REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université du 20 Août 1955 Skikda

Faculté de Technologie

Département de Génie Mécanique

N° d'ordre : D012125032D

THESE

Présentée en vue de l'obtention du diplôme de

DOCTEUR

Filière : Electromécanique

Spécialité : Electromécanique

Par :

M. Abdelmoumene HECHIFA

Développement d'un système d'information pour le diagnostic et la prédiction des défauts des transformateurs électriques

Soutenue publiquement le : 19/07/2025

Devant le Jury composé par :

Nom et prénom	Grade	Qualité	Affiliation
Toufik SEBBAGH	MCA	Président	Université du 20 août 1955-Skikda
Abdelaziz LAKEHAL	Professeur	Encadreur	Université de Souk Ahras
Ridha KELAIAIA	Professeur	Co-encadreur	Université du 20 août 1955-Skikda
Adlen KERBOUA	MCA	Examinateur	Université du 20 août 1955-Skikda
Tarek KHOUALDIA	MCA	Examinateur	Université de Souk Ahras
Redha MENECEUR	Professeur	Examinateur	Université d'El Oued
Chouaib LABIOD	MCA	Invité	Université d'El Oued

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

The Ministry of Higher Education and Scientific Research

University of 20 August 1955-Skikda

Faculty of Technology

Department of Mechanical Engineering

N° of order : D012125032D

THESIS

Presented with a view to obtaining the diploma of

DOCTOR

Sector: Electromechanical

Specialty: Electromechanical

By:

Mr. Abdelmoumene HECHIFA

Development of an Information System for Fault Diagnosis and Prediction in Electrical Power Transformers

Publicly defended on: 19/07/2025

Jury:

First and last name	Grade	Quality	Affiliation
Toufik SEBBAGH	MCA	President	University of August 20, 1955-Skikda
Abdelaziz LAKEHAL	Professor	Supervisor	University of Souk Ahras
Ridha KELAIAIA	Professor	Co-supervisor	University of August 20, 1955-Skikda
Adlen KERBOUA	MCA	Examiner	University of August 20, 1955-Skikda
Tarek KHOUALDIA	MCA	Examiner	University of Souk Ahras
Redha MENECEUR	Professor	Examiner	University of El Oued
Chouaib LABIOD	MCA	Invited	University of El Oued

Dedication

In the Name of Allah, the Most Gracious, the Most Merciful

I dedicate this humble work to:

My dear parents May Allah reward them for their endless love, prayers, and sacrifices.

My brothers, sisters, and entire family for their unwavering support and encouragement.

And to a soul that left this world but remains alive in my heart
my friend, Dr. Laib Abdelmoumen
May Allah forgive him and grant him the highest place in Jannat al-Firdaus.

Acknowledgments

In the name of Allah, the Most Gracious, the Most Merciful. Alhamdulillah (All praise is due to Allah) for granting me the strength, patience, and blessings to complete this thesis.

I would like to express my deepest gratitude to my PhD supervisors, ***Prof. Abdelaziz Lakehal*** *(Souk Ahras University) and* ***Prof. Ridha Kelaiaia*** *(Skikda University), for their invaluable guidance, unwavering encouragement, and continuous support throughout my doctoral journey. Their mentorship has been instrumental in shaping my research and academic growth.*

I am especially grateful to ***Dr. Chouaib Labiod*** *(El Oued University) for his constant support, insightful feedback, and crucial assistance throughout my research. His encouragement and dedication have been a source of great inspiration, and his contributions played a vital role in the successful completion of this work.*

My sincere appreciation extends to ***Dr. Saurabh Dutta*** *(Manchester University) and* ***Prof. Hazlee Azil Illias*** *(Malaya University) for their generous support, thoughtful comments, and significant contributions during my internship in Malaysia. Additionally, I am indebted to* ***Dr. Arnaud Nafak*** *(Douala University), whose expertise was invaluable in setting up experiments, programming, and implementing algorithms.*

I am thankful to ***Dr. Mohammed Senoussaoui*** *(Mascara University) for his contributions and to* ***Dr. Said Djaballah*** *(Chlef University) for his guidance and encouragement, which were pivotal during critical stages of my research.*

A special note of gratitude goes to my brother and colleague, ***Ali Bebboukha*** *(El Oued University), for his unwavering support, guidance, and contributions to my success. His encouragement kept me motivated throughout this journey.*

I would also like to acknowledge ***Skikda University*** *for its support during my doctoral studies, as well as* ***El Oued University*** *for providing a nurturing academic environment throughout this time.*

Finally, I express my heartfelt thanks to the members of the jury, ***Dr. Toufik Sebbagh*** *,* ***Dr. Adlen Kerboua*** *(Skikda University),* ***Prof. Redha Meneceur*** *(El Oued University) and* ***Dr. Tarek Khoualdia*** *(Souk Ahras University) for their valuable time, thoughtful evaluation, and constructive feedback.*

List of Publications

Publications in refereed conferences

- **A. Hechifa**, A. Lakehal, L. Saidi, A. Nanfak, R. Kelaiaia, and M. E. A. Senoussaoui, "Improvement of Transformer Fault Diagnosis using Fuzzy Rule and Decision Tree Based on Dissolved Gas Analysis," in 2023 1st International Conference on Renewable Solutions for Ecosystems: Towards a Sustainable Energy Transition (ICRSEtoSET), 2023: **IEEE**, pp. 1-6.
- **A. Hechifa**, A. Lakehal, C. Labiod, A. Nanfak, and S. S. Ghoneim, "The effect of source data on graphical pentagons DGA methods for detecting incipient faults in power transformers," in 2023 International Conference on Decision Aid Sciences and Applications (DASA), 2023: **IEEE**, pp. 152-157.
- A. Lakehal, L. Saidi, **A. Hechifa**, and A. Nanfak, "A Methodology for Operational Fault Diagnosis in Electrical Power Transformer: Practical Application," in International Conference on Green Energy Conversion System, 2023: **Springer**, pp. 631-642.
- **A. Hechifa**, A. Lakehal, and R. Kelaiaia, "Diagnostics of Power Transformer Faults Using Prediction Fusion of Neural Network (MLP) with Naïve Bayes Algorithm Based on DGA Data." The 4th International Conference on Electromechanical Engineering (ICEE2022). 2022: **University of Skikda**, Algeria.
- **A. Hechifa**, A. Lakehal, C. Labiod, and R. Kelaiaia, " Improving the fault diagnosis of oil-filled transformers based on feature selection of multiple input vectors." Journée d'étude sur l'intelligence artificielle en électromécanique et ses applications industrielles (JEIAM'23). 2023: **University of El Oued**, Algeria.
- **A. Hechifa**, A. Lakehal, C. Labiod, R. Meneceur, A. Bebboukha, and S. Djaballah, " A comparative study for diagnosing and detecting faults in power transformers." 1st International Seminar on Mechatronics Innovation Materials, Renewable Energy, and Artificial Intelligence (ISMIMREAI'24). 2024 : **University of Tipaza**, Algeria.

Publications in refereed journals

- **A. Hechifa**, C. Labiod, A. Lakehal, A. Nanfak, and D.-E. A. Mansour, "A novel graphical method for interpretating dissolved gases and fault diagnosis in power transformer based on dynamique axes in circular form," **IEEE Transactions on Power Delivery**, 2024.
- **A. Hechifa**, A. Lakehal, L. Saidi, A. Nanfak, C. Labiod, D.-E. A. Mansour, and D. Said., "Improved intelligent methods for power transformer fault diagnosis based on tree ensemble learning and multiple feature vector analysis," **Electrical Engineering**, vol. 106, no. 3, pp. 2575-2594, 2024.
- **A. Hechifa**, S. Dutta, A. Lakehal, H. A. Illias, A. Nanfak, and C. Labiod, "Enhancing power transformer health assessment through dimensional reduction and ensemble approaches in Dissolved Gas Analysis," **IET Nanodielectrics**, 2024.
- **A. Hechifa**, A. Lakehal, A. Nanfak, L. Saidi, and C. Labiod, "Machine Learning Algorithms Fusion Based on DGA Data for Improving Fault Diagnosis of Electrical Power Transformer," **The Scientific Bulletin of Electrical Engineering Faculty**, vol. 23, no. 2, pp. 5-11.
- A. Nanfak, **A. Hechifa**, S. Eke, A. Lakehal, C. H. Kom, and S. S. Ghoneim, "A combined technique for power transformer fault diagnosis based on k-means clustering and support vector machine," **IET Nanodielectrics**, vol. 7, no. 3, pp. 175-187, 2024.

Abstract

Fault diagnosis and prevention in oil-immersed transformers are critical for maintaining reliable power systems. While dissolved gas analysis (DGA) is a widely used diagnostic method, traditional techniques often struggle with overlapping fault regions. This thesis proposes innovative solutions that integrate machine learning, deep learning, and optimization algorithms to address these limitations. The research evaluates DGA inputs vectors and combines them with tree-based algorithms for enhanced accuracy. To further enhance performance, principal component analysis (PCA) is applied to extract key features, which are then combined with ensemble learning techniques such as boosting, bagging, and decorate. These approaches are systematically compared with conventional techniques, demonstrating their superior diagnostic accuracy. The novel contribution of this thesis is to introduce Circle Methods 1 and 2, are proposed, with dynamic axes and variable angles. These methods innovatively base their analysis on the convergence of samples with similar characteristics around their centers and the spacing of these centers for samples with differing characteristics exploit optimization algorithms to find the best fault zone distribution. Validated against real-world datasets, including the IEC TC10 database, the Circle Methods deliver superior diagnostic accuracy compared to existing approaches. These advancements significantly enhance the interpretability of DGA and establish a robust framework for predictive maintenance and fault prevention in power transformers.

Keywords: Dissolved Gas Analysis (DGA), Power Transformer Diagnostics, Machine Learning, Ensemble Learning, PCA, Optimization Algorithms, Predictive Maintenance, IEC TC10, Circle Methods.

Abstract (in French)

Le diagnostic et la prévention des défauts dans les transformateurs immergés dans l'huile sont essentiels pour garantir la fiabilité des systèmes électriques. Bien que l'analyse des gaz dissous (AGD) soit une méthode de diagnostic largement utilisée, les techniques traditionnelles rencontrent souvent des difficultés face aux zones de défauts qui se chevauchent. Cette thèse propose des solutions novatrices intégrant l'apprentissage automatique, l'apprentissage profond et les algorithmes d'optimisation pour surmonter ces limites. La recherche évalue les vecteurs d'entrée de la DGA et les combine avec des algorithmes basés sur des arbres pour améliorer la précision. Afin d'optimiser davantage les performances, l'analyse en composantes principales (ACP) est appliquée pour extraire les caractéristiques clés, qui sont ensuite combinées avec des techniques d'apprentissage ensembliste telles que le boosting, le bagging et le decorate. Ces approches sont systématiquement comparées aux techniques conventionnelles, démontrant leur supériorité en termes de précision diagnostique. La contribution novatrice de cette thèse est l'introduction des Méthodes du Cercle 1 et 2, qui sont proposées avec des axes dynamiques et des angles variables. Ces méthodes innovantes reposent sur la convergence des échantillons ayant des caractéristiques similaires autour de leurs centres et sur l'espacement de ces centres pour des échantillons présentant des caractéristiques différentes, exploitant des algorithmes d'optimisation pour trouver la meilleure distribution des zones de défauts. Validées sur des ensembles de données réels, y compris la base de données IEC TC10, les Méthodes du Cercle offrent une précision diagnostique supérieure par rapport aux approches existantes. Ces avancées améliorent significativement l'interprétabilité de la DGA et établissent un cadre robuste pour la maintenance prédictive et la prévention des défauts dans les transformateurs électriques.

Mots-clés : Analyse des Gaz Dissous (AGD), Diagnostic des Transformateurs, Apprentissage Automatique, Apprentissage ensembliste, ACP, Algorithmes d'Optimisation, Maintenance Prédictive, IEC TC10, Méthodes du Cercle.

Abstract (in Arabic)

الملخص

يُعد تشخيص الأعطال ومنعها في المحولات المغمورة بالزيت أمرًا بالغ الأهمية للحفاظ على موثوقية أنظمة الطاقة. وعلى الرغم من أن تحليل الغازات المذابة (DGA) هو طريقة تشخيصية مستخدمة على نطاق واسع، إلا أن التقنيات التقليدية غالبًا ما تواجه صعوبة في التعامل مع تداخل مناطق الأعطال. تقترح هذه الأطروحة حلولًا مبتكرة تجمع بين التعلم الآلي، والتعلم العميق، وخوارزميات التحسين لمعالجة هذه القيود. يتم تقييم متجهات مدخلات تحليل الغازات المذابة ودمجها مع خوارزميات قائمة على الأشجار لتحسين الدقة. ولتعزيز الأداء بشكل أكبر، يتم تطبيق تحليل المكونات الرئيسية (PCA) لاستخلاص الميزات الرئيسية التي يتم دمجها بعد ذلك مع تقنيات التعلم التجميعي مثل (Boosting)، (Bagging)، و(Decorate). يتم مقارنة هذه الأساليب منهجيًا مع التقنيات التقليدية، مما يظهر تفوقها في دقة التشخيص. تتمثل الإضافة الجديدة لهذه الأطروحة في تقديم طريقتي الدائرة 1و2، اللتين تقترحان استخدام محاور ديناميكية وزوايا متغيرة. تعتمد هذه الطرق المبتكرة على تقارب العينات ذات الخصائص المتشابهة حول مراكزها، وتباعد هذه المراكز للعينة ذات الخصائص المختلفة، مستغلة خوارزميات التحسين لتحديد أفضل توزيع لمناطق الأعطال. وقد تم التحقق من صحة هذه الطرق باستخدام مجموعات بيانات واقعية، بما في ذلك قاعدة بيانات IEC TC10، حيث أظهرت طرق الدائرة تفوقًا في دقة التشخيص مقارنة بالنهج الحالية. تسهم هذه التطورات بشكل كبير في تحسين فهم تحليل الغازات المذابة، وتوفر إطارًا قويًا للصيانة التنبؤية ومنع الأعطال في المحولات الكهربائية.

الكلمات المفتاحية: تحليل الغازات المذابة (DGA)، تشخيص المحولات الكهربائية، التعلم الآلي، التعلم التجميعي، تحليل المكونات الرئيسية (PCA)، خوارزميات التحسين، الصيانة التنبؤية، IEC TC10، طرق الدائرة.

Table of Contents

Chapter IV: Application and Evaluation of AI Technique for Diagnosis

Chapter V: New technique based on dynamic axes for DGA interpretation

Appendices

List of Figures

List of Tables

Nomenclature

Acronym	
RCM	Reliability Centered Maintenance
DGA	Dissolved Gas Analysis
IEC	International Electrotechnical Commission
PD	Partial Discharge
D1	Low Energy Discharges
D2	High Energy Discharges
T1	Faults (<300°C)
T2	Faults (300°C-700°C)
T3	Faults (>700°C)
FRA	Frequency Response Analysis
IFT	Interfacial Tension
IR	Insulation Resistance
DP	Degree Of Polymerization
DPC	Polarization And Depolarization Current
BDV	Breakdown Voltage
RVM	Recovery Voltage Measurement
HST	Hot Spot Temperature
GC	Gas Chromatography
CGE	Characteristic Gas Ensemble
TRT	Three Ratio Technique
DTM	Duval Triangle Method
DPM	Duval Pentagon Method
MPM	Mansour Pentagon Method
GTM	Gouda Triangle Method
DRM	Dornenburg Ratios Method
RFRM	Rogers' Four Ratios Method
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
ES	Expert Systems
EL	Ensemble learning

RF	Random Forest
TE	Tree Ensemble
GBT	Gradient Boosted Tree
PCA	Principal Component Analysis
MBC	Maximum Boost Control
LSTM	Long Short-Term Memory
DT	Decision Trees
SVM	Support Vector Machine
KNN	K Nearest Neighbor
RBFN	Radial Basis Function Networks
GWO	Gray Wolf Optimization
WOA	Whale Optimization Algorithm

Chapter I. Introduction

I.1 General overview

The electrical energy industry is typically divided into three functional sections: production, transmission, and distribution. Power transformers, tap changers, circuit breakers, measurement transformers, substation batteries, and connection equipment are the main devices in a transmission and distribution infrastructure, enabling the transfer of energy from power plants to consumers [1]. These devices are critical assets, and their failure can lead to electricity supply interruptions, personal and environmental risks, and high repair costs. Therefore, these critical assets must be closely and continuously monitored to assess their operational conditions and ensure maximum availability [2]. In particular, large oil-immersed power transformers are among the costliest assets in electricity transmission and distribution networks. In terms of investment, this equipment represents nearly 60% of the cost of a substation. Consequently, as major equipment in electrical systems, their proper functioning is essential for the efficient and reliable operation of power grids [3]. Given these facts and current economic constraints, electric utilities are therefore compelled to focus on power transformer maintenance. To address this issue, these companies have changed their maintenance strategies and the management of their transformer fleets, abandoning regular maintenance in favor of reliability-centered maintenance (RCM) or condition-based maintenance, with the goal of extending transformer service life. In other words, the objective is to maximize the service life of transformers with minimal maintenance while planning for their eventual shutdown [4].

The main goal of the condition-based maintenance strategy for transformers is to detect early signs of faults and track the aging process. In this approach, a diagnosis can only be made when the transformer shows an anomaly, which presents as symptoms. This is achieved through "off-line" techniques [5]. For continuous monitoring, "on-line" methods are used to determine whether the transformer is operating correctly, if there is a fault, how severe the fault is, and whether the transformer can remain in service. Reliable monitoring and diagnostic tools, therefore, provide an economical means of planning electrical grid maintenance [6].

In recent years, rapid changes and developments have been observed in the field of transformer condition monitoring and assessment. The performance and reliability of transformers can be significantly improved by using advanced on-line and off-line fault diagnostic systems.

Oil analysis, similar to blood tests for humans, particularly the analysis of dissolved gases, represents one of the most accurate and information-rich techniques. It can be considered an off-line diagnostic tool through detailed analysis and interpretation of the various gases, or on-line by triggering an alarm when the acceptable limits for the total or individual gas quantities are exceeded.

Today, thanks to advances in computer hardware and data mining software primarily based on artificial intelligence, in-depth analysis of various phenomena affecting the operation of transformers has become possible.

To develop a reliable and general model for assessing the condition of power transformers, several intelligent and hybrid algorithms can be considered to create a set of optimal models. Consequently, an intelligent approach focused on the development and improvement of preventive measures can be adopted, which can be used by electric utilities and energy

companies, regardless of the type and severity of the fault and the technical conditions of the transformer. However, it is important to note that the use of such models requires a thorough understanding of the algorithm's functioning and its integration with traditional methods to properly calibrate it and ensure the reliability of electrical systems.

I.2 Thesis objective

The objective of this thesis is to develop an intelligent information system for fault diagnosis and prediction in electrical power transformers. To achieve this, the potential of intelligent technologies, including classification and optimization algorithms, is utilized for fault classification, diagnosis, decision-making, and data preprocessing, with the goal of extracting valuable insights from raw data.

Furthermore, a novel technique based on dynamic axes has been developed, which addresses the longstanding debate surrounding traditional methods and effectively overcomes their limitations.

I.3 Thesis contribution

The research presented in this thesis has led to the development of an advanced tool for assessing the condition of power transformers. Several approaches have been proposed and implemented:

- A fault diagnosis algorithm for power transformers, based on dissolved gas analysis (DGA), has been developed using tree-based ensemble learning algorithms. This approach is integrated with various traditional method vectors, where the effect of each vector was analyzed and the most suitable one selected.
- A novel approach was introduced by combining machine learning algorithms with ensemble techniques. This method leverages the core information from DGA graphical method vectors and generates synthetic data using deep learning algorithms. This contribution is particularly valuable for researchers and engineers, addressing the issue of limited availability of open DGA datasets.
- A new graphical technique for interpreting dissolved gases and diagnosing faults in power transformers has been implemented. This method, based on dynamic axes in a circular form, merges field methods with intelligent optimization techniques, offering a more effective approach for fault interpretation.

