))/std(Xlog(:,3));

Xlog(:,4)= (Xlog(:,4)-mean(Xlog(:,4)))/std(Xlog(:,4));

Xlog(:,5)= (Xlog(:,5)-mean(Xlog(:,5)))/std(Xlog(:,5));

%calculating PCA

[psc,scrs,~,~,pexp]=pca(Xlog);

%[psc,scrs,~,~,pexp]=pca(X);

pareto(pexp)

xlabel('individual selected component')

ylabel('Cumulative Explained variance [%]')

title("plot of explained variance")

% [idx c sumd]=kmeans(Xnorm,5)

% gscatter(scrs(:,1),scrs(:,2),idx)

figure, scatter(scrs(:,1),scrs(:,2),20,t)

figure, scatter3(scrs(:,1),scrs(:,2),scrs(:,3),20,t)

xlabel('Principal component 1')

ylabel('Principal component 2')

zlabel('Principal component 3')

title("Plot of Principal Component Analysis")

%axis([-0.5 0.5 -0.5 0.5 -0.5 0.5])

%legend(classOrigin)

%ANN Code

% Solve a Pattern Recognition Problem with a Neural Network
```

```
% Script generated by Neural Pattern Recognition app

% Created 13-Jan-2023 15:03:08

%

% This script assumes these variables are defined:

%

% features - input data.

% faultcode - target data.

%defining features and outputs

All_Features = scrs(:,1:3)';

%collecting perf data for repeatability

collect_perf_data = [];

collect_err_data = [];

%initialising a loop for repeatability of the results

init = 1;

while init >0

% Choose a Training Function

% For a list of all training functions type: help nntrain

% 'trainlm' is usually fastest.

% 'trainbr' takes longer but may be better for challenging problems.

% 'trainscg' uses less memory. Suitable in low memory situations.

trainFcn = 'trainbr'; % Scaled conjugate gradient backpropagation.

% Create a Pattern Recognition Network

hiddenLayerSize = 8;

net = patternnet(hiddenLayerSize, trainFcn);

%net = feedforwardnet(hiddenLayerSize,'trainscg');

% Choose Input and Output Pre/Post-Processing Functions
```

```
% For a list of all processing functions type: help nnprocess

net.input.processFcns = {'removeconstantrows','mapminmax'};

net.trainParam.epochs = 1000;

net.trainParam.goal = 0;

% Setup Division of Data for Training, Validation, Testing

% For a list of all data division functions type: help nndivision

net.divideFcn = 'dividerand'; % Divide data randomly

net.divideMode = 'sample'; % Divide up every sample

net.divideParam.trainRatio = 70/100;

net.divideParam.valRatio = 20/100;

net.divideParam.testRatio = 10/100;

% Choose a Performance Function

% For a list of all performance functions type: help nnperformance

net.performFcn = 'mse'; % Cross Entropy

% Choose Plot Functions

% For a list of all plot functions type: help nnplot

net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...

'plotconfusion', 'plotroc'};

x= All_Features;

t= All_Labels';

% 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);

dummyvar(yind)

y_use_later= onehotdecode(dummyvar(yind),classOrigin,2)

percentErrors = sum(tind ~= yind)/numel(tind);

%accuraccy of prediction

test_index=tr.testInd;

testperf = tind(tr.testInd);

%testperfx = testing(tr.testInd)

testperf1 = yind(tr.testInd);

testperf2 = sum(testperf == testperf1)/numel(tind)*100;

% Recalculate Training, Validation and Test Performance

trainTargets = t .* tr.trainMask{1};

valTargets = t .* tr.valMask{1};

testTargets = t .* tr.testMask{1};

trainPerformance = perform(net,trainTargets,y);

valPerformance = perform(net,valTargets,y);

testPerformance = perform(net,testTargets,y);

collect_err_data = [collect_err_data percentErrors];

collect_perf_data = [collect_perf_data performance];

init=init-1;

% View the Network

end

performance_final = sum(collect_perf_data)/100;

error_final = sum(collect_err_data)/100;

view(net)
```

```
% Plots

% Uncomment these lines to enable various plots.

figure, plotperform(tr)

figure, plottrainstate(tr)

figure, ploterrhist(e)

figure, cm = confusionchart(testperf1,testperf);

cm.Title= "Fault Type Confusion Matrix";

cm.XLabel='predicted Class';

cm.YLabel='True Class';

cm.ColumnSummary = 'column-normalized';

cm.RowSummary = 'row-normalized';

cm.ClassLabels = {'Pd','D','TL','TH','N'};

figure, plotroc(t,y);

% Deployment

% Change the (false) values to (true) to enable the following code blocks.