I.4 Thesis outline

Chapter II presents an overview of power transformers, including fault classification and the causes of faults. It also provides a comprehensive description of the various techniques commonly employed for monitoring and diagnosing faults in oil-immersed transformers using dissolved gas analysis (DGA).

Chapter III introduces 22 feature vectors derived from traditional DGA methods, which serve as input vectors for four tree-based ensemble algorithms: random forest, tree ensemble, gradient boosted tree, and extreme gradient tree. The results indicate that the best performance is achieved using the 16-feature vector, composed of the gas ratios from the Rogers four-ratio method and the three-ratio technique. This represents a novel approach that combines traditional method input vectors with tree-based learning algorithms.

Chapter II. Literature review on internal faults diagnosis in power transformers

II.1 Introduction

In electrical system network power transformers represents the most important components. Regular maintenance through condition monitoring gives this equipment a longer life, which is essential for continuous power supply. Their failure and instability can cause serious financial and environmental consequences. Therefore, understand the faults causes are much recommended. These faults can be effectively prevented through condition assessment techniques such as monitoring and diagnosing methods.

This chapter introduces concepts about power transformers, fault classification and their causes. In addition, it provides an overview of the various techniques most commonly used in monitoring and diagnosing immersed oil transformer faults by using dissolved gas analysis (DGA).

II.2 Power Transformers

The electrical grid is primarily divided into two subsystems: transmission and distribution. The transmission system is responsible for the transfer of electrical energy over long distances, ensuring the efficient delivery of power from generation plants to substations. On the other hand, the distribution network delivers this energy at appropriate voltage levels to end consumers, including residential, commercial, and industrial locations [4, 7].

A critical component in both transmission and distribution systems is the power transformer. According to the International Electrotechnical Commission (IEC), a power transformer is a static device equipped with two or more windings [8]. Through the process of electromagnetic induction, it converts an alternating voltage and current system into another system, typically of different voltage and current values, while maintaining the same frequency. This transformation is essential for efficient power transmission, allowing energy to be delivered safely and reliably over vast distances and at various voltage levels suited to consumer needs [9].

II.2.1 Transformer types:

The International Electrotechnical Commission (IEC) classifies oil-immersed power transformers into three categories based on their apparent power capacity [10]:

- **Distribution transformer:** up to 2,500 kVA for three-phase and 833 kVA for single-phase transformers.
- **Medium power transformer:** ranging from 2.5 MVA to 100 MVA for three-phase and 33.3 MVA for single-phase transformers.
- **Large power transformer:** above 100 MVA for three-phase and above 33.3 MVA for single-phase transformers.

II.2.2 Structure:

Regardless of its category, a power transformer consists essentially of the following components [11]:

- **Active part:** Comprising the iron core, this channels the magnetic flux, and the low- and medium-voltage windings. Their role is to respectively produce a variable flux in the primary winding and create an induced voltage in the secondary winding. This assembly is housed inside a metallic tank.
- **Bushings:** These are used to connect the transformer windings to external power lines through the tank.
- **Tap changer:** Employed to adjust the voltage according to the load on the network by adding or removing winding turns.
- **Cooling system:** Its purpose is to dissipate heat and maintain the transformer at acceptable temperatures. The cooling method varies from one transformer to another, depending on power levels, client requirements, and usage conditions.
- **Insulation system:** Typically composed of oil and paper.
- **Various accessories:** These include thermometers, Buchholz relays, desiccators, expansion tanks, and more.

Figure II.1 showes the power transformer components. These include (1) the bushings and porcelain terminals, (2) the tap changer, (3) the magnetic core, (4) the cooling and protection system, (5) the insulation between windings and ground, (6) the tank/grounding system, (7) the winding connections, (8) the winding structure, (9) the winding geometry, (10) the insulation between turns, and (11) the expansion tank.

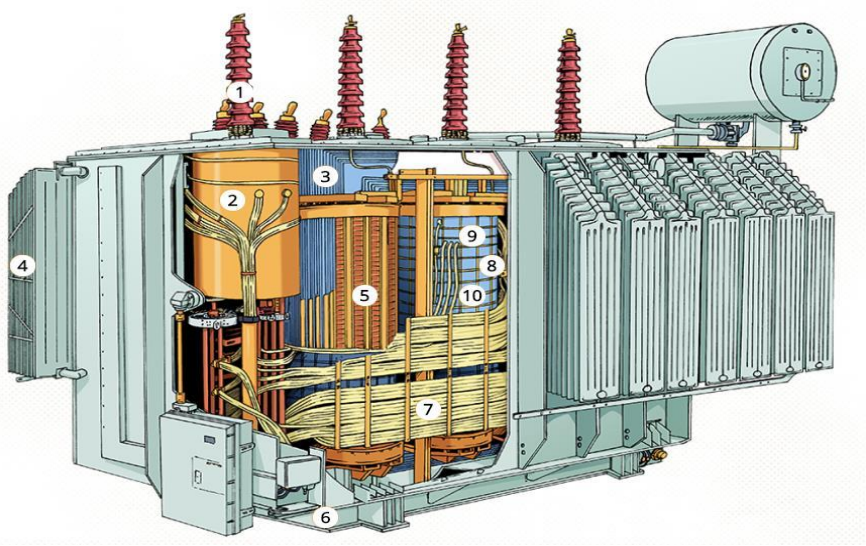


Figure II.1 Structural Overview of a Power Transformer [10]

II.3 Incipient faults classification in power transformers

Power transformers are subject to failure as a result of electrical, thermal, or mechanical stresses that arise during regular operation. Although the failure rate attributable to stress-related faults remains relatively low and stable throughout most of a transformer's operational lifespan, these faults warrant attention, as their underlying causes can often be identified prior to a complete failure. The following sections outline several key fault causes investigated in this study.

II.3.1 Electrical faults:

Current flows between points that are normally insulated from each other can generate electrical faults. Also, electrical faults can result from short circuits between laminated windings and loose connections. Electrical faults are further classified as high energy discharges or arcs (D2) and low energy discharges (D1). The classification is based on the energy content

associated with them. Electrical discharges may puncture solid insulation and form carbon particles in liquid insulation. As the energy associated with discharges increases, gas alarms may be raised and equipment may fail [12].

II.3.2 Partial discharge:

Partial discharge (PD) occurs when the insulation between conductors only partially bonds. Corona, a type of PD, typically happens in the gaseous medium around conductors distant from solid or liquid insulation. While the ionization from PD does not create a low-resistance path for high current flow, it can still affect transformer performance, often going undetected until a fault occurs. PD typically becomes problematic at several thousand volts and can increase noise levels in sensitive equipment. Detection methods exist but are primarily qualitative and provide limited information for setting acceptable limits, as the complexity varies with voltage levels and insulation types [13].

II.3.3 Thermal faults:

Thermal stress, such as overheating, can cause hot spots in insulation, leading to thermal faults, which are classified by temperature: T1 faults (<300°C), T2 faults (300°C-700°C), and T3 faults (>700°C). Bad connections, circulating currents, or leakage flux can also cause localized overheating. These thermal faults raise the transformer liquid temperature and generate more fault gases compared to electrical faults over the same duration. The type and amount of gases produced vary with temperature, making them useful for diagnosing transformer faults [14].

II.4 Comprehensive review of condition monitoring methods

Maintenance and monitoring of transformers have important roles in the energy and system management, where the condition of transformers plays an effective role in the transmission of electricity and energy, and any malfunction in the transformer would affect the entire network [15]. Among the methods for evaluating the condition of a transformer, frequency response analysis (FRA), or as it is called, the transfer function method, is an effective method for identifying mechanical faults, including winding deformation [16-18]. PD analysis is also used for monitor power transformer. PD method is a discharge of electrical insulation, whether liquid or solid, where early detection of partial discharge reduces aging and extends the life of the transformer [19-21].Alongside, these two methods we find vibration analysis. Vibration is usually the result of a mechanical movement, and this explains the use of vibration analysis to detect mechanical faults such as looseness, imbalance, misalignment [22, 23].This method is not limited to mechanical systems only, but in recent years vibration analysis is used to monitor power transformer [24]. Interfacial tension (IFT) is other method. It is a chemical analysis test for the detection of acidic contaminants in oil as a result of the oxidative decomposition of paper [25, 26].Insulation resistance (IR) is to evaluate the insulation deterioration in the windings, based on the insulation resistance test, where the Polarization Index helps in reading the insulation resistance information [27].Thermography analysis is a significant method using a thermal imaging camera to detect hot spot temperature (HST) in infrared. Red and white colors represent hot areas and blue and black colors represent cold areas [28-30].Degree of polymerization (DP) is a useful way to increase the reliability of paper insulation, as the mechanical strength of the paper depends on the cellulose particles and their good bonding [31, 32].Polarization and Depolarization Current (PDC) is a time analysis for evaluating high-voltage insulation as it has the ability to determine the moisture value of oil and paper insulation and reduce deterioration [33, 34].Breakdown voltage (BDV) is as per IEC 60156 standard that the dielectric strength of the oil is determined by measuring the breakdown voltage and is conducted at room temperatures [35, 36].Recovery voltage measurement (RVM) is considered

an indicator of the deterioration of the insulation, and although it is an old method, it is an effective method in determining the condition of the oil insulation [37-39].Voltage and current measurements are a method that depends on the voltage difference and the correlation between the incoming and outgoing current. To know the extent of the different types and severity of faults, simulation, experimentation and the online detection must be done [40-42].

Finally, dissolved gas analysis method (DGA) is the best technique in determining the state of a transformer over the ages, owing to its significance in diagnosing faults in oil-filled transformers and is considered a continuous study because it always seeks to develop methods for interpreting the method of DGA [43], whether by traditional methods, new and old, or by using artificial intelligence (AI). Figure II.2 presents methods for evaluating power transformers and methods for interpreting DGA in particular.

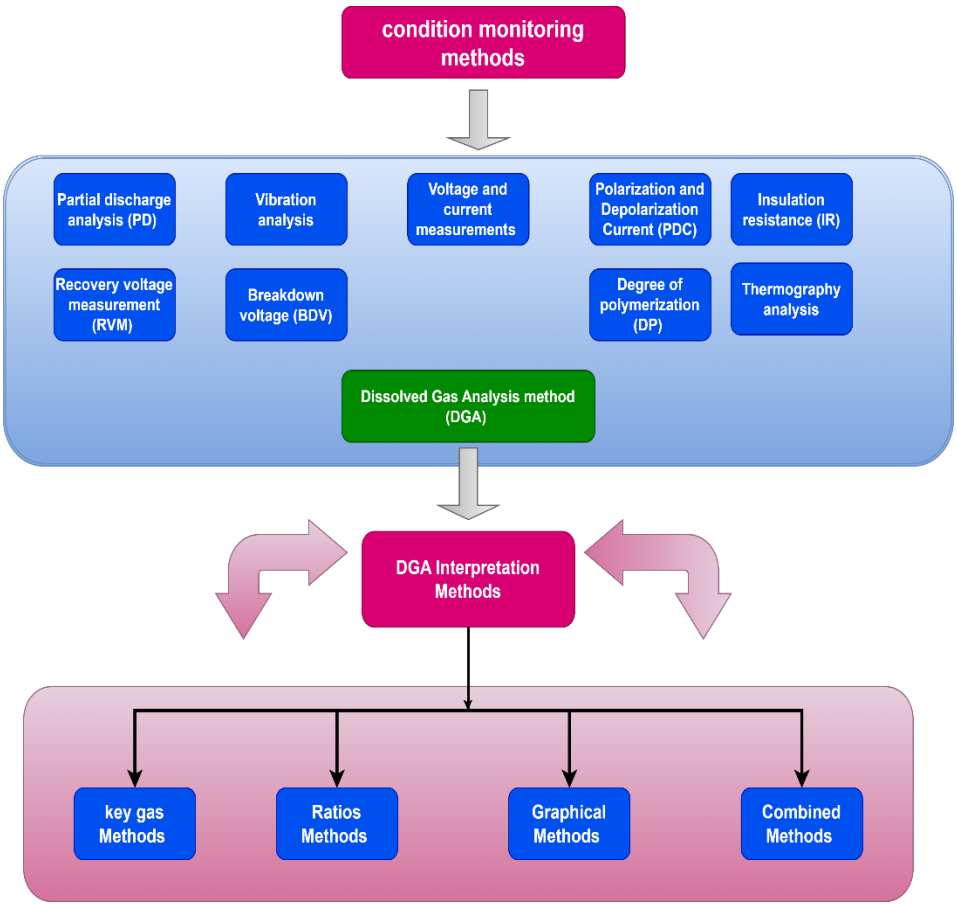


Figure II.2 Power transformer condition assessment and DGA interpretation methods

II.5 Dissolved gas analysis methods

DGA is a critical diagnostic tool for detecting incipient faults in oil-filled power transformers [44]. DGA is commonly employed to analyze the gases dissolved in insulating oil, which are generated due to thermal and electrical stresses affecting the operational integrity of transformers [45, 46]. The primary gases analyzed in DGA include hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), and acetylene (C_2H_2). In cases of cellulose degradation, carbon monoxide (CO) and carbon dioxide (CO_2) are also significant indicators [47, 48].

The Buchholz relay, developed in 1928 [49], was an early innovation aimed at protecting equipment by responding to gas formation due to faults. As shown in Figure II.3, the relay's operation relies on detecting gas accumulation in a pipeline connected to an expansion tank.

The gas flow triggers the relay, which then activates an alarm or disconnects the transformer to prevent damage.

By the late 1950s, advancements in analytical techniques allowed for the sampling and analysis of gases and liquids, leading to the development of a field chromatography analyser by H.H. Wagner in 1959 [50]. Gas chromatography (GC) method is widely used for identifying and quantifying dissolved gases in transformer oil, particularly those produced by electrical faults [51].

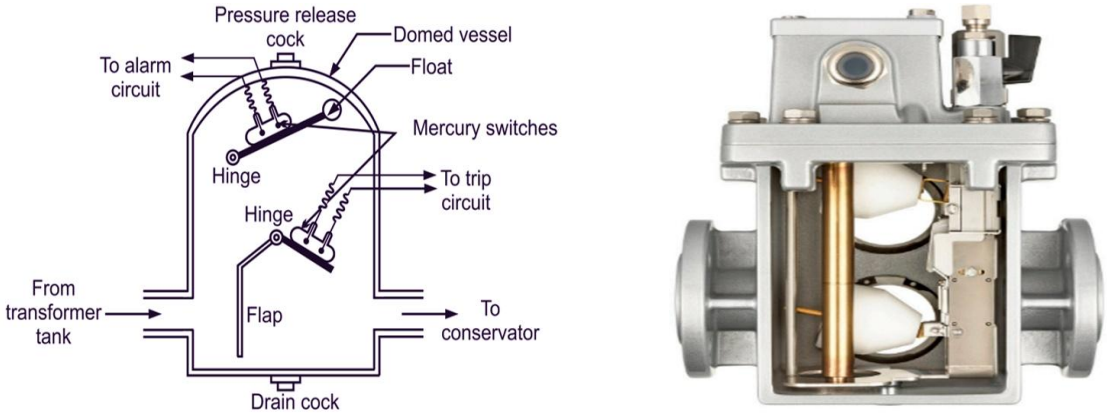


Figure II.3 Buchholz relay, (a) electrical schematic; (b) an example of a Buchholz relay

There are several steps in the method of chromatographic analysis of dissolved gases in oil. Sampling is the first step, which can be done using flexible aluminium bottles or a syringe. Gases extracted from the oil sample can also be temporarily stored in the Buchholz relay [52]. Following sample collection, the gases are extracted and analyzed using a gas chromatograph, which separates the various gases, allowing for their identification and quantification [53]. Figure II.4 delineates the systematic process of gas chromatography as applied in DGA, encompassing sample collection, gas extraction, and analytical separation and identification of dissolved gases. This methodical approach ensures accurate diagnostics, enabling timely maintenance and prevention of transformer failures.

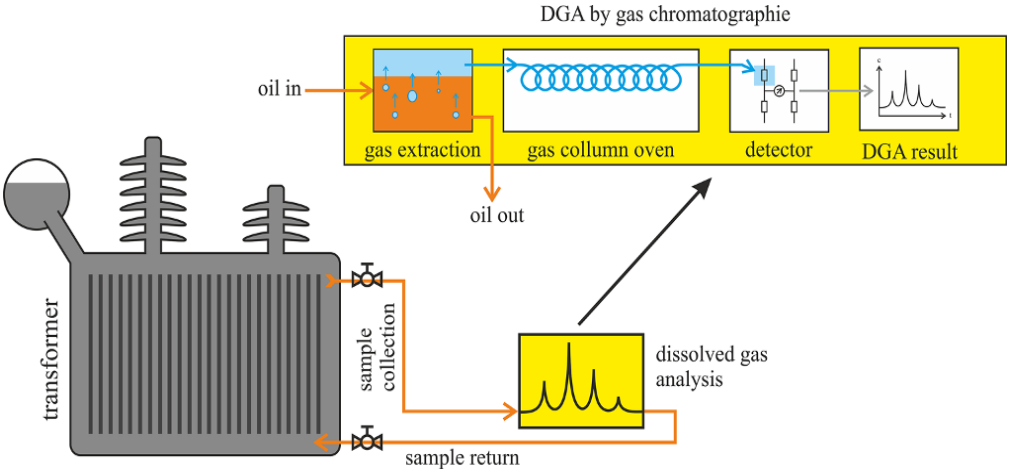


Figure II.4 Steps for gas chromatography analysis [54]

II.6 The most used DGA methods

This section provides an overview of the different types of DGA methods, highlighting those that have gained prominence in recent years due to their reliability, accuracy, and widespread adoption in the industry. By focusing on the most widely used methods, this discussion aims to shed light on the current trends and best practices in transformer diagnostics.

II.6.1 Key gas Methods:

II.6.1.1 IEEE key gas:

The IEEE Key Gas method, originally conceptualized in 1973 by the Doble Engineering Company, forms the basis for assessing power transformer conditions through the analysis of gas composition resulting from thermal and electrical stresses. This diagnostic technique was first introduced by David Pugh in 1974 [55], and its refined version was later incorporated into the Standard IEEE C57.104. This method employs the concentrations of specific gases, including hydrogen, hydrocarbons, and carbon monoxide, generated during the decomposition of both oil and paper. Table II.1 outlines the principal gases utilized alongside their associated faults, while Table II.2 details the gas ratios corresponding to the 4 main fault types identified through this method. The percentages referenced in Table II.1 and Table II.2 are based on the empirical observations of numerous experts in the field [56]. However, the IEEE Key Gas method is not without limitations; it can sometimes produce ambiguous or incorrect diagnoses. Additionally, in some cases, the dominant gas detected is not one of the primary gases considered by this method, and carbon monoxide is mistakenly interpreted as an indication that paper insulation is involved in the malfunction [57].

Table II.1 Key gas generated according to a fault type [56]

Key gas	Fault type	Typical proportions
C_2H_4	Overheating, oil	Mainly C_2H_4; Smaller proportions of C_2H_6, CH_4, and H_2; Traces of C_2H_2 at very high fault temperature
CO	Overheating, paper and oil	Mainly CO; Much smaller quantities of hydrocarbon gases (Predominantly C_2H_4 with smaller proportions of C_2H_6, CH_4, and H_2)
H_2	Partial discharge	Mainly H_2; Small quantities of CH_4; Traces of C_2H_4 and C_2H_6
H_2& C_2H_2	Arcing	Mainly H_2 and C_2H_2; Minor traces of CH_4, C_2H_4, and C_2H_6; Also, CO if cellulose is involved

Table II.2 Key gas proportions of fault types of IEEE key gas method [56]

	Key gas percentages (%)					
Fault type	**H_2**	**CH_4**	**C_2H_6**	**C_2H_4**	**C_2H_2**	**CO**
Partial discharges	85	13	1	1	-	-
Arcing	60	5	2	3	30	-
Overheating, oil	2	16	19	63	-	-
Overheating, paper	-	-	-	-	-	92

II.6.1.2 CSUS Key gas:

The CSUS Key Gas method, also referred to as the Key Gas Quantity method, was developed through a collaboration between California State University in Sacramento, and the Pacific Gas and Electric Company [58]. This diagnostic technique involves directly and individually comparing the concentrations of hydrocarbon and carbonaceous gas against their respective threshold values. As shown in Table II.3, standard and non-standard limit values are provided for each gas.

Table II.3 Determination of the type of fault according to the CSUS method [58]

	Limit concentrations		
Key gas	**Normal**	**Anormal**	**Fault type**
H_2	150	1000	Corona discharges, arcing
CH_4	25	80	Corona discharges
C_2H_6	10	35	Local overheating
C_2H_4	20	150	Severe overheating
C_2H_2	15	70	arcing
CO	500	1000	Overload, paper decomposition
CO_2	10 000	15 000	Overload, paper decomposition

If all recorded concentrations fall lower the specified normal values, the transformer is deemed to be in satisfactory condition. Fault types are identified through a direct, individual comparison of gas concentrations to these limits. However, despite its straightforward application, the CSUS Key Gas method suffers from low diagnostic accuracy and is often susceptible to diagnosing multiple faults simultaneously, which is a notable limitation.

II.6.1.3 CIGRE Key gas:

Direct comparison is based on the CIGRE key gas method between measured concentrations of key gases and their permissible levels. This approach, suggested in CIGRE Task Force 15.01.01 [59], extends beyond the three primary fault categories by also enabling the detection of faults related to paper degradation. Table II.4 presents the typical concentration thresholds for key gases, along with the specific malfunctions associated with exceeding these levels. Although this method provides valuable diagnostic insights, it is rarely used in isolation for transformer condition assessment. Instead, the CIGRE TF 15.01.01 guideline recommends incorporating gas ratios into the interpretation process. Similar to the CSUS method, the CIGRE Key Gas method's primary limitation lies in its tendency to identify multiple faults simultaneously, which can complicate accurate diagnosis.

II.6.1.4 Fallou Key gas:

The 1975 publication by CIGRE [60] introduced a notable gas detection technique based on data on the behavior of gas in mineral oils in electrical and thermal fault conditions. This method, developed through experiments conducted by Fallou, focuses on determining the concentration in percentage of each gas produced during these conditions. Specifically, the

approach considers carbon dioxide and hydrocarbons containing three carbon atoms as key indicators, as outlined in Table II.5.

Table II.4 Determination of the type of fault according to the CIGRE method [59]

Key gas	Limit	Fault
C_2H_2	> 20	Arcing
H_2	> 100	Partial discharges
$\sum C_xH_y$	> 1 000	Thermal fault
$\sum CO_x$	> 10 000	Cellulosic degradation

Table II.5 Key gas method proposed by Fallou [60]

	Key gas percentage (%)							
Fault type	**H_2**	**CH_4**	**C_2H_6**	**C_2H_4**	**C_2H_2**	**C_3H_x**	**CO**	**CO_2**
Low-energy discharge (PD1)	88	7	2	-	-	1	1	1
High-energy discharge (PD2)	55	7	5	6	15	6	1	6
Arcing	39	10	-	6	35	3	4	2
Overheating, oil 300°C	-	37	13	19	-	31	-	-
Overheating, oil 500°C	17	25	8	25	-	23	-	-
Overheating, oil 800°C	16	16	6	41	-	21	-	-
Overheating, paper 300°C	26	1	-	-	-	-	-	73
Overheating, paper 500°C	6	1	-	-	-	-	33	59
Overheating, paper 800°C	9	8	1	4	-	1	50	25

II.6.1.5 Characteristic gas ensemble:

The latest advancement in key gas analysis was presented in 2018 [61] by researchers Davidenko and Ovchinnikov. This advancement is the Characteristic Gas Ensemble (CGE) method, which evaluates the relative concentrations of the main gases through the following equation:

$$c_r^i = C^i / C_{lim}^i \qquad \text{(II.1)}$$

Where c_r^i , C^i and C_{lim}^i denote relative, observed, and permissible concentration of the ith key gas.

CGE method is implemented through a three-step process. Initially, the relative concentrations of various key gases are determined. Next, these relative concentrations are used to develop a letter-code system for the key gases, which is based on the significance of each gas

within the ensemble. The meanings of the letters in this coding system are clarified in Table II.6. Finally, in this step, identify the faults using the letter codes specified in Table II.7.

Table II.6 Alphanumeric coding of major gases [61]

Letter	Fault type
A	Primary gas with the highest relative concentration
B	Gas with the second-highest relative concentration $c_r^i \geq 1$
C	Gas at the second or third level with $c_r^i < 1$
D	Corresponds to all other gases not categorized under the previous letters

Table II.7 Determine the fault of the method of gas ensembles characteristic [61]

	Key gas				
Fault type	**H_2**	**CH_4**	**C_2H_6**	**C_2H_4**	**C_2H_2**
Overheating, T < 300°C	C, D	B	A	C, D	D
Overheating, 300°C < T < 700°C	D	A	C	B	D
Overheating, T > 700°C	D	B	C	A	D
Partial discharges	A	C	D	D	D
Discharges of low energy	A	C	D	D	B
Discharges of high energy	B	C	D	D	A
A set of faults with electrical fault propagation	D	B	D	C	A
A set of faults with thermal fault propagation	C, D	A	B	D	C, D

II.6.2 Ratios Methods DGA:

II.6.2.1 Dornenburg ratios method:

Introduced in 1974, the Dornenburg ratios method become one of the earliest approaches developed for analyzing dissolved gases [62]. This method facilitates the analysis of three types of faults: thermal decay, corona, and arc Table II.8. For the method to be considered valid, the concentration of at least one of the major gases hydrogen (H_2), methane (CH_4), ethylene (C_2H_4), or acetylene (C_2H_2) must be at least a concentration of it has twice the number limit Table II.9. In addition, this method generates cases where ethane (C_2H_3) or carbon monoxide (CO) concentrations exceed prescribed threshold values [63].

Table II.8 Fault classification by Dornenburg Ratio Method [62]

Fault	$R_1 = \frac{CH_4}{H_2}$	$R_2 = \frac{C_2H_2}{C_2H_4}$	$R_3 = \frac{C_2H_4}{C_2H_6}$	$R_4 = \frac{C_2H_2}{CH_4}$
Thermal decomposition	>1.0	<0.75	<0.3	>0.4
Corona	<0.1	/	<0.3	>0.4
Arcing	0.1–1.0	> 0.75	>0.3	<0.4

Table II.9 Permitted limit for Dornenburg Ratio Method [62]

Gas	H_2	CH_4	C_2H_6	C_2H_2	C_2H_4	CO
Limit (ppm)	100	120	65	50	1	350

II.6.2.2 The four ratios method of Rogers:

The Rogers' Four Ratio Method (RFRM) [64], as the name suggests, detects transformer faults by analyzing four specific gas ratios: methane-hydrogen (CH_4/H_2), ethane-methane (C_2H_6/CH_4), ethylene-ethane (C_2H_4/C_2H_3), and from acetylene to ethylene (C_2H_2/C_2H_4). These parameters are determined by considering them in defined areas and assigning rules ranging from zero to five, as detailed in Table II.10. This method allows the identification of 12 possible faults, which are shown in Table II.11 [65].

Table II.10 Rogers Ratio Code [64]

Gas Ratio	Ranges	Code
$R_1 = \frac{CH_4}{H_2}$	< 0.1	5
	0.1- 1	0
	1--3	1
	> 3	2
$R_2 = \frac{C_2H_2}{C_2H_4}$	< 1	0
	> 1	1
$R_3 = \frac{C_2H_4}{C_2H_6}$	< 1	0
	1-3	1
	> 3	2
$R_4 = \frac{C_2H_6}{CH_4}$	< 0.5	0
	0.5-3	1
	> 3	2

Table II.11 Fault Types as stated in Rogers Ratio Method [64]

No	CH_4/H_2	C_2H_6/CH_4	C_2H_4/C_2H_6	C_2H_2/C_2H_4	Fault type
1	0	0	0	0	No fault
2	1-2	0	0	0	<150 ºC thermal fault
3	1-2	1	0	0	150–200 ºC thermal fault
4	0	1	0	0	200–300 ºC thermal fault
5	0	0	1	0	General conductor overheating
6	1	0	1	0	Winding circulating currents
7	1	0	2	0	Core and tank circulating currents overheated joints
8	5	0	0	0	Partial discharge
9	5	0	0	1-2	Partial discharge with tracking
10	0	0	0	1	Flashover without power follow through
11	0	0	1-2	1-2	Arc with power follow- through
12	0	0	2	2	Continuous sparking to floating potential

II.6.2.3 IEC Ratio Method:

The IEC Ratio Method is one of the methods used to interpret DGA [66]. It is similar to the four-step Rogers method but omits the ratio of ethane to methane (C_2H_3/CH_4). Instead, conventional gas concentration limits and their corresponding codes for three basic gaseous parameters are used for fault diagnosis: acetylene and ethylene (C_2H_2/C_2H_4), methane and hydrogen (CH_4/H_2), and ethylene and ethane (C_2H_4/C_2H_2) are given in Table II.12, while Table II.13 provides a detailed description of each fault based on these ratio codes.

Table II.12 IEC code [66]

Ratio Codes	Range	Code
R1 = CH_4/H_2	$R1<0.1$	1
	$0.1 \leq R1 \leq 1$	0
	$R1>1$	2
R2 = C_2H_2/C_2H_4	$R2<0.1$	0
	$0.1 \leq R2 \leq 3$	1
	$R2>3$	2
R3 = C_2H_4/C_2H_6`	$R3<1$	0

	$1 \leq R3 \leq 3$	1
	$R3 > 3$	2

Table II.13 Fault diagnosis using IEC code [66]

N	Fault type	code		
1	No Fault	0	0	0
2	Partial discharge with low energy density	0	1	0
3	Partial discharge with high energy density	1	1	0
4	discharge of low energy	1.2	0	1.2
5	discharge of high energy	1	0	2
6	Overheating T<150 °C	0	1	0
7	Overheating 150<T<300 °C	2	1	0
8	Overheating 300≤ T≤ 700 °C	2	1	0
9	Overheating ≥ 700 °C	2	2	2

II.6.2.4 CIGRE Ratio Method:

The CIGRE ratio method [59] proposed by CIGRE Working Group 15.01.01 uses five specific gas ratios acetylene-ethane (C_2H_2/C_2H_3), hydrogen-methane (H_2/CH_4), ethylene-ethane (C_2H_4/C_2H_6), acetylene- hydrogen (C_2H_2/H_2), and carbon dioxide to carbon monoxide (CO_2/CO) to detect transformer failures, including electrical faults, thermal problems, and solid insulation problems. These ratios are outlined in detail in Table II.14.

Table II.14 Determination of the type of fault according to the CIGRE ratio method [67]

Ratio	limit	Fault type
C_2H_2/C_2H_6	≥ 1.0	Arcing
H_2/CH_4	≥ 10.0	Partial discharges
C_2H_4/C_2H_6	≥ 1.0	Thermal fault
C_2H_2/H_2	≥ 2.0	Discharges in OLTC
CO_2/CO	≥ 10.0	Overheating, paper
	< 3.0	Cellulose decomposition due to electrical fault

When all the gas ratios fall within the specified limits provided in Table II.14 [67], the transformer is deemed to be in normal condition. CIGRE ratio method is often used in conjunction with the CIGRE Key Gas Method, with their combined identification framework described in Table II.15 [67]. The CIGRE methods employ a letter-coding system to classify

gas concentrations and ratios. "K1" indicates that all key gases are below the concentration limits in Table II.4, while "K2" shows that at least one gas exceeds these limits. For gas ratios, "R1" means all values are within the limits in Table II.14, and "R2" signals that at least one ratio surpasses these limits.

Table II.15 identification procedures using CIGRE ratios and major gas methods [67]

Ratios	Key gases	Explanation
R1	K1	No action, transformer most probably healthy
R2	K2	Transformer most probably faulty, additional analyses needed
R2	K1	Possible incipient failure, additional analyses needed
R1	K2	Possibility of more than one failure, further investigations needed

II.6.2.5 SOU-N EE 46.501: 2006 method:

The SOU-N EE 46.501:2006 method, established as a standard in 2007, serves as a guideline for conducting DGA within the Ukrainian power engineering sector [68]. This method allows for the identification of various fault conditions through the analysis of gas ratio ranges, as detailed in Table II.16 [68].

Table II.16 Identification scheme according to SOU-N EE 46.501:2006 [68]

Fault type	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
No fault	no significant	0.1 – 1.0	< 0.2
PD	no significant	< 0.1	< 0.2
D1	> 1.0	0.1 – 0.5	> 1.0
D2	> 1.0	0.1 – 1.0	> 2.0
D3	< 1.0	0.3 – 0.5	> 5.0
T12	Not significant	> 1.0	< 1.0
T2	Not significant	> 1.0	1.0 – 4.0
T3	< 0.2	> 1.0	> 4.0

In addition to detecting 6 fundamental faults defined by the IEC, the method also identifies the normal operating condition and creeping discharges (D3), which are discharges occurring on the surface and within the solid insulation. Thise ratios used in this method are based on those initially proposed in the Russian Standard RD 153-34.0-46.302-00, introduced in 1989 and later revised in the 2001 standard. The diagnostic criteria employed in the SOU-N EE 46.501:2006 method align with the ratio-based approach of the RRM/IEC methods, offering nine diagnostic outcomes. However, the specific ratio ranges for these diagnoses differ from those used in the RRM/IEC approach.

II.6.2.6 ETRA ratios method:

ETRA Ratios Method employs three specific gas ratios, which are found from concentrations of ethane (C_2H_6), ethylene (C_2H_4), and acetylene (C_2H_2), to diagnose the six fundamental faults identified by the IEC. The coding system and the suggested diagnoses associated with this method are detailed in Table II.17 and Table II.18, as referenced in [69]. This approach provides a streamlined framework for interpreting DGA results in transformer fault diagnostics.

Table II.17 Gas ratio coding using ETRA ratio method [69]

Ratio	Range	Code
$R_1 = C_2H_2/C_2H_4$	$R_1 \leq 0.01$	0
	$0.01 < R_1 < 0.2$	1
	$R_1 \geq 0.2$	2
$R_2 = C_2H_2/C_2H_6$	$R_2 \leq 0.01$	0
	$0.01 < R_2 < 1.0$	1
	$1.0 \leq R_2 < 10.0$	2
	$R_2 \geq 10.0$	3
$R_3 = C_2H_4/C_2H_6$	$R_3 < 1.0$	0
	$1.0 \leq R_3 < 4.0$	1
	$4.0 \leq R_3 < 10.0$	2
	$R_3 \geq 10.0$	3

Table II.18 Reference explanation using ETRA ratio method [69]

Fault type	R_1	R_2	R_3
Thermal fault, $T > 700°C$	0	0/1	2/3
Thermal fault, $T < 300°C$	0	0	0
Thermal fault, $300°C < T < 700°C$	0	0/1	1
Partial discharge	2	1	0/1/2
Low-energy discharge	2	2	0/1/2
High-energy discharge	2	2	3
High-energy discharge	2	3	/

II.6.2.7 Hyosun cooperation ratio:

The Hyosung Corporation Ratios Method [70], introduced in 2013 by a group of Korean researchers from HYOSUNG Company with the scientist Duval, represents a significant advancement in transformer fault diagnostics. The method employs 6 specific gas ratios, as outlined in Table II.19, to identify the six fundamental IEC faults.

Table II.19 Gas ratios used in the HYOSUN Corporation gas ratio method [70]

Ratio	R1	R2	R3	R4	R5	R6
Expression	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_2}{CH_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_4}{C_2H_6}$	$\frac{C_2H_4}{CH_4}$

The core concept of this technique involves using these gas ratios to distinguish between various types of transformer faults. The diagnostic process begins with the use of two gas ratios (R1 and R2) to distinguish thermal faults and electrical faults. Once this distinction is made, each fault group can be analyzed independently. For thermal faults, the T1 faults are separated from other thermal faults by gas ratios R5 and R6. When T1 faults are detected, the method uses gas ratio pairs (R2, R5) to differentiate between T2 and T3 faults. In power faults, the same two ratios (R2, R5) are used to differentiate partial discharge (PD), D1, and D2 faults. The overall diagnostic process is visually illustrated in the flow chart shown in Figure II.5 [70].

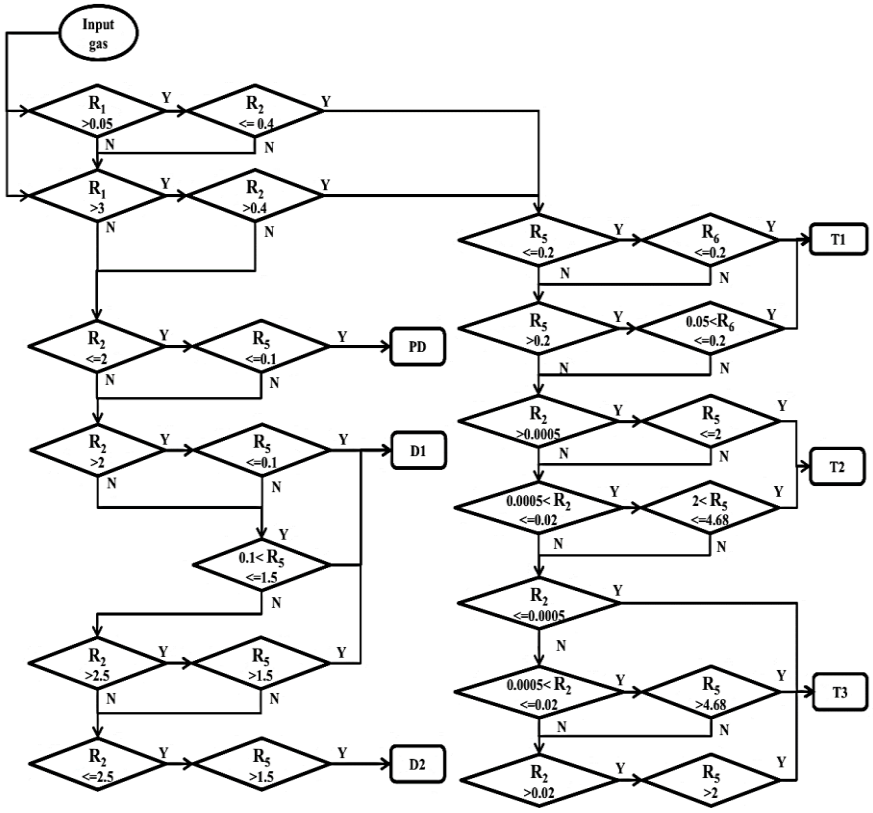


Figure II.5 Flowchart of gas ratio method of HYOSUN Corporation [70]

II.6.2.8 Three Ratios Techniques:

Gouda et al have developed an enhanced diagnostic method called Three-Ratio Technique (TRT) [71], based on previous gas analysis interpretation methods. In this method, three new gas parameters are used that are specifically designed to improve the fault detection and magnitude analysis, as detailed in Table II.20 [72].

Table II.20 TRT codes [72]

Ratio Codes	Range	Code
$R_1 = \frac{C_2H_2}{C_2H_4}$	R1 < 0.05	0
	0.05 ≤ R1 ≤ 0.9	1
	R1 > 0.9	2
$R_2 = \frac{(C_2H_6 + C_2H_4)}{(H_2 + C_2H_2)}$	R2 < 1	0
	1 ≤ R2 ≤ 3.5	1
	R1 > 3.5	2
$R_3 = \frac{(CH_4 + C_2H_2)}{C_2H_4}$	R3 < 0.05	0
	0.05 ≤ R3 ≤ 0.5	1
	R3 > 0.5	2

A first-order fault detection method based on different rules is presented in Table II.21 [72]. Indicate the faults and and severity. The TRT method works on the principle that if the concentration of one gas is significantly higher than the other, it still remains within normal limits [72].The normal limits and conditions are shown in Table II.22 [72]. This method provides an analytical framework an impressive increase in reliability and fault detection accuracy in power transformers.