% See the help for each generation function for more information.

if (false)

% Generate MATLAB function for neural network for application

% deployment in MATLAB scripts or with MATLAB Compiler and Builder

% tools, or simply to examine the calculations your trained neural

% network performs.

genFunction(net,'myNeuralNetworkFunction');

y = myNeuralNetworkFunction(x);

end

if (false)

% Generate a matrix-only MATLAB function for neural network code
```

```
% generation with MATLAB Coder tools.
genFunction(net,'myNeuralNetworkFunction','MatrixOnly','yes');
y = myNeuralNetworkFunction(x);
end
if (false)
% Generate a Simulink diagram for simulation or deployment with.
% Simulink Coder tools.
gensim(net);
end
my=my_data(:,[1:5]);
the_wrong_features=my(test_index(testperf ~= testperf1),:);
the_wrong_label = tx(test_index(testperf ~= testperf1),:);
the_wrong_prediction = y_use_later(test_index(testperf ~= testperf1),:);
analyse = array2table(the_wrong_features,...
'VariableNames',{'H2','CH4','C2H6','C2H4','C2H2',});
Label = array2table(the_wrong_label,...
'VariableNames',{'True label'});
prediction = array2table(the_wrong_prediction,...
'VariableNames',{'Prediction'});
analyse(:,6:7)= [Label prediction];
analyse.Properties.VariableNames = {'H2','CH4','C2H6','C2H4','C2H2','True label','Prediction'
};
%calculation for Rogers method
g = 1;
collect_ROGER_data = [];
while g <274
```

```
R1 = All_Features_DGA(g,5)/All_Features_DGA(g,4);
if R1 < 0.1
R1 = 0;
elseif R1<=3
R1 = 1;
else R1 = 2;
end
R2 = All_Features_DGA(g,2)/All_Features_DGA(g,1);
if R2 < 0.1
R2 = 1;
elseif R2 <= 1
R2 = 0;
else R2 = 2;
end
R3 = All_Features_DGA(g,4)/All_Features_DGA(g,3);
if R3 < 1
R3 = 0;
elseif R3 <= 3
R3 = 1;
else R3 = 2;
end
if (R1 == 0) && (R2 == 0) && (R3 == 0)
ROGER_fault = "N";
elseif (R1 == 0) && (R2 == 1) && (R3 == 0)
ROGER_fault = "PD";
elseif (R1 == 1) && (R2 == 0) && (R3 == 2)
```

```
ROGER_fault = "D";
elseif (R1 == 0) && (R2 == 0) && (R3 == 1)
ROGER_fault = "TL";
elseif ((R1 == 0) && (R2 == 2) && (R3 == 1)) || ((R1 == 0) && (R2 == 2) && (R3 == 2))
ROGER_fault = "TH";
else
ROGER_fault = "NA";
end
collect_ROGER_data = [collect_ROGER_data; R1 R2 R3 ROGER_fault];
g=g+1;
end
%calculation for IEC method
a = 1;
collect_IEC_data = [];
while a <274
R1 = All_Features_DGA(a,5)/All_Features_DGA(a,4);
if R1 < 0.1
R1 = 0;
elseif R1<=3
R1 = 1;
else R1 = 2;
end
R2 = All_Features_DGA(a,2)/All_Features_DGA(a,1);
if R2 < 0.1
R2 = 1;
elseif R2 <= 1
```

```
R2 = 0;
else R2 = 2;
end
R3 = All_Features_DGA(a,4)/All_Features_DGA(a,3);
if R3 < 1
R3 = 0;
elseif R3 <= 3
R3 = 1;
else R3 = 2;
end
if (R1 == 0) && (R2 == 0) && (R3 == 0)
IEC_fault = "N";
elseif ((R1 == 0) && (R2 == 1) && (R3 == 0)) || ((R1 == 0) && (R2 == 1) && (R3 == 1))
IEC_fault = "PD";
elseif ((R1 == 1) && (R2 == 0) && (R3 == 1)) || ((R1 == 1) && (R2 == 0) && (R3 == 2)) ||...
((R1 == 2) && (R2 == 0) && (R3 == 2)) || ((R1 == 2) && (R2 == 0) && (R3 == 1))
IEC_fault = "D";
elseif ((R1 == 1) && (R2 == 0) && (R3 == 0)) || ((R1 == 0) && (R2 == 2) && (R3 == 0))
IEC_fault = "TL";
elseif ((R1 == 0) && (R2 == 2) && (R3 == 0)) || ((R1 == 2) && (R2 == 2) && (R3 == 0))
IEC_fault = "TL";
else
IEC_fault = "NA";
end
collect_IEC_data = [collect_IEC_data; R1 R2 R3 IEC_fault];
a=a+1;
```