Table II.21 Identify fault with TRT [72]

Fault	type	R1	R1	R3
T > 700◦C	T3	1 or 2	0	0 or 1
300◦C < T < 700◦C	T2	1 or 2	1	0 or 1
150◦C < T < 300◦C	T1	1 or 2	2	0 or 1
Low temperature thermal T < 150◦C	T0	1	/	1
Low partial discharge	PD1	0	1 or 2	0 or 1
High partial discharge	PD1	0	1 or 2	2
High arcingdischarge	D2	0 or 1	0 or 1	2
Low arcingdischarge	D1	1 or 2	2	2
Mix of electrical and thermal fault	DT	2	0 or 1	2

Table II.22 Limit of dissolved gases concentrations for the TRT method application [72]

Gas	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	CO	CO_2
Limt ppm	100	120	65	50	1	350	2500

II.6.3 graphical methods:

II.6.3.1 Duval triangle method:

DTM is the first alternative graphic method for ratios methods to interpret the analysis of dissolved gases. This method was proposed by the Canadian scientist Michel Duval in 1974 [73]. This method allows the use of three ratios R1, R2, R3. These ratios represent a point in a graph in the form of a triangle, as shown in Figure II.6. A triangle is divided into seven zones, each zone representing the faults as follows: PD, D1, D2, DT, T1, T2 and T3. The fault is diagnosed based on the region to which the point belongs [74]. Michel Duval also developed triangle 1 in identifying faults for other cases, adding six Duval triangles [75], including triangle 4 and triangle 5, where triangle 4 is for low temperatures and triangle 5 is for high temperatures [76]. As shown in Figure II.7.

$$R1 = 100 * CH4/CH4 + C2H4 + C2H2 \quad \text{(II.2)}$$

$$R2 = 100 * C2H2/CH4 + C2H4 + C2H2 \quad \text{(II.3)}$$

$$R3 = 100 * C2H4/CH4 + C2H4 + C2H2 \quad \text{(II.4)}$$

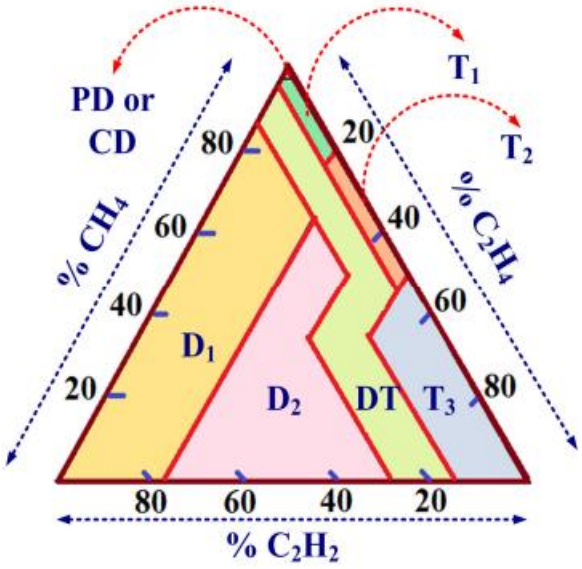


Figure II.6 Duval Triangle 1 [73]

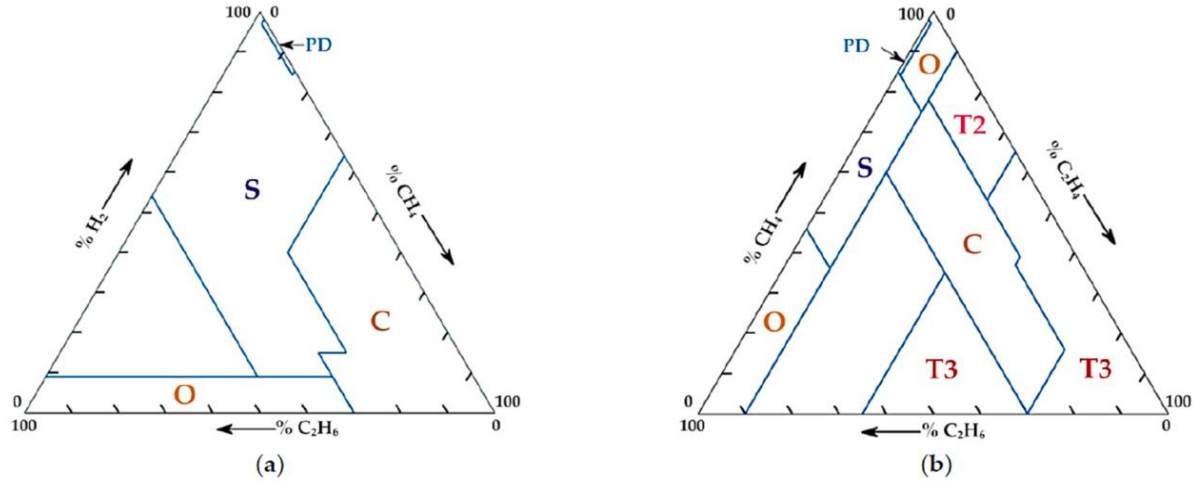


Figure II.7 (a) Duval 4 for a low-temperature fault (b) Duval 5 for a high-temperature fault [75]

II.6.3.2 Duval Pentagon Method:

The Duval Pentagon method developed by Duval and Lamar [77] provides a graphical method for interpreting the results of DGA see Figure II.8. This method uses a pentagon chart with five axes, each representing the relative concentrations of the five major combustion gases: hydrogen (H_2), methane (CH_4), ethane (C_2H_3), ethylene (C_2H_4), and acetylene (C_2H_4) of any axis from 0% to 100%. The Duval Pentagon shows six distinct zones corresponding to thermal and electrical failures, as well as a separate zone for stray gases. The method is supported by a set of equations, which are used to calculate the position of the pentagon, and facilitate the accurate interpretation of the DGA data.

$$\begin{cases} x_0 = \%H_2 \cos\left(\frac{\pi}{2}\right) \\ x_1 = \%C_2H_6 \cos\left(\frac{\pi}{2}+\frac{2\pi}{5}\right) \\ x_2 = \%CH_4 \cos\left(\frac{\pi}{2}+\frac{4\pi}{5}\right) \\ x_3 = \%C_2H_4 \cos\left(\frac{\pi}{2}+\frac{6\pi}{5}\right) \\ x_4 = \%C_2H_2 \cos\left(\frac{\pi}{2}+\frac{8\pi}{5}\right) \end{cases} \quad \begin{cases} y_0 = \%H_2 \sin\left(\frac{\pi}{2}\right) \\ y_1 = \%C_2H_6 \sin\left(\frac{\pi}{2}+\frac{2\pi}{5}\right) \\ y_2 = \%CH_4 \sin\left(\frac{\pi}{2}+\frac{4\pi}{5}\right) \\ y_3 = \%C_2H_4 \sin\left(\frac{\pi}{2}+\frac{6\pi}{5}\right) \\ y_4 = \%C_2H_2 \sin\left(\frac{\pi}{2}+\frac{8\pi}{5}\right) \end{cases} \quad \text{(II.5)}$$

$$\mathrm{A} = \frac{1}{2}\sum_{i=0}^{4}(x_i y_{i+1} - x_{i+1} y_i) \quad \text{(II.6)}$$

$$\begin{cases} X_C = \frac{\sum_{i=0}^{4}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{6A} \\ Y_C = \frac{\sum_{i=0}^{4}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{6A} \end{cases} \quad \text{(II.7)}$$

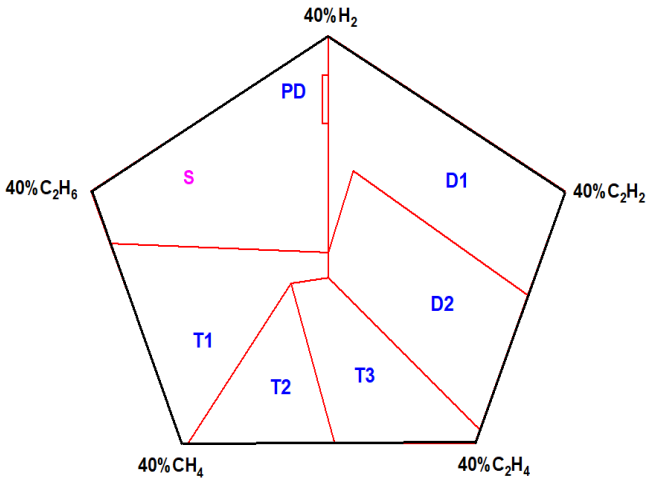


Figure II.8 Duval Pentagon Method

II.6.3.3 Mansour Pentagon Method:

In 2012, Mansour engaged the first graphical method in pentagon form for fault diagnosis [78]. This approach, further developed in subsequent work [79], involves constructing a pentagon where each vertex shows the percentage concentration of a specific gas relative to the total gases. Similar to the Duval Pentagon, the Mansour Pentagon identifies six distinct zones corresponding to various electrical and thermal faults, with the primary difference lying in the coding system used to classify these faults, as depicted in Figure II.9. The method is governed by a series of equations that define the relationships between the gas concentrations and their respective positions on the pentagon, facilitating accurate fault diagnosis.

$$s = \sum_{i=5}^{5} P_i \tag{II.8}$$

$$P_{\%i} = \frac{P_i}{s} * 100 \tag{II.9}$$

$$\begin{cases} x_m = \dfrac{\sum_{i=1}^{n} P_{\%i} x_i}{P_{\%i}} \\ y_m = \dfrac{\sum_{i=1}^{n} P_{\%i} y_i}{P_{\%i}} \end{cases} \tag{II.10}$$

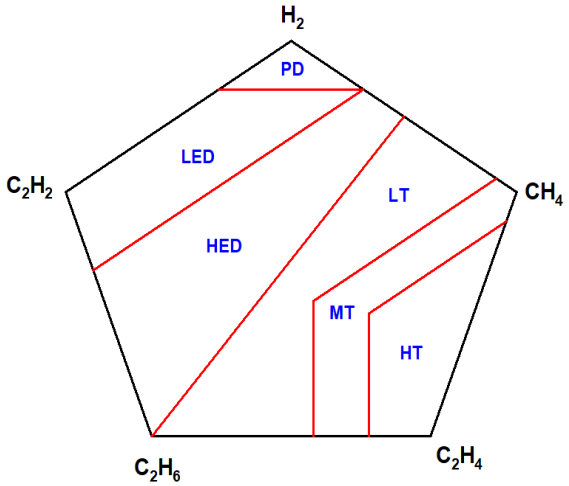


Figure II.9 Mansour Pentagon Method

II.6.3.4 Gouda Triangle Method:

The Gouda Triangle Method (GTM) is an advanced graphical diagnostic method introduced by Gouda and a team of scientists [80]. This method was developed as an improvement upon the Duval Triangle Method, addressing a key limitation of the Duval method its reliance on only three gases for fault detection. The Gouda method enhances fault diagnosis by incorporating two additional gases, ethane (C_2H_6) and hydrogen (H_2), into the analysis.In GTM, three ratios, R1, R2, and R3, are calculated and then converted into percentages, P1, P2, and P3. Each of these percentages represents a vertex of a triangle, forming a parallelogram structure. This approach is similar to the Duval method in its use of a triangular graphical representation and in defining fault zones, as illustrated in Figure II.10. The addition of more gases allows for a more comprehensive and accurate identification of transformer faults.

$$\begin{cases} R_1 = \dfrac{CH_4}{CH_4 + C_2H_4 + C_2H_6 + C_2H_2} \\ R_2 = \dfrac{C_2H_2}{H_2 + C_2H_6 + CH_4 + C_2H_2} \\ R_3 = \dfrac{C_2H_4}{H_2 + C_2H_6 + CH_4 + C_2H_2} \end{cases} \quad \text{(II.11)}$$

$$\begin{cases} P_1 = \dfrac{R_1 \times 100}{R_3 + R_2 + R_1} \\ P_2 = \dfrac{R_2 \times 100}{R_3 + R_2 + R_1} \\ P_3 = \dfrac{R_1 \times 100}{R_3 + R_2 + R_1} \end{cases} \quad \text{(II.12)}$$

With

$$\begin{cases} x = 100 - P2 - P3\cos\left(\dfrac{\pi}{6}\right)\cot\left(\dfrac{\pi}{3}\right) \\ y = P3\cos\left(\dfrac{\pi}{6}\right) \end{cases} \quad \text{(II.13)}$$

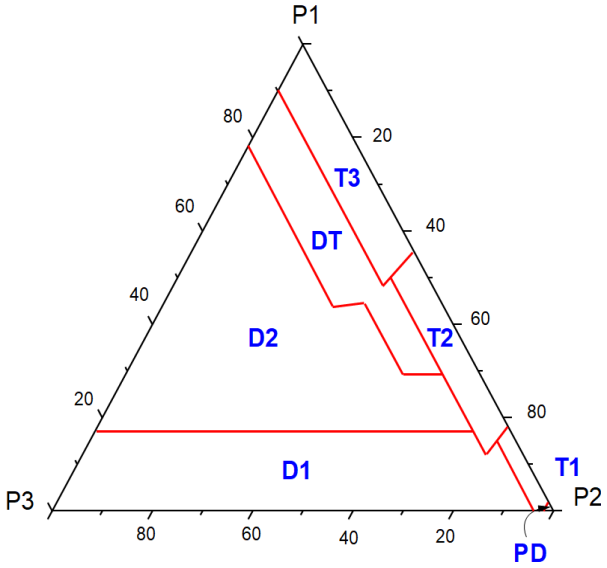


Figure II.10 Gouda Triangle method

II.6.3.5 Gouda heptagon method:

The Gouda Heptagon Method is a graphical diagnostic technique introduced by Gouda et al. in 2018 [81] for identifying faults in immersed electrical power transformers. This technique utilizes a heptagon, where each side represents the proportional percentage of one of the seven key gases: hydrogen (H_2), methane (CH_4), ethylene (C_2H_4), ethane (C_2H_6), acetylene (C_2H_2), carbon monoxide (CO), and carbon dioxide (CO_2). The percentages are weighted according to factors provided in Table II.23 [81].The Gouda Heptagon able for the diagnosis of 3 levels of cellulose degeneration high concentration (HCCD), medium concentration (MCCD), and low concentration (LCCD) as well as the six basic IEC faults and combinations of electrical and thermal faults. Figure II.11 illustrates how these degradation levels and fault types are identified using this method, offering a comprehensive tool for transformer fault analysis.

Table II.23 Main gas weighting factors in Gouda Heptagon Method [81]

Key gas	H_2	C_2H_4	C_2H_6	CH_4	CO_2	CO	C_2H_2
Weighting factor	3.5	7	5.3846	2.9167	0.14	1	116.6667

$$\mathbf{H_2} = \frac{\mathbf{H_2 \times 3.5 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.14)}$$

$$\mathbf{CH_4} = \frac{\mathbf{CH_4 \times 2.9167 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.15)}$$

$$\mathbf{C_2H_6} = \frac{\mathbf{C_2H_6 \times 5.3846 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.16)}$$

$$\mathbf{C_2H_4} = \frac{\mathbf{C_2H_4 \times 7 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.17)}$$

$$\mathbf{C_2H_2} = \frac{\mathbf{C_2H_2 \times 116.6667 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.18)}$$

$$\mathbf{CO_2} = \frac{\mathbf{CO_2 \times 0.14 \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.19)}$$

$$\mathbf{CO} = \frac{\mathbf{CO \times 100}}{\mathbf{(3.5 \times H_2) + (2.9167 \times CH_4) + (5.3846 \times C_2H_6) + (7 \times C_2H_4) + (116.6667 \times C_2H_2) + CO + (0.14 \times CO_2)}} \quad \text{(II.20)}$$

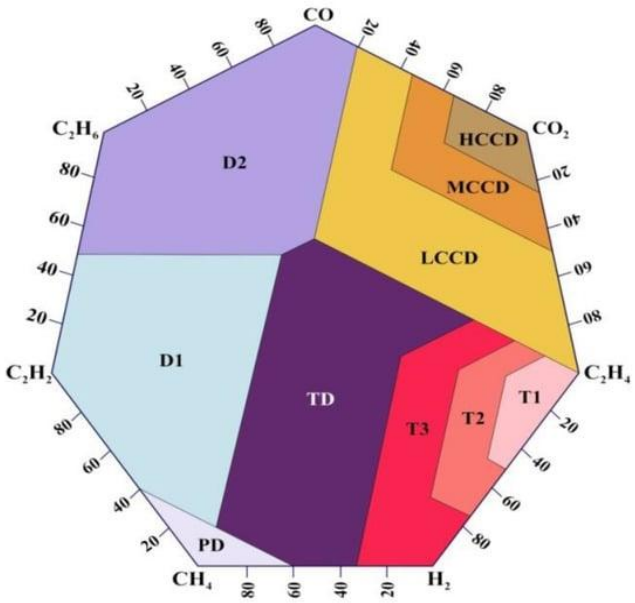


Figure II.11 Gouda's heptagon Method [81]

I.6.3.6 Cartesian graphical method:

The Cartesian graphical method [82], proposed by Fathallah et al., is the most recent advancement in DGA and is recognized as a cutting-edge approach for diagnosing transformer faults. This method utilizes the concentrations of four key gases: hydrogen (H_2), methane (CH_4), ethylene (C_2H_4), and acetylene (C_2H_2). The technique involves the calculation of three specific ratios that incorporate the concentration of hydrogen (H_2) and consider the impact of ethane (C_2H_6). The ratios are defined as follows:

$$P_1 = \frac{CH_4 + H_2}{2 \times H_2 + 2 \times CH_4 + C_2H_2 + C_2H_4} \quad \text{(II.21)}$$

$$P_2 = \frac{C_2H_2 + H_2}{2 \times H_2 + 2 \times CH_4 + C_2H_2 + C_2H_4} \quad \text{(II.22)}$$

$$P_3 = \frac{C_2H_4 + CH_4}{2 \times H_2 + 2 \times CH_4 + C_2H_2 + C_2H_4} \quad \text{(II.23)}$$

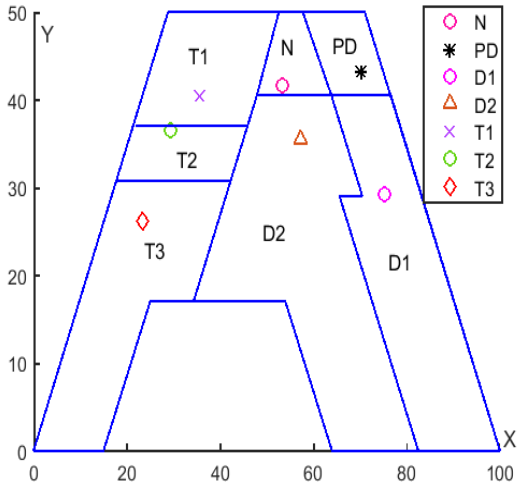


Figure II.12 Cartesian graphical Method [82]

These ratios are subsequently used to determine the Cartesian coordinates $(x,y)$for graphical representation. The coordinates are computed as follows:

$$x_{\text{point}} = (P_2 + P_1 \times \sin(30°)) \times L \quad \text{(II.24)}$$

$$y_{\text{point}} = (P_1 \times \sin(60°)) \times L \quad \text{(II.25)}$$

Where L is a scaling factor.

The resulting coordinates allow for the visualization of fault types within a Cartesian graph, as shown in Figure II.12. This graph is splited into several zones wich represents different fault categories, including partial discharge (PD), low-energy discharge (D1), high-energy discharge (D2), and thermal faults (T1, T2, T3), along with a normal condition (N) zone. This method provides a refined and precise means of diagnosing faults through a straightforward graphical approach.

II.6.4 Combined DGA methods:

II.6.4.1 Clustering method:

The two-step clustering method, introduced in 2016 by [83], provides a structured approach to diagnosing transformer faults through DGA. In the first step, a matching subset for a specific sample is identified by applying a set of criteria based on the relative concentrations of various fault related gases, as detailed in Table II.24. This initial selection of a subgroup offers critical insights into potential issues within the sample. In the next stage, gas ratios are employed to distinguish the "true" malfunction from other potential issues identified in the first step. Table II.25 lists the two key gas ratios used in this phase, along with the principles applied during the initial step to identify possible faults. The criteria for determining the precise fault among the possible faults is outlined in Table II.26. The entire diagnostic process is visually summarized in a flowchart, as shown in Figure II.13, which details the sequential steps involved in accurately identifying transformer faults using this method.

Table II.24 Types of faults for percentage limits of main gases [83]

Fault type	$\%C_2H_2$	$\%CH_4$	$\%C_2H_6$	$\%H_2$	$\%C_2H_4$
D_1	≤ 40	≤ 14.5	≤ 42	10 – 96	≤ 15
D_2	≤ 80	≤ 20	≤ 70	≤ 61	≤ 35
PD	≤ 2.5	≤ 18	≤ 66	30 – 98	≤ 13
T_1	≤ 4	≤ 80	≤ 100	≤ 50	≤ 40
T_2	≤ 2	≤ 83	4 – 90	≤ 25	10 – 70
T_3	≤ 12	≤ 50	≤ 20	≤ 35	40 – 100

Table II.25 Gas ratios of clustering method [83]

Ratio	Expression	Ratio	Expression
R1	C_2H_2/H_2	R5	C_2H_4/ CH_4
R2	C_2H_2/CH_4	R6	R_4+R_5
R3	C_2H_2/C_2H_6	R7	C_2H_4/C_2H_6
R4	C_2H_4/H_2		

Table II.26 Rule gases in the clustering method [83]

	Fault type					
	PD	D_1	D_2	T_1	T_2	T_3
Rule	$R_2 \leq 0.2$	$R_1 \leq 2$	$R_2 \geq 0.1$	$R_6 \leq 5$	$R_7 \leq 4$	
	$R_4 \leq 0.16$	$R_3 \geq 0.7$	$R_3 \geq 0.14$	$R_7 \leq 0.8$		
		$R_4 \leq 0.7$	$R_4 \leq 1.8$			

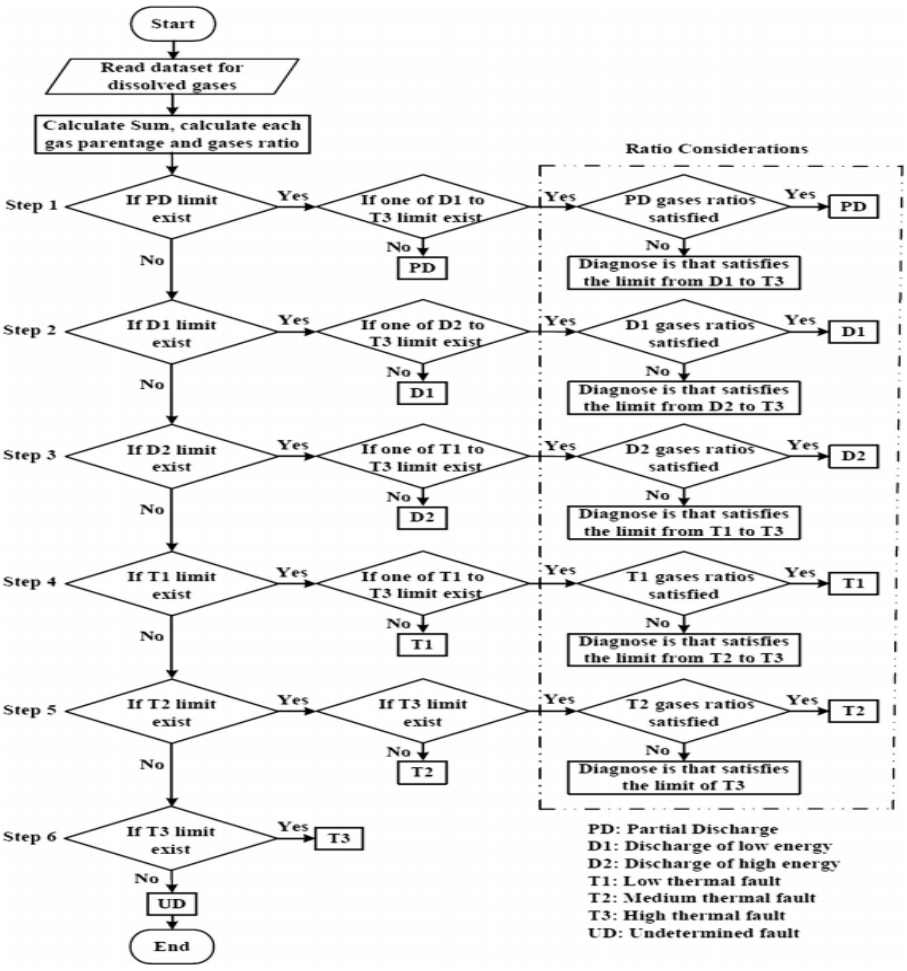


Figure II.13 Flowchart of clustering method [83]

II.6.4.2 Graphical method with two-shapes:

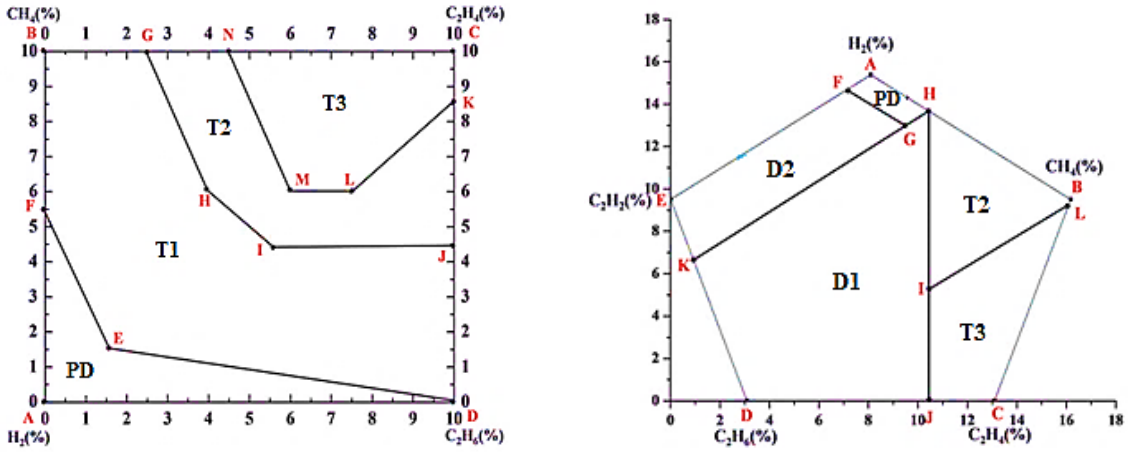
The graphical method with two-shapes [84] is an alternative approach proposed for discriminating between different fault types in transformers based on the percentage of acetylene derived from DGA in Figure II.14. This method utilizes two geometric shapes: a square for low thermal faults, where acetylene levels are negligible, and a pentagon for electrical discharge faults, where acetylene production is significant. The square shape primarily estimates low thermal faults by considering four gases (H_2, CH_4, C_2H_4, and C_2H_6), while the pentagon incorporates all five gases (H_2, CH_4, C_2H_4, C_2H_6, and C_2H_2) to diagnose electrical discharge faults and other fault types. The fault diagnosis is determined by locating the decision point using the center of mass for the relevant gas percentages, calculated through specific equations.

$$S = \sum_{i=1}^{5} (P_i) \qquad \text{(II.26)}$$

$$P_{\%i}\& = \frac{P_i}{S} * 100 \qquad \text{(II.27)}$$

$$x_m = \frac{1}{\sum_{i=1}^{n} P_{\%i}} \sum_{i=1}^{n} (P_{\%i} x_i) \qquad \text{(II.28)}$$

$$y_m = \frac{1}{\sum_{i=1}^{n} P_{\%i}} \sum_{i=1}^{n} (P_{\%i} y_i) \qquad \text{(II.29)}$$



A B

Figure II.14 Two-shapes graphical method [84]

II.6.4.3 The subset analysis method:

The subset analysis method, introduced by Nanfak et al. [85], presents a robust two-step approach for diagnosing power transformer faults. This method begins by grouping samples based on the maximum and minimum concentrations of fault-related gases. After this initial grouping, a specific diagnsis sub-model is implemented to each subset. The method leverages six gas ratios, derived from five key hydrocarbon gases present in transformer oil: hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), and acetylene (C_2H_2). By breaking down

the dataset into 75 distinct subsets, including those subsets dominated by hydrogen, this method facilitates the development of flexible and highly targeted diagnostic models. Figure II.15 and Table II.27 provide a detailed overview of the implementation process, as well as the gas ratios employed for fault discrimination. This subset analysis approach allows for a more nuanced and precise diagnosis of transformer faults by tailoring the diagnostic model to the specific characteristics of each subset.

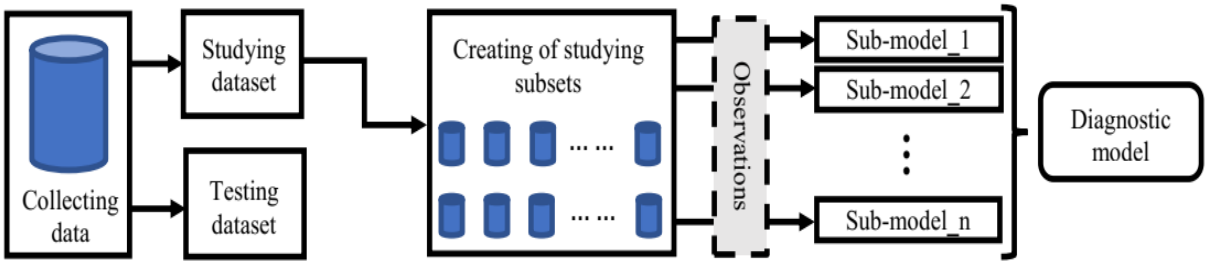


Figure II.15 Schematic view of the subset analysis method [85]

Table II.27 Gas ratios used in the subset analysis method [85]

Ratio	Expression
R_1	$(CH_4 + C_2H_6)/ (H_2 + CH_4 + C_2H_6 + C_2H_4 + C_2H_2)$
R_2	$(CH_4 + C_2H_4)/ (H_2 + CH_4 + C_2H_6 + C_2H_4 + C_2H_2)$
R_3	$C_2H_6/(CH_4 + C_2H_4)$
R_4	$(CH_4 + H_2)/ (H_2 + CH_4 + C_2H_6 + C_2H_4 + C_2H_2)$
R_5	$(C_2H_4 + C_2H_2)/ (H_2 + CH_4 + C_2H_6 + C_2H_4 + C_2H_2)$
R_6	C_2H_2/C_2H_4

II.6.4.4 Combined technique Ward:

The Combined Technique No. 1, introduced by Ward et al. in 2021, represents a state-of-the-art approach in transformer fault diagnostics. This method synergizes the strengths of three established DGA-based diagnostic techniques to develop a highest accurate and reliable model [86]. The combined technique integrates the diagnostic results obtained from the Duval Triangle Method, the modified Rogers Ratios Method (RRM) as adapted by the Central Electricity Generating Board (CEGB) [87], and the modified RRM according to IEC standards [87].The process of integrating these methods is systematically illustrated in the flowchart presented in Figure II.16. This flowchart outlines the sequential steps for processing the diagnostic outputs from each of the three methods to produce a consolidated and precise diagnosis. The results from the DTM, modified RRM/CEGB and RRM/IEC are combined, ensuring a comprehensive analysis that leverages the advantages of each technique, ultimately enhancing the accuracy of transformer fault detection.

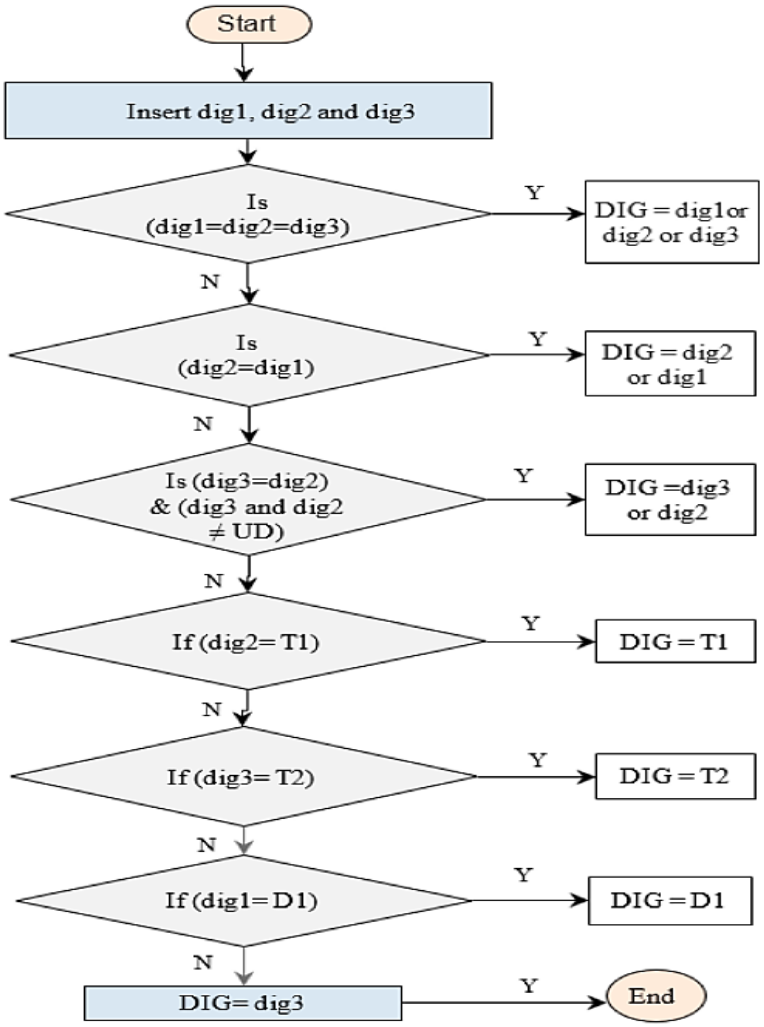


Figure II.16 Flowchart of Combined technique Ward [86]

II.7 Conclusion

The precise diagnostic of primary faults in power transformers is critical for developing a maintenance strategy that aligns with the equipment's operational condition. Conventional diagnostic methods, while widely utilized by transformer maintenance professionals, exhibit certain limitations as highlighted in the literature reviewed in this chapter. The analysis presented is structured into four techniques: key gas methods, gas ratio methods, graphical methods, and Combined methods . For each type, both the strengths and limitations of these techniques are critically examined, providing a comprehensive overview of their applicability and effectiveness in transformer fault diagnosis.

Chapter III. Tree ensemble learning and feature analysis for fault diagnosis

III.1 Introduction

Following the historical development of DGA, it is evident that all interpretation models developed over the years, as discussed in detail in the previous chapter, have contributed in some way to improving transformer diagnostics. However, none of these models can be used to make definitive decisions, as they do not always provide consistent and appropriate interpretations. This issue arises from the complexity of the gas generation process, which leads to complex combinations of gases, as well as from the uncertainty in DGA data due to sampling procedures and uncontrollable chromatographic analysis conditions.

As a result, no defined relationships exist between the characteristic quantity of faults and fault assumptions, making fault diagnosis more challenging. To address the complex decision-making issues, overcome the limitations of traditional approaches, and extract additional information from DGA data, various techniques have recently been developed, typically involving artificial intelligence (AI) methods.

In this chapter, traditional methods are combined with AI techniques to discuss the effect of feature input vector on the performance of smart power transformer fault diagnosis methods based on DGA.

III.2 A comprehensive overview of the use of artificial intelligence technique in diagnosing power transformers

The use of AI in diagnosing DGA for power transformers has become a crucial tool in modern power systems management. DGA is a primary method for assessing the condition of power transformers, as it helps detect faults by analyzing the gases dissolved in the transformer's insulating oil. Here is a comprehensive overview of how AI techniques are employed in diagnosing DGA power transformers:

III.2.1 AI Techniques in DGA Diagnosis

The integration of AI helps address the limitations of conventional DGA analysis, such as subjective interpretation and low accuracy in complex fault scenarios. Key AI techniques include:

III.2.1.1 Machine Learning (ML)

Machine learning algorithms, such as neural network (NN) [88, 89], support vector machine (SVM) [90], k-nearest neighbor (KNN) [91], and decision tree (DT) [92], have been widely studied for DGA diagnosis. Research suggests that machine learning can effectively classify fault types based on gas concentration patterns, improving accuracy compared to traditional methods. In addition, it provides interpretable models that can highlight important features in the data, making it useful for diagnosis.

III.2.1.2 Deep Learning (DL)

Deep learning techniques, particularly Convolutional Neural Networks (CNNs) [93] and Recurrent Neural Networks (RNNs) [94], are increasingly applied to DGA diagnosis. Studies show that CNNs can automatically extract features from raw data, leading to higher classification accuracy in fault detection. RNNs, on the other hand, are effective in analyzing time-series data, capturing temporal dependencies in gas concentration changes over time.

III.2.1.3 Expert Systems (ES)

Expert systems utilize knowledge-based approaches to mimic human reasoning in fault diagnosis. Research on ES in DGA has demonstrated their ability to combine expert knowledge with data analysis, providing reliable fault diagnosis and recommendations for maintenance actions. Algorithms such as Rule-Based Systems [95] and Fuzzy Logic [96] have been implemented to enhance decision-making under uncertainty.

III.2.1.4 Ensemble learning (EL)

Ensemble learning combines multiple models to improve predictive performance. Techniques such as Random Forest [97] and Gradient Boosting [98] have shown significant improvements in accuracy for DGA diagnosis. Studies indicate that these methods reduce overfitting and enhance robustness against noisy data by aggregating predictions from several base learners, thereby improving overall diagnostic capabilities.

III.3 Tree-based ensemble learning algorithms

III.3.1 Random Forest

Introduced by Breiman in 2001 [99], the Random Forest (RF) algorithm is an ensemble method rooted in the bagging technique. It is widely regarded as one of the most robust and well-known ensemble algorithms due to its efficiency and straightforward implementation [100, 101]. In the RF model, each tree within the forest is generated independently using a random vector derived from the input data and follows the same distribution as the other trees. The final predictions of the forest are obtained by averaging outputs through bootstrap aggregation and selecting features randomly [102]. The classification prediction in Random Forest can be defined as:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \qquad \text{(III.1)}$$

Where:

T is the total number of trees.

$h_t(x)$ is the prediction of the $t\text{-}th$ tree for input x.

The final prediction $\hat{y}$ is obtained by averaging the predictions of all trees (for regression) or by majority voting (for classification).

III.3.2 Tree ensemble

The Tree Ensemble (TE) algorithm operates on the principle that combining multiple decision trees yields better results than relying on a single tree [103]. By utilizing both rows (records) and columns (attributes) to construct each tree, the algorithm creates a robust and effective ensemble model [104]. During the partitioning process, a Gini Index is often employed to measure impurity:

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \qquad \text{(III.2)}$$

Where:

p_i is the probability of a sample belonging to class i.

The Gini Index is calculated for each split and helps select the best split, similar to the RF algorithm [105].

The final prediction is derived by combining the outputs of individual trees, just as in Random Forest:

$$\hat{y} = \frac{1}{M} \sum_{m=1}^{M} f_m(x) \qquad \text{(III.3)}$$

Where:

M is the total number of trees.
$f_m(x)$ is the prediction of the $m\text{-}th$ tree.

III.3.3 Gradient boosted tree

Gradient Boosted Tree (GBT) is an algorithm rooted in the boosting technique, where multiple weak learners are trained sequentially to minimize the errors of previous decision trees and enhance the overall performance of the trained Tree Ensemble (TE). The process begins by focusing on reducing bias to minimize error [106]. Unlike traditional boosting, GBT reduces the residual error progressively in each iteration, with each new model adjusted to address the gradient of the residuals from the prior model [107]. The general form of the model is:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \qquad \text{(III.4)}$$

Where:

$F_m(x)$ is the ensemble model at step mmm.

$h_m(x)$ is the $m\text{-}th$ weak learner (DT).

γ_m is the learning rate or step size that controls the contribution of each learner.

The model is updated iteratively in the direction of the negative gradient to minimize the loss function $L(y, \hat{y})$:

$$L(y, \hat{y}) = \sum_{i=1}^{n} \left(y_i - F_m(x_i)\right)^2 \qquad \text{(III.5)}$$

III.3.4 XGBoost Tree Ensemble

XGBoost is a parallel tree-boosting framework engineered for exceptional efficiency and flexibility [108]. It is capable of constructing a powerful learner while significantly reducing computational costs and mitigating overfitting [109]. Unlike GBT, which typically employs the first-order Taylor expansion, XGBoost enhances the loss function by incorporating the second-order Taylor expansion. Additionally, it utilizes a regularized objective function to further simplify the method and prevent overfitting [110]. The objective function in XGBoost can be written as:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \qquad \text{(III.6)}$$

Additionally, XGBoost uses the second-order Taylor expansion for optimizing the objective:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[g_i f(x_i) + \frac{1}{2} h_i f^2(x_i) \right] + \Omega(f) \qquad \text{(III.7)}$$

Where:

g_i and h_i are the first and second derivatives of the loss function w.r.t. the predictions.