```
end

%calculation for Duval method

d = 1;

collect_DUVAL_data = [];

while d <274

R1_CH4 = (All_Features_DGA(d,2)*100)/(All_Features_DGA(d,2)+ All_Features_DGA(d,5) +
All_Features_DGA(d,4));

R2_C2H4 = (All_Features_DGA(d,4)*100)/(All_Features_DGA(d,2)+ All_Features_DGA(d,5) +
All_Features_DGA(d,4));

R3_C2H2 = (All_Features_DGA(d,5)*100)/(All_Features_DGA(d,2)+ All_Features_DGA(d,5) +
All_Features_DGA(d,4));

if R1_CH4 > 98

DUVAL_fault = "PD";

elseif (R1_CH4 < 98) && (R2_C2H4 < 20) && (R3_C2H2 < 4)

DUVAL_fault = "TL";

elseif (R2_C2H4 >= 20) && (R3_C2H2 < 15)

DUVAL_fault = "TH";

elseif (R2_C2H4 >= 40) && (R3_C2H2 >= 4) && (R3_C2H2 < 29)

DUVAL_fault = "DT";

else

DUVAL_fault = "D";

end

collect_DUVAL_data = [collect_DUVAL_data; R1_CH4 R2_C2H4 R3_C2H2 DUVAL_fault];

d=d+1;

end

%calculation for Dornenburg method

b = 1;
```

```
collect_DORN_data = [];

while b <274

R1 = All_Features_DGA(b,2)/All_Features_DGA(b,1);

if R1 < 0.1

R1 = 0;

elseif R1<=1

R1 = 1;

else R1 = 2;

end

R2 = All_Features_DGA(b,5)/All_Features_DGA(b,4);

if R2 < 0.75

R2 = 0;

else R2 = 1;

end

R3 = All_Features_DGA(b,5)/All_Features_DGA(b,2);

if R3 < 0.3

R3 = 0;

else R3 = 1;

end

R4 = All_Features_DGA(b,3)/All_Features_DGA(b,5);

if R4 < 0.4

R4 = 0;

else R4 = 1;

end

if (R1 == 2) && (R2 == 0) && (R3 == 0) && (R4 == 1)

DORN_fault = "TH";
```

```
elseif (R1 == 0) && (R3 == 0) && (R4 == 1)

DORN_fault = "PD";

elseif (R1 == 1) && (R2 == 1) && (R3 == 1) && (R4 == 0)

DORN_fault = "D";

else

DORN_fault = "NA";

end

collect_DORN_data = [collect_DORN_data; R1 R2 R3 R4 DORN_fault];

b=b+1;

end

% Preparing output DGA dataset and fault analysis per method

IEC_x = array2table(categorical(collect_IEC_data(test_index,:)),...

'VariableNames',{'R1','R2','R3','IEC_fault',});

DORN_x = array2table(categorical(collect_DORN_data(test_index,:)),...

'VariableNames',{'R1','R2','R3','R4','DORN_fault',});

DUVAL_x = array2table(categorical(collect_DUVAL_data(test_index,:)),...

'VariableNames',{'R1','R2','R3','DUVAL_fault',});

ROGER_x = array2table(categorical(collect_ROGER_data(test_index,:)),...

'VariableNames',{'R1','R2','R3','ROGER_fault',});

true_label = array2table(tx(test_index),'variablenames',{'actual_fault',});

features_4_testing = my(test_index,:);

compare_data = array2table(features_4_testing,...

'VariableNames',{'H2','CH4','C2H6','C2H4','C2H2',});

Predictor = array2table(y_use_later(test_index),'variablenames',{'Predicted_fault',});

i =1;

sample_collector = []
```

```
while i < 274

sample = "Sample " + (i);

sample_collector = [sample_collector sample];

i = i + 1;

end

sample_collector = sample_collector';

case_number = array2table(categorical(sample_collector(test_index)),...

'VariableNames',{'Case No',});

%comparing DGA against actual fault and predictor

compare_data{:,6:11}= [IEC_x.IEC_fault DORN_x.DORN_fault ...

DUVAL_x.DUVAL_fault ROGER_x.ROGER_fault true_label.actual_fault Predictor.Predicted_fault];

compare_data.Properties.VariableNames = {'H2','CH4','C2H6','C2H4','C2H2',...

'IEC_fault','DORN_fault','DUVAL_fault','ROGER_fault','Actual_fault','Predicted_fault'};

compare_data = [case_number compare_data]
```