III.4 Input features vectors

III.4.1 Gas standard methods

To facilitate the learning process, the following input feature vectors are utilized:

- **Vector 1:** X = [H_2, C_2H_4, C_2H_6, CH_4, C_2H_2]

This feature vector comprises the concentrations of five gases, measured in parts per million (ppm). The original measured data remains unaltered.

- **Vector 2:** X = [%H_2, %C_2H_4, %C_2H_6, %CH_4, %C_2H_2]

To mitigate the impact of outliers and address gas overlap, the data has been normalized into percentages. The components of this feature vector are calculated using the following equation [61]:

$$X(i) = \frac{C_i}{\sum_{j=1}^{5} C_j} \qquad \text{(III.8)}$$

Where C_1 to C_5 represent the concentrations (in ppm) of H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6, respectively.

- **Vector 3:** X = [R1, R2, R3, R4, R5]

This vector represents the relative concentrations of critical gases derived from characteristic gas ensemble methods [61]. The components are calculated as follows:

$$\begin{cases} R_1 = H_2 / H_{2\lim} \\ R_2 = CH_4 / CH_{4\lim} \\ R_3 = C_2H_2 / C_2H_{2\lim} \\ R_4 = C_2H_4 / C_2H_{4\lim} \\ R_5 = C_2H_6 / C_2H_{6\lim} \end{cases} \quad \text{(III.9)}$$

- **Vector 4:** X = [R_1, R_2, R_3, R_4, R_5]

The components of this feature vector are derived using the following equations:

$$\begin{cases} R_1 = H_2 / C_{\max} \\ R_2 = CH_4 / C_{\max} \\ R_3 = C_2H_2 / C_{\max} \\ R_4 = C_2H_4 / C_{\max} \\ R_5 = C_2H_6 / C_{\max} \end{cases} \quad \text{(III.10)}$$

Here, $C_{\max} = \max(H_2, CH_4, C_2H_2, C_2H_4, C_2H_6)$ and R_1 to R_5 are relative concentrations of key gases are from the Dekyoken method [111].

III.4.2 Gas ratios methods

- **Vector 5:** X = [R_1, R_2, R_3, R_4]

The components of this feature vector are derived using the following equations:

$$\begin{cases} R_1 = CH_4 / H_2 \\ R_2 = C_2H_2 / C_2H_4 \\ R_3 = C_2H_6 / C_2H_2 \\ R_4 = C_2H_2 / CH_4 \end{cases} \quad \text{(III.11)}$$

R_1 to R_4 are from the Dornenburg ratios method DRM [112].

- **Vector 6:** X = [R_1, R_2, R_3]

The components of this feature vector are derived using the following equations:

$$\begin{cases} R_1 = CH_4 / H_2 \\ R_2 = C_2H_2 / C_2H_4 \\ R_3 = C_2H_4 / C_2H_6 \end{cases} \quad \text{(III.12)}$$

R_1 to R_3 are from the Rogers/IEC ratios method [113].

- **Vector 7:** X = [R_1, R_2, R_3, R_4]

The components of this feature vector are derived using the following equations:

$$\begin{cases} R_1 = CH_4 / H_2 \\ R_2 = C_2H_2 / C_2H_4 \\ R_3 = C_2H_4 / C_2H_6 \\ R_4 = C_2H_6 / CH_4 \end{cases} \quad \text{(III.13)}$$

R_1 to R_4 are from Rogers' four ratios method RFRM [64].

- **Vector 8:** X = [R_1, R_2, R_3]

The components of this feature vector are derived using the following equations:

$$\begin{cases} R_1 = C_2H_2/C_2H_4 \\ R_2 = (C_2H_6 + C_2H_4)/(H_2 + C_2H_2) \\ R_3 = (CH_4 + C_2H_2)/C_2H_4 \end{cases} \quad \text{(III.14)}$$

R_1 to R_3 are from the three ratios techniques TRT [16].

III.4.3 Graphical methods

- **Vector 9:** X = [*x*, *y*]

The components of this feature vector are derived using the following equations:

$$\begin{cases} x = 100 - \%C_2H_2 - \%CH_4 \cos\left(\frac{\pi}{6}\right)\cot\left(\frac{\pi}{3}\right) \\ y = \%CH_4 \cos\left(\frac{\pi}{6}\right) \end{cases} \quad \text{(III.15)}$$

With

$$\begin{cases} \%CH_4 = CH_4/(CH_4 + C_2H_4 + C_2H_2) \\ \%C_2H_4 = C_2H_4/(CH_4 + C_2H_4 + C_2H_2) \\ \%C_2H_2 = C_2H_2/(CH_4 + C_2H_4 + C_2H_2) \end{cases} \quad \text{(III.16)}$$

This feature vector is based on the Duval triangle method DTM [114].

- **Vector 10:** X = [*x*, *y*]

The components of this feature vector are derived using the following equations:

$$\begin{cases} x = 100 - \%R_2 - \%R_3 \cos\left(\frac{\pi}{6}\right)\cot\left(\frac{\pi}{3}\right) \\ y = \%R_3 \cos\left(\frac{\pi}{6}\right) \end{cases} \quad \text{(III.17)}$$

With

$$\begin{cases} R_1 = \dfrac{CH_4}{(CH_4 + C_2H_6 + C_2H_4 + C_2H_2)} \\ R_2 = \dfrac{C_2H_2}{(H_2 + CH_4 + C_2H_6 + C_2H_4)} \\ R_1 = \dfrac{C_2H_4}{(H_2 + CH_4 + C_2H_6 + C_2H_2)} \end{cases} \quad \begin{cases} \%R_1 = \dfrac{100R_1}{(R_1 + R_2 + R_3)} \\ \%R_2 = \dfrac{100R_2}{(R_1 + R_2 + R_3)} \\ \%R_3 = \dfrac{100R_3}{(R_1 + R_2 + R_3)} \end{cases} \quad \text{(III.18)}$$

This feature vector is based on the Gouda' triangle method GTM [80].

- **Vector 11:** X = [X(1), X(2)]

The components of this feature vector are derived using the following equations:

$$\begin{cases} X(1) = \dfrac{1}{6}\dfrac{\sum_{i=0}^{4}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{\frac{1}{2}\sum_{i=0}^{4}(x_i y_{i+1} - x_{i+1} y_i)} \\ X(2) = \dfrac{1}{6}\dfrac{\sum_{i=0}^{4}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{\frac{1}{2}\sum_{i=0}^{4}(x_i y_{i+1} - x_{i+1} y_i)} \end{cases} \quad \text{(III.19)}$$

With

$$\begin{cases} x_0 = \%H_2 \cos\left(\dfrac{\pi}{2}\right) \\ x_1 = \%C_2H_6 \cos\left(\dfrac{\pi}{2} + \dfrac{2\pi}{5}\right) \\ x_2 = \%CH_4 \cos\left(\dfrac{\pi}{2} + \dfrac{4\pi}{5}\right) \\ x_3 = \%C_2H_4 \cos\left(\dfrac{\pi}{2} + \dfrac{6\pi}{5}\right) \\ x_4 = \%C_2H_2 \cos\left(\dfrac{\pi}{2} + \dfrac{8\pi}{5}\right) \end{cases} \quad \text{(III.20)}$$

The parameter y_i can be found by replacing the cosine with the sine in the x_i expressions. This feature vector is based on the Duval pentagon method DPM [77].

- **Vector 12:** X = [x_m, y_m]

The components of this feature vector are derived using the following equations:

$$\begin{cases} x_m = \dfrac{\sum_{i=1}^{n} m_i x_i}{100} \\ y_m = \dfrac{\sum_{i=1}^{n} m_i y_i}{100} \end{cases} \quad \text{(III.21)}$$

Each of ($x_i y_i$) represents the coordinates of the vertices of the Pentagon and m_i are the concentrations of all Pentagon gases in percentage,n thousand is the number of five combustible gases. This feature vector is based on the Mansour pentagon method MPM [79].

III.4.4 Compilation of gas ratios methods

- **Vector 13:** DRM combined with RFRM

The Dornenburg ratios method and Rogers' four ratios method are integrated into a single input vector by concatenating the corresponding vectors derived from each method. This results in Vector 13, which is formed by joining Vector 5 and Vector 7.

- **Vector 14:** DRM combined with TRT

Vector 14 is created by combining the Dornenburg ratios method from Vector 5 with the three ratios technique from Vector 8, effectively linking the two vectors together.

- **Vector 15:** IEC combined with TRT

To form Vector 15, the IEC method represented by Vector 6 is combined with the three ratios technique using Vector 8. These vectors are merged to generate the input.

- **Vector 16:** TRT combined with RFRM

Vector 16 is a combination of the three ratios technique method and Rogers' four ratios method using Vector 7 and Vector 8.

III.4.5 Compilation of graphical methods

- **Vector 17:** DTM combined with GTM

The Duval Triangle Method (Vector 9) and Gouda's Triangle Method (Vector 10) are combined into a single input vector, Vector 17, by concatenating the respective vectors.

- **Vector 18:** DPM combined with MPM

To form Vector 18, the Duval Pentagon Method (Vector 11) is merged with the Mansour Pentagon Method (Vector 12) by linking the two vectors.

- **Vector 19:** DTM combined with DPM

Vector 19 is created by connecting the two Duval Methods, one corresponding to the triangle (Vector 9) and the other to the pentagon (Vector 11), resulting in a combined vector.

- **Vector 20:** GTM combined with MPM

In this case, Vector 20 is formed by merging Gouda's Triangle Method (Vector 10) with the Mansour Pentagon Method (Vector 12).

- **Vector 21:** DTM combined with MPM

Vector 21 represents the combination of the Duval Triangle Method (Vector 9) with the Mansour Pentagon Method (Vector 12).

- **Vector 22:** GTM combined with DPM

Finally, Vector 22 is the result of combining Gouda's Triangle Method (Vector 10) with the Duval Pentagon Method (Vector 11).

III.5 Statistical Significance Tests in Non-Parametric Analysis

III.5.1 Friedman Test

The Friedman test [115] is a non-parametric statistical test used to detect differences in treatments across multiple attempts in repeated measures experiments. It is particularly useful when the assumptions of normality required by parametric tests, such as the repeated-measures ANOVA, are not met. The test ranks the data for each subject or experimental unit across different conditions and compares the mean ranks to determine if there is a significant difference between them.

III.5.1.1 Hypotheses:

- **Null Hypothesis (H_0):** There are no statistically significant differences between the conditions or treatments (i.e., the median ranks are the same across groups).
- **Alternative Hypothesis (H_1):** There is a statistically significant difference between at least one conditions compared to the others.

II.5.1.2 Test Procedure:

- **Rank the Data:** Each subject's data is ranked across the conditions from lowest to highest.
- **Sum Ranks for Each Condition:** The ranks for each condition are summed across all subjects.
- **Calculate the Friedman Test Statistic:** The test statistic is computed from the rank sums and evaluated using a chi-square distribution with $k-1$ degrees of freedom, where k is the number of conditions.
- **Interpretation:** If the test statistic exceeds the critical value (or the p-value is below a predetermined significance level, usually 0.05), the null hypothesis is rejected, indicating that at least one condition differs from the others.

The Friedman test does not specify where the differences lie; it only indicates that differences exist among the conditions.

III.5.1.3 Use Cases:

The test is applicable in experiments where:

- The same subjects are measured across different conditions.
- The data is ordinal or continuous but not normally distributed.
- There are repeated measurements taken from dependent samples.

III.5.2 Wilcoxon Signed-Ranks Test

The Wilcoxon signed-rank test [116] is a non-parametric statistical method designed to compare two related samples, matched pairs, or repeated measurements within a single sample. Commonly employed as a post-hoc analysis following the Friedman test, it identifies specific pairs of conditions that exhibit significant differences. The test operates by ranking the differences between paired observations and assessing whether the median of these differences significantly deviates from zero.

III.5.2.1 Hypotheses:

- **Null Hypothesis (H_0):** There is no difference between the two related samples or conditions.
- **Alternative Hypothesis (H_1):** There is a statistically significant difference between the two related samples or conditions.

III.5.2.2 Test Procedure:

- Calculate Differences: For each subject, compute the difference between the two conditions.

- Rank Absolute Differences: Rank the absolute values of the differences, assigning ranks from lowest to highest.
- Sum the Ranks: Sum the ranks of the positive and negative differences separately.
- Calculate the Test Statistic: The Wilcoxon test statistic is based on these rank sums.
- Evaluate Significance: The test statistic is compared to a critical value from the Wilcoxon distribution, or the p-value is calculated. If the p-value is less than a significance level (e.g., 0.05), the null hypothesis is rejected, indicating a significant difference between the two conditions.

III.5.2.3 Test Criteria:

A p-value less than the significance level (usually 0.05) and a z-value greater than -1.96 and 1.96 (for a 95% confidence level) confirm statistical significance in the performance between the conditions.

III.5.2.3 Use Cases:

The Wilcoxon signed-ranks test is useful when:

- Comparing two related or matched groups.
- The data are ordinal or not normally distributed.
- A nonparametric alternative to the paired t-test is needed to compare two dependent samples.

III.6 Experimental and discussion

III.6.1 Model evaluation

The performance of the proposed model is assessed using various metrics that emphasize reliability. When comparing the effectiveness of tree-based classifiers, different performance metrics may yield contrasting evaluations. While performance is typically gauged using accuracy, this study extends the evaluation to include additional metrics such as recall, F-measure, precision, Cohen's kappa, specificity, and the confusion matrix, which incorporates four key elements: True Negative (TN), False Positive (FP), True Positive (TP), and False Negative (FN) In this context, Po represents the observed relative agreement between two annotators (i.e., accuracy), while Pe denotes the hypothetical probability of agreement by chance.

$$Specificity = \frac{TN}{TN + FN} \tag{III.22}$$

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{III.23}$$

$$Precision = \frac{TP}{TP + FP} \tag{III.24}$$

$$Sensitivity = \mathrm{Re}call = \frac{TP}{TP + FN} \tag{III.25}$$

$$F_{measure} = \frac{2\left(\mathrm{Pr}ecision \times \mathrm{Re}call\right)}{\mathrm{Pr}ecision + \mathrm{Re}call} \tag{III.26}$$

$$Cohen's\ Kappa = \frac{P_o - P_e}{1 - P_e} \quad \text{(III.27)}$$

III.6.2 Data collection and experimental setup

Researchers and expert engineers in maintenance are continuously working to develop models that enhance the durability and precision of power system equipment diagnostics by utilizing data mining techniques and advanced classification models [117]. One widely used approach is DGA performed on transformer oil through gas chromatography, which involves collecting and analyzing oil samples to determine the type of fault present in a power transformer. The resulting sample data is compiled into an Excel file for further analysis. Table III.1 displays the sources and distribution of the training and testing data utilized in the study.

The next step involves organizing and inputting the data into MATLAB, where it is converted into suggested input vectors. These vectors are designed to isolate interfering gases and help identify the most accurate inputs for the analysis. After preprocessing, the data is transferred to the KNIME analytics platform, where tree-based learning algorithms are employed to predict transformer faults. Figure III.1 outlines the model's structure for diagnosing power transformer issues.

KNIME serves as a versatile analytics platform for data mining, supporting various libraries and software tools. It addresses the common challenges researchers face, such as time consumption and errors, by streamlining the preparation of the programming environment. Moreover, KNIME enables researchers and engineers to share workflows easily, a feature that proves beneficial as many of them lack specialized programming expertise. The KNIME workflow contract was designed using programming languages such as Python and Java, further simplifying its application [118]. Figure III.2 illustrates the proposed model within the KNIME platform and provides an overview of the workflow steps.

Table III.1 Data Distribution for Training and Testing

Ref.		PD	D1	D2	T1	T2	T3	Total
[119], [120],[121]	Training (70%)	23	41	69	60	33	58	**284**
	Testing (30%)	9	18	30	26	15	25	**123**
	Total	32	59	99	86	48	83	**407**

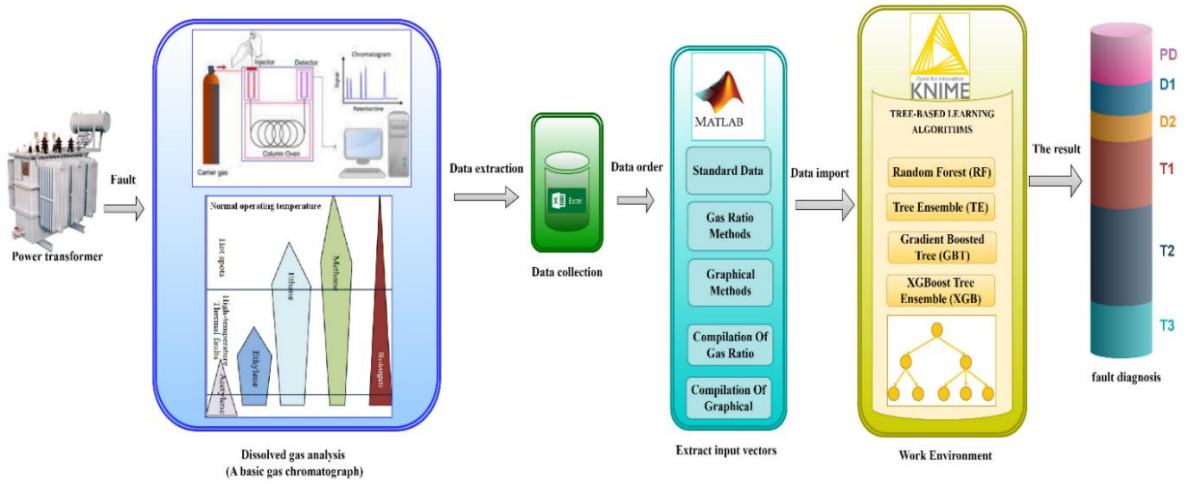


Figure III.1 Schematic representation of the implementation of the proposed diagnostic models

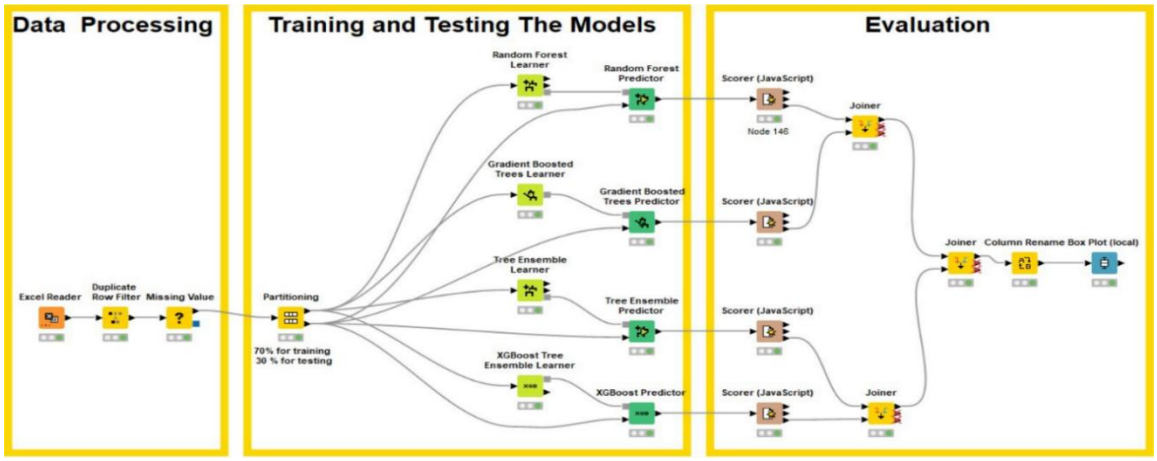


Figure III.2 KNIME workflow for enhancing DGA interpretation

- Data Processing: The first step involves reading the input file using the reader node, which scans and identifies the number and types of columns in the data. Next, the duplicate row node eliminates any redundant rows from the input table. Missing values are then processed and handled accordingly before further analysis.
- Training and Testing the Model: The data is split into 70% for training and 30% for testing using the partitioning node. The training process is conducted using various learner nodes (Random Forest, Gradient Boosted Trees, etc.), and the performance of the models is then evaluated with predictor nodes.
- Evaluation: The effectiveness of the developed models is evaluated using several metrics, including accuracy, recall, F-measure, precision, Cohen's kappa, specificity, and a confusion matrix, all generated by the scorer node.

In Figure III.3, the data preprocessing steps are further illustrated, along with a simplified overview of the tree-based ensemble learning process. The workflow begins with data collection, splitting it into 70% training and 30% testing datasets. Different ensemble learning methods, such as Random Forest (RF), Tree Ensemble (TE), Gradient Boosted Trees (GBT),

and XGBoost (XGB), are used for training. After model evaluation, the best results are selected to make a final decision.

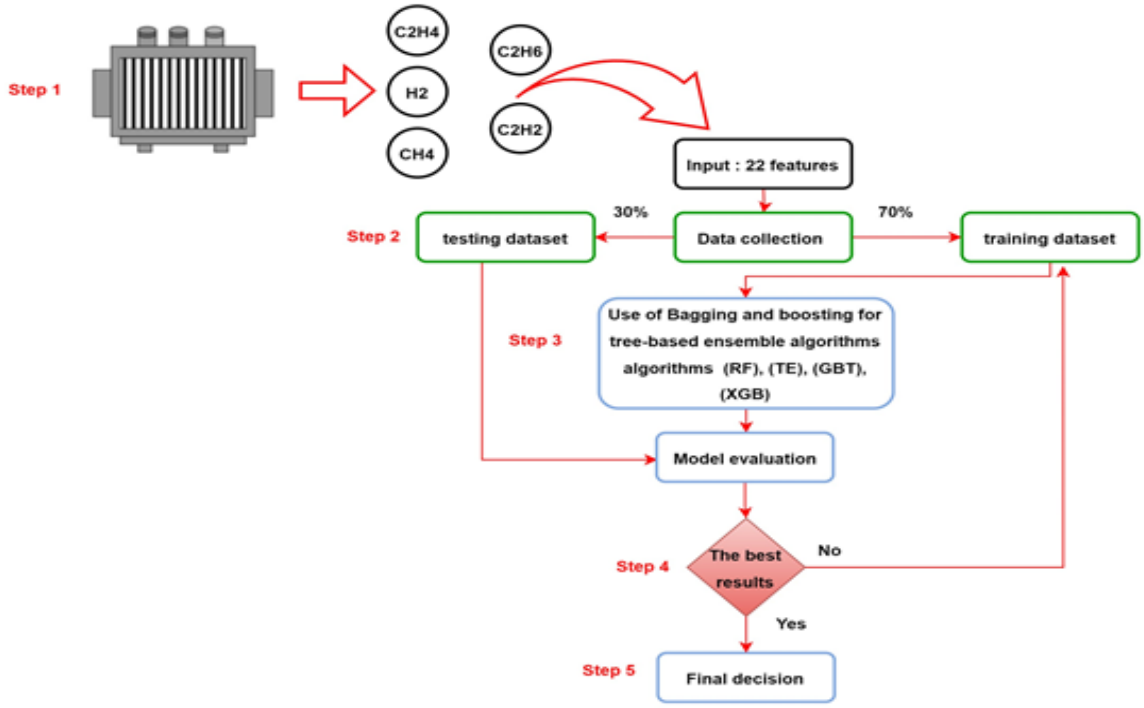


Figure III.3 Flowchart for implementation of the tree-based ensemble learning

III.6.3 Results and discussion

In this study, a total of 407 samples were used to assess the performance of the proposed tree-based model. Of these, 284 samples were randomly allocated for the training phase, while the remaining 123 samples were reserved for testing. The input vectors included various gas analysis methods: standard gas methods, gas ratio methods, graphical representation techniques, and combinations of both gas ratio and graphical methods, all serving as inputs to the tree-based algorithm.

As shown in Table III.2, the input vectors ranged from 1 to 22, and each vector corresponds to a specific accuracy rate for the different tree-based algorithms. The variation in accuracy between vectors, depending on the type of fault, can be attributed to the degree of compatibility between the input vector and the algorithm. The most suitable input vector achieves the highest accuracy, offering a robust approach for improving the early detection of power transformer faults.

Table III.2 The diagnosis accuracies of tree-based ensemble classifiers for all input vectors

Input vectors DGA technique		Model Accuracy (%)			
		RF	TE	GBT	XGB
Gas standard methods	**Vector 1**	82.93	78.86	78.05	81.30
	Vector 2	95.93	93.50	92.68	94.31
	Vector 3	87.80	86.99	80.49	82.93
	Vector 4	97.56	95.12	94.31	95.93
Gas ratios methods	**Vector 5**	85.37	85.37	84.55	86.99
	Vector 6	91.06	92.68	86.99	88.62
	Vector 7	95.93	94.93	92.68	95.12

	Vector 8	86.99	88.62	83.74	86.18
Graphical methods	**Vector 9**	81.30	78.05	75.61	78.86
	Vector 10	79.67	82.11	81.30	82.93
	Vector 11	91.06	91.87	86.18	90.24
	Vector 12	80.49	77.24	78.05	82.93
Combination of gas ratios methods	**Vector 13**	95.93	92.68	87.80	92.68
	Vector 14	90.24	92.68	90.24	88.62
	Vector 15	96.75	95.93	91.87	91.87
	Vector 16	98.37	96.75	95.93	97.56
Combination of graphical methods	**Vector 17**	86.99	84.55	84.55	87.80
	Vector 18	93.50	91.87	91.06	94.31
	Vector 19	93.50	92.68	90.24	94.31
	Vector 20	94.31	93.50	92.68	91.87
	Vector 21	95.12	93.50	91.06	94.31
	Vector 22	95.93	94.31	95.12	96.75

Figure III.4 presents the diagnostic accuracies of the 22 feature input vectors utilized in the study.

PD Fault (Figure III.4 (A)): The vectors 6, 8, 16, 17, 20, and 22 provide 100% diagnostic accuracy across all four algorithms. For the Gradient Boosted Trees (GBT) algorithm, vectors 8, 10, 14, 16, 17, 20, 21, and 22 achieve perfect diagnostic accuracy, with vectors derived from graphical methods performing the best. In contrast, for the Random Forest (RF) algorithm, only vectors 2, 9, 11, 12, 13, 18, and 22 achieve accuracies below 100%, with graphical methods being less effective for identifying partial discharge faults. The Tree Ensemble (TE) algorithm, on the other hand, achieves 100% accuracy with vectors 3, 4, 5, 6, 8, 10, 14, 15, 16, 17, 20, and 22, with ratio-based methods proving most effective. For the XGBoost (XGB) algorithm, vectors 5, 6, 8, 16, 17, 20, and 22 give 100% accuracy, including those generated from a combination of graphical and ratio methods.

D1 Fault (Figure III.4 (B)): For this fault, only the TE and RF algorithms achieve 100% accuracy. The GBT algorithm produces accuracy over 90% for vectors 4, 14, 15, 17, and 21, with a majority of these vectors being combined methods. Similarly, for the RF algorithm, vectors 4, 9, 15, 16, 17, and 21 achieve more than 90% accuracy, with combined vectors again dominating. The TE algorithm attains over 90% accuracy with the combined vectors 14, 15, 16, and 17, while for the XGB algorithm, vectors 14, 15, 17, and 21 surpass 90% accuracy, with combined methods showing superior performance across all algorithms.

D2 Fault (Figure III.4 (C)): The diagnostic performance for D2 faults is notably better than for D1. The GBT algorithm achieves accuracy greater than or equal to 90% for all vectors except 1, 6, 12, and 21. For the RF algorithm, vectors 1, 9, and 17 show less than 90% accuracy, while vectors 2, 13, 20, and 22 achieve perfect accuracy. The TE algorithm delivers accuracy below 90% for vectors 1, 9, 12, and 17, with vector 7, alongside vectors 13 and 22, reaching 100% accuracy. For the XGB algorithm, only vectors 1, 6, and 9 have accuracies below 90%.

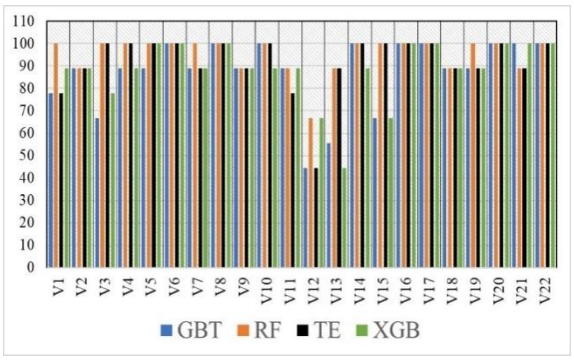
T1 Fault (Figure III.4 (D)): he diagnostic performance across all four algorithms is generally consistent. Vectors 5, 8, 9, 10, 12, and 17 achieve accuracy rates below 90%, while all other vectors surpass this threshold. For the Random Forest (RF) algorithm, vectors 2, 4, 7, 11, 15, 16, 18, and 19 achieve perfect accuracy (100%). The Tree Ensemble (TE) algorithm also attains 100% accuracy with vectors 3, 11, 15, 18, and 19. Similarly, the XGBoost (XGB) algorithm

reaches 100% accuracy with vectors 2, 3, 7, 16, 18, and 19. In summary, the combined vectors outperform individual methods in diagnosing T1 faults.

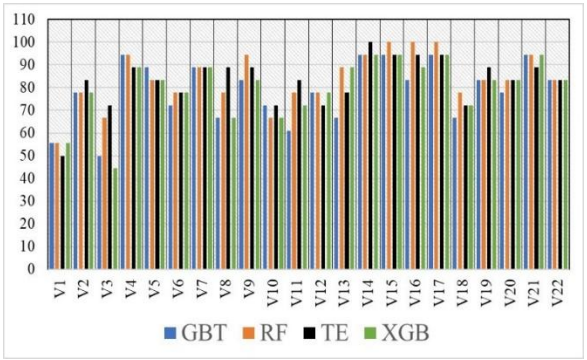
T2 Fault (Figure III.4 (E)): Vectors 4, 16, and 22 achieve 100% diagnostic accuracy with all four algorithms. Vectors derived from graphical methods are less effective individually but show improved accuracy when combined, as demonstrated by vectors 18 to 22, which average 90% accuracy.

T3 Fault (Figure III.4 (F)): The diagnostic accuracies of the four algorithms are closely aligned for each vector. Only vector 10 falls below 90% accuracy for the RF algorithm, while vectors 2, 8, 13, and 14 achieve 100%. For the XGB algorithm, all vectors except 17 exceed 90% accuracy, with vectors 2, 6, 7, 13, 14, and 16 performing the best. Both the GBT and TE algorithms achieve over 90% accuracy using combined vectors (vectors 13 to 22), while individual vectors show more variability, with graphical methods such as Duval and Gouda triangle vectors performing the worst.

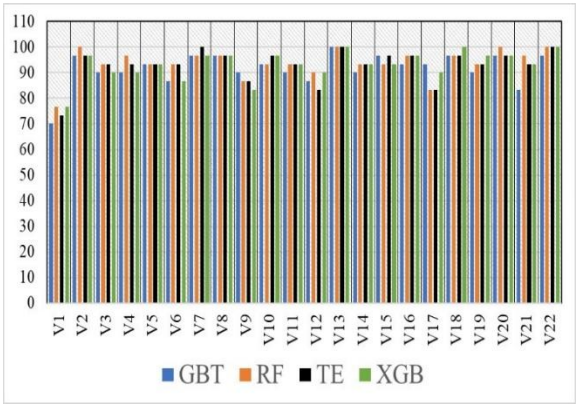
These results highlight the significant impact of input feature vectors on the classification performance of the algorithms used. While a specific input vector may excel in identifying one set of faults, it might be less effective for another. This variability also explains the subpar performance of certain intelligent methods when relying on a single feature vector for fault detection.



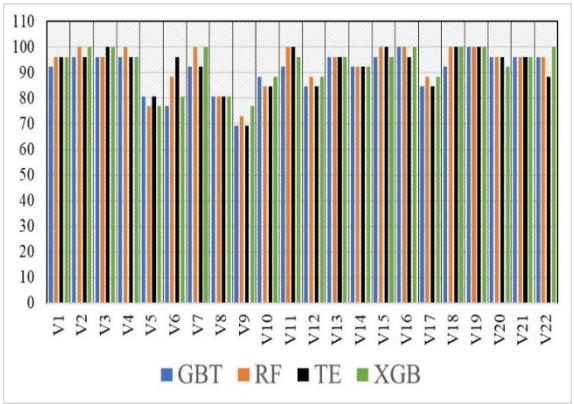
(A) PD



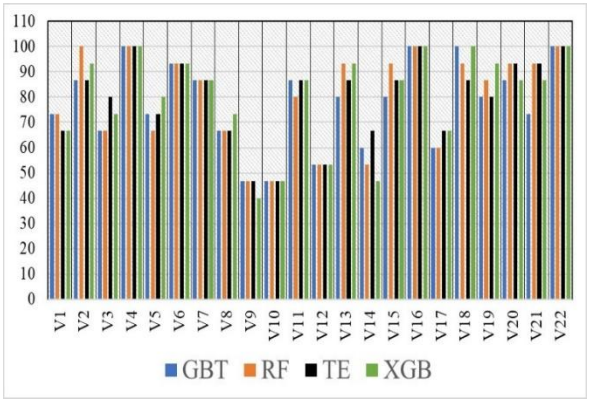
(B) D1



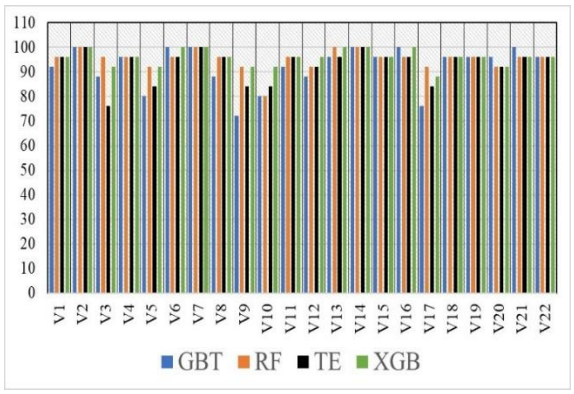
(C) D2



(D) T1



(E) T2



(F) T3

Figure III.4 The diagnosis accuracies according to fault type

Based on the obtained results, the most accurate input vector was identified for each classifier to enhance the DGA interpretation. Table III.3 summarizes the optimal input vectors for each classifier under the proposed DGA technique. For the gas standard methods, Vector 4, based on the Denkyoken method, was determined to be the best input. In the gas ratio methods, Vector 7, corresponding to Roger's Four-Ratio Method (RFRM), was preferred. For graphical representation methods, Vector 11, derived from the Duval Pentagon Method (DPM), performed best. When compiling gas ratio methods, Vector 16, representing a combination of the Three Ratios Technique (TRT) with Roger's Four-Ratio Method (RFRM), provided the highest accuracy. Additionally, for the compilation of graphical methods, Vector 22, combining the Duval Pentagon Method (DPM) and the Gouda Triangle Method (GTM), showed superior performance.

The results demonstrate that Vector 16, which combines the Three Ratios Technique (TRT) with Rogers' Four-Ratio Method (RFRM), is the most effective input vector. When evaluated on the testing dataset, this vector achieved accuracies of 98.37%, 96.75%, 95.93%, and 97.56% using the Random Forest (RF), Tree Ensemble (TE), Gradient Boosted Trees (GBT), and XGBoost (XGB) algorithms, respectively.

Table III.3 Tree-based learning classification for all input vectors

DGA technique	Model Accuracy (%)				
	Vectors	RF	TE	GBT	XGB
Gas standard methods	Vector 4 (Denkyoken)	97.56	95.12	94.31	95.93
Gas ratios methods	Vector 7 (RFRM)	95.93	94.31	92.68	95.12
Graphical methods	Vector 11 (DPM)	91.06	91.87	86.18	90.24
Combination of gas ratios methods	Vector 16 (RFRM with TRT)	98.37	96.75	95.93	97.56
Combination of graphical methods	Vector 22 (DPM with GTM)	95.93	94.31	95.12	96.75

III.6.4 Statistical significance tests for the 22 input vectors

To validate the performance of RF, TE, GBT, and XGB algorithms with 22 input vectors, the Friedman test was employed to detect statistically significant differences between them. The test operates with two hypotheses: the null (H0) assumes no significant differences in performance, while the alternative (H1) suggests otherwise. If the p-value is less than 0.05 and the Chi-Square exceeds 3.841, H0 is rejected, indicating a significant performance difference.

Since the Friedman test only identifies if a difference exists, the Wilcoxon signed-ranks test was used for pairwise comparisons to determine which algorithm performed better. A p-value below 0.05 and a z-value outside the range of -1.96 to 1.96 confirm statistical significance in performance differences.

Table III.4 Friedman test for mean ranks

Algorithms	Mean Rank
RF	3.39
TE	2.48
GBT	1.30
XGB	2.84

Nonparametric tests were carried out based on the diagnostic accuracy outcomes of the tree-based algorithms for all 22 vectors listed in Table III.2. The mean rank analysis using the Friedman test, which evaluates the accuracy performance of the tree-based algorithms, is summarized in Table III.4.

Table III.4 indicates significant differences between the mean ranks of the algorithms, and the Friedman test rejects the null hypothesis (Table III.5), as the p-value of 0.00000054 is below the 0.05 threshold, and the Chi-Square statistic of 31.935 surpasses the critical value of 3.88. This demonstrates that the performance differences among the algorithms are statistically significant.

Table III.5 Friedman test statistics

N	**22**
Chi-Square	31.935
df	3
Asymp. Sig.	0.00000054

To identify which algorithms, outperform the others, a post-hoc Wilcoxon signed-ranks test was performed on the algorithms, including pairwise comparisons such as RF vs TE, RF vs GBT, and RF vs XGB.

Table III.6 presents the Wilcoxon signed-ranks test results, which show the Random Forest algorithm consistently outperforming the others in terms of mean ranks. Table III.7 provides the Wilcoxon rank test statistics, confirming the statistical significance of the algorithms' performance differences.

Table III.6 Wilcoxon signed ranks test results

		N	Mean Rank	Sum of Ranks
RF vs TE	**Negative Ranks**	5[a]	10.60	53.00
	Positive Ranks	16[b]	11.13	178.00
	Ties	1[c]		
	Total	22		
RF vs GBT	**Negative Ranks**	1[d]	3.50	3.50
	Positive Ranks	20[e]	11.38	227.50
	Ties	1[f]		
	Total	22		
RF vs XGB	**Negative Ranks**	6[g]	8.42	50.50
	Positive Ranks	15[h]	12.03	180.50
	Ties	1[i]		
	Total	22		

a. RF < TE ; b. RF > TE ; c. RF = TE ; e. RF > GBT ; f. RF = GBT ; g. RF < XGB ; h. RF > XGB ; i. RF = XGB

Table III.7 Evaluation of the statistical significance of performance using the two-tailed Wilcoxon signed-rank test

	RF vs TE	RF vs GBT	RF vs XGB
Z	-2.177	-3.900	-2.271
Asymp. Sig. (2-tailed)	0.029	0.000096	0.023
Statistical significance	Yes	Yes	Yes

The Friedman and Wilcoxon signed-ranks tests, both non-parametric methods, confirmed that the tree-based algorithms exhibit statistically significant differences in performance. The Random Forest algorithm demonstrated superior performance over the others, with further validation expected through the selection of the best input vector results.

II.6.5 Performance analysis of the optimal input vector

The confusion matrix is a crucial tool for assessing the accuracy of a model by comparing actual fault outcomes with predicted results (PD, D1, D2, T1, T2, and T3). This evaluation is performed using the input vector, Vector 16, which incorporates both the Three Ratios Technique and Roger's Ratios.

	D1 (Predicted)	D2 (Predicted)	PD (Predicted)	T1 (Predicted)	T2 (Predicted)	T3 (Predicted)	
D1 (Actual)	18	0	0	0	0	0	100.00%
D2 (Actual)	0	29	0	1	0	0	96.67%
PD (Actual)	0	0	9	0	0	0	100.00%
T1 (Actual)	0	0	0	26	0	0	100.00%
T2 (Actual)	0	0	0	0	15	0	100.00%
T3 (Actual)	0	0	0	0	1	24	96.00%
	100.00%	100.00%	100.00%	96.30%	93.75%	100.00%	

(A) RF

	D1 (Predicted)	D2 (Predicted)	PD (Predicted)	T1 (Predicted)	T2 (Predicted)	T3 (Predicted)	
D1 (Actual)	17	1	0	0	0	0	94.44%
D2 (Actual)	1	29	0	0	0	0	96.67%
PD (Actual)	0	0	9	0	0	0	100.00%
T1 (Actual)	1	0	0	25	0	0	96.15%
T2 (Actual)	0	0	0	0	15	0	100.00%
T3 (Actual)	0	0	0	0	1	24	96.00%
	89.47%	96.67%	100.00%	100.00%	93.75%	100.00%	

(B) TE

	D1 (Predicted)	D2 (Predicted)	PD (Predicted)	T1 (Predicted)	T2 (Predicted)	T3 (Predicted)	
D1 (Actual)	15	2	0	0	1	0	83.33%
D2 (Actual)	2	28	0	0	0	0	93.33%
PD (Actual)	0	0	9	0	0	0	100.00%
T1 (Actual)	0	0	0	26	0	0	100.00%
T2 (Actual)	0	0	0	0	15	0	100.00%
T3 (Actual)	0	0	0	0	0	25	100.00%
	88.24%	93.33%	100.00%	100.00%	93.75%	100.00%	

(C) GBT

	D1 (Predicted)	D2 (Predicted)	PD (Predicted)	T1 (Predicted)	T2 (Predicted)	T3 (Predicted)	
D1 (Actual)	16	1	0	0	1	0	88.89%
D2 (Actual)	1	29	0	0	0	0	96.67%
PD (Actual)	0	0	9	0	0	0	100.00%
T1 (Actual)	0	0	0	26	0	0	100.00%
T2 (Actual)	0	0	0	0	15	0	100.00%
T3 (Actual)	0	0	0	0	0	25	100.00%
	94.12%	96.67%	100.00%	100.00%	93.75%	100.00%	

(D) XGB

Figure III.5 Confusion matrix for the proposed models using Vector 16

Table III.8 Performance evaluation of tree-based learning classifiers using the best input vector (Vector 16)

Model	Class	TP	FP	TN	FN	Recall (%)	Precision (%)	Specificity (%)	F-measur (%)	Cohen's Kappa	Accuracy Overall (%)
RF	PD	9	0	114	0	100	100	100	100	0.980	98.37
	D1	18	0	105	0	100	100	100	100		
	D2	29	0	93	1	96.67	100	100.00	98.31		
	T1	26	1	96	0	100	96.30	98.97	98.11		
	T2	15	1	107	0	100	93.75	99.07	96.77		
	T3	24	0	98	1	96.00	100	100	97.96		
ET	PD	9	0	114	0	100	100	100	100	0.960	96.75
	D1	17	2	103	1	94.44	89.47	98.10	91.89		
	D2	29	1	92	1	96.67	96.67	98.92	96.67		
	T1	25	0	97	1	96.15	100	100	98.04		
	T2	15	1	107	0	100	93.75	99.07	96.77		
	T3	24	0	98	1	96.00	10	100	97.96		
GBT	PD	9	0	114	0	100	100	100	100	0.950	95.93
	D1	15	2	103	3	83.33	88.24	98.10	85.71		
	D2	28	2	91	2	93.33	93.33	97.85	93.33		
	T1	26	0	97	0	100	100	100	100		
	T2	15	1	107	0	100	93.75	99.07	96.77		
	T3	25	0	98	0	100	100	100	100		
XGB	PD	9	0	114	0	100	100	100	100	0.970	97.56
	D1	16	1	104	2	88.89	94.12	99.05	91.43		
	D2	29	1	92	1	96.67	96.67	98.92	96.77		
	T1	26	0	97	0	100	100	100	100		
	T2	15	1	107	0	100	93.75	99.07	96.77		
	T3	25	0	98	0	100	100	100	100		

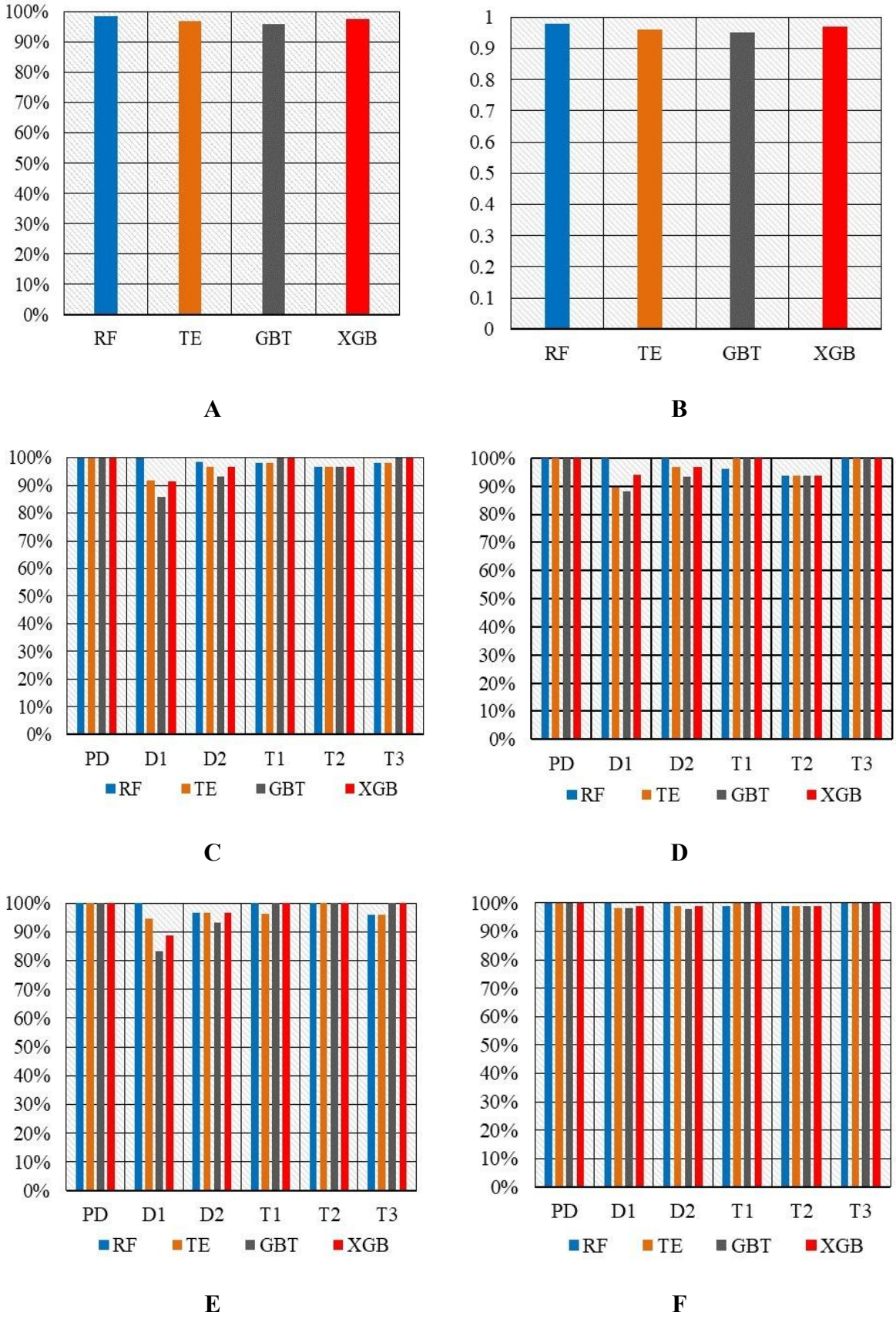


Figure III.6 Evaluation metrics: (A) Accuracy, (B) Cohen's Kappa, (C) F-measure, (D) Precision, (E) Recall, and (F) Specificity

In Figure III.5, the blue diagonal cells represent correctly classified data, while the off-diagonal cells indicate misclassifications by the classification algorithms. Figure III.5(A) demonstrates that all faults (PD, D1, T1, and T2) were correctly classified, with two exceptions: fault D2 was misclassified as T1, and fault T3 as T2. Figure III.5(B) shows correct classification for faults PD and T2; however, faults D1 and T3 were misclassified, with D1 labeled as D2, T1 as D1, and T3 as T2. Figure III.5(C) indicates that fault D1 was misclassified twice as D2 and once as T2, while D2 was misclassified twice as D1. The remaining faults (PD, T1, T2, and T3) were correctly classified. Similarly, Figure III.5(D) reveals that D1 was misclassified once as D2 and once as T2, while D2 was incorrectly classified as D1. The remaining faults (PD, T1, T2, and T3) were accurately classified.

Table III.9 Distribution of validation samples and comparison according to Types of fault and references

Ref.	PD	D1	D2	T1	T2	T3	Total
[120]	4	3	4	1	0	0	12
[122]	0	0	2	0	0	0	2
[123]	0	3	1	2	0	1	7
[124]	1	0	4	4	0	2	11
[125]	0	1	4	2	1	0	8
[126]	1	0	0	4	2	14	21
[127]	0	2	2	0	0	0	4
[128]	0	0	0	0	1	2	3
[129]	1	4	2	0	2	8	17
[130]	1	0	0	0	1	0	2
[131]	0	0	0	0	0	2	2
Total	8	13	19	13	7	29	89

Table III.7 illustrates the classification performance of the proposed models based on several evaluation metrics. The tree-based learning algorithms applied to the input vector 6 TRT combined with RFRM, as shown in Table III.8, exhibit significant efficiency across these metrics. Nonetheless, slight variations are observed in metrics such as recall, F-measure, precision, and specificity, with the random forest algorithm demonstrating a marginal advantage in terms of accuracy and Cohen's kappa. Figure III.6 visually depicts these evaluation metrics.

To validate and compare the proposed model, a dataset of 89 samples was employed. This data was used to evaluate the model's effectiveness and compare its performance with previous studies, as detailed in Table III.9, which shows the data distribution based on the reference for each sample.

Table III.10 presents a comparison of the proposed model with earlier research, focusing on the severity of faults (PD, D1, D2, T1, T2, and T3) for various algorithms, including Random Forest (RF), Tree Ensemble (TE), Gradient Boosted Tree (GBT), and XGBoost Tree Ensemble

(XGB). The respective accuracies of these algorithms were 92.13%, 91.01%, 89.89%, and 91.01%, with tree-based learning methods consistently outperforming other techniques.

Table III.10 Comparison between the proposed method and other methods

		Model Accuracy (%)						
	Ref.	**PD**	**D1**	**D2**	**T1**	**T2**	**T3**	**TOTAL**
Previous studies	[113]	37.50	53.85	15.79	69.23	71.43	82.76	57.30
	[87]	87.50	61.54	84.21	84.62	57.14	89.66	80.90
	[86]	100.00	61.54	84.21	84.62	57.14	89.66	82.02
	[87]	100.00	61.54	84.21	76.92	57.14	86.21	79.78
	[71]	75.00	46.15	89.47	69.23	71.43	93.10	78.65
	[64]	75.00	0.00	63.16	76.92	14.29	55.17	50.56
	[83]	100.00	69.23	84.21	92.31	42.86	89.66	83.15
	[130]	75.00	53.85	78.95	53.85	57.14	89.66	73.03
	[80]	87.50	69.23	84.21	53.85	57.14	93.10	78.65
	[85]	75.00	76.92	84.21	84.62	57.14	82.76	79.78
Proposed Model	RF [132]	100.00	92.31	89.47	92.31	85.71	93.10	92.13
	TE [132]	100.00	84.62	89.47	92.31	85.71	96.43	91.01
	GB [132]	100.00	84.62	89.47	84.62	71.43	96.55	89.89
	XGB [132]	100.00	92.31	89.47	92.31	71.43	93.10	91.01

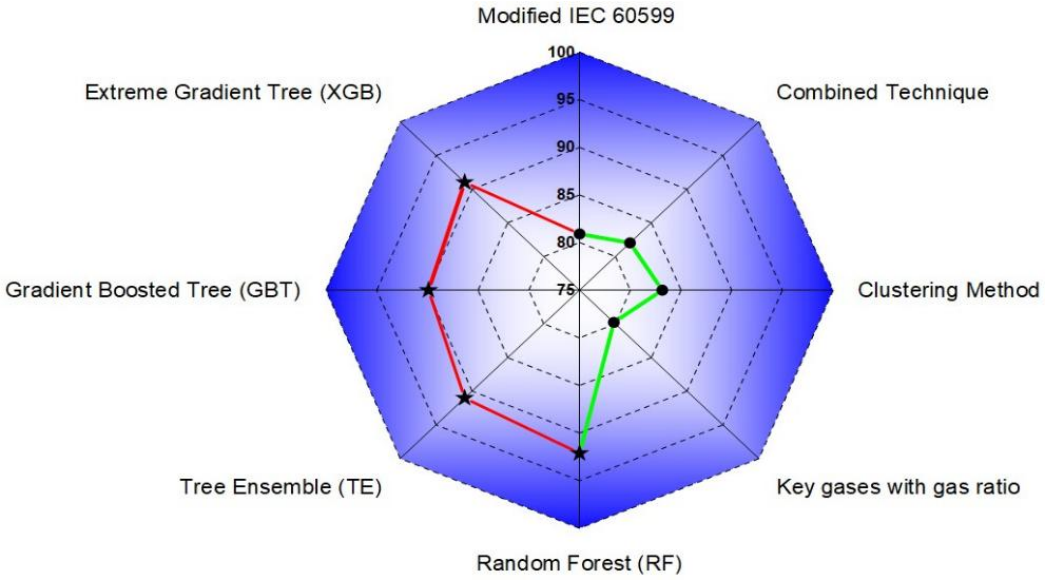


Figure III.7 Comparison of the best results of the previous methods with the proposed methods

Figure III.7 presents a comprehensive theoretical framework for comparing the proposed method with the best results from previous research. Specifically, the RF, TE, GBT, and XGBoost models achieved diagnostic accuracies of 92.13%, 91.01%, 89.89%, and 91.01%, respectively. These values show a significant improvement over earlier methods, including the Clustering Method (83.15%), the combined technique (82.02%), the modified IEC 60599 (80.90%), and Key Gases with Gas Ratios (79.78%). This marked increase in accuracy highlights the effectiveness of using input vectors in combination with advanced algorithms, enabling more precise transformer fault diagnosis and addressing the shortcomings of existing methodologies.

III.7 Conclusion

This chapter presents several intelligent models for diagnosing faults in power transformers. The proposed diagnostic models are based on decision tree-based ensemble learning techniques and utilize a range of feature input vectors. A total of 22 feature vectors, derived from traditional diagnostic methods, serve as inputs for four tree-based ensemble algorithms: Random Forest, Tree Ensemble, Gradient Boosted Trees, and Extreme Gradient Boosting. The models are implemented and evaluated using 407 labeled samples, representing six distinct fault types.

The results demonstrate that the optimal performance is achieved with a 16-feature vector, composed of gas ratios from Rogers' four-ratio method and the three-ratio technique. This combination surpasses the accuracy of existing methods documented in the literature.

Chapter IV. Application and Evaluation of AI Technique for Diagnosis

IV.1 Introduction

This chapter presents a novel approach to power transformer fault diagnosis by leveraging machine learning algorithms in combination with ensemble techniques. The proposed methodology integrates dimension-reduced DGA input features, obtained through Principal Component Analysis (PCA), with ensemble models such as Bagging, Decorate, and Boosting, ensuring that only principal components are utilized. To address data imbalance, synthetic data is generated using the Long Short-Term Memory (LSTM) algorithm. The findings underscore the efficacy of this approach, which not only outperforms conventional DGA-AI combination techniques but also enhances input vector consistency, smooths data integration with ensemble methods, and mitigates previous methods limitations.

IV.2 AI Techniques

This section explores various artificial intelligence techniques utilized in the proposed model. These include a range of algorithms categorized into machine learning, deep learning, and ensemble learning methods. Additionally, Principal Component Analysis (PCA) is employed as a preprocessing step to enhance the performance of these algorithms.

IV.2.1 ML Techniques

IV.2.1.1 Decision Trees (DT)

Decision Trees (DT) are a widely recognized classification algorithm known for their relatively simple structure. They excel at accurately analyzing substantial amounts of data within a short timeframe, making them particularly well-suited for efficient mass data processing [133]. The DT classifier is constructed using a combination of internal nodes and leaf nodes, where internal nodes represent decision thresholds based on feature values, and leaf nodes correspond to the predicted class labels as in Figure IV.1. The training process involves recursively identifying the best partition at each node to minimize uncertainty about the class labels, commonly assessed through metrics such as Entropy and Gini Impurity [134].

Entropy is calculated as:

$$H(S) = -\sum_{i=1}^{c} p_i \log_2 \ p_i \qquad \text{(IV.1)}$$

Gini Index is defined as:

$$IG(S,A) = H(S) - \sum_{v \in A} \frac{|S_v|}{|S|} H(S_v) \quad \text{(IV.2)}$$

where p_i is the probability of class i in set S.

Information Gain is used to determine the best feature to split on:

$$\text{Gini}(S) = 1 - \sum_{i=1}^{c} (p_i)^2 \quad \text{(IV.3)}$$

This process continues until the dataset is fully analyzed, ensuring that each path through the tree leads to a definitive classification outcome.

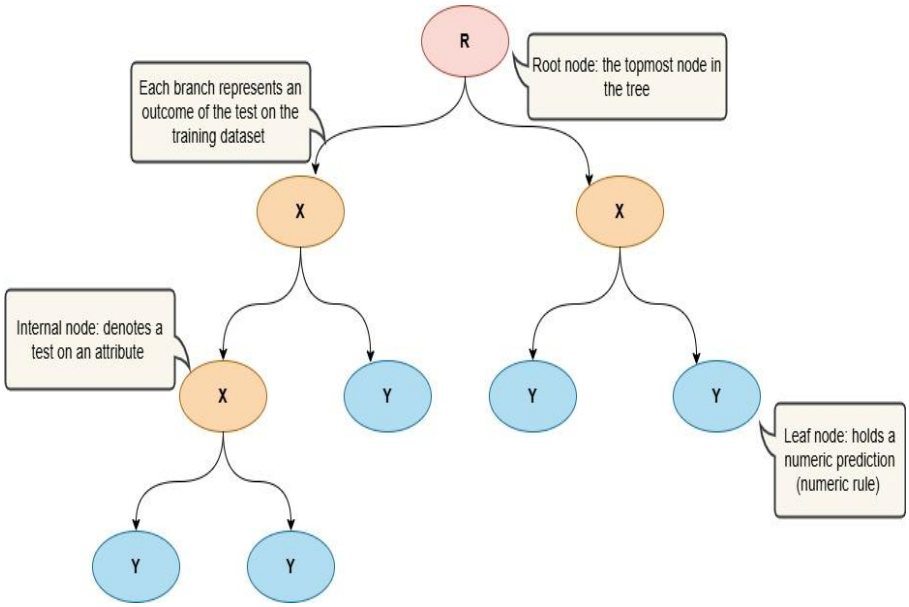


Figure IV.1 Decision tree structure

IV.2.1.2 Support Vector Machine (SVM)

Support Vector Machines (SVM) are sophisticated classifiers that originated from the work of Vapnik in the 1990s, utilizing a linear discriminant function to effectively separate classes within a dataset. Renowned for their accuracy and robustness, SVMs have become popular in data mining applications, proficiently classifying both linear and nonlinear data [135]. Figure IV.2 illustrates the SVM algorithm. The algorithm identifies the optimal hyperplane defined by the equation:

$$f(x) = w^T x + b \quad \text{(IV.4)}$$

Where w is the weight vector, x is the input feature vector, and b is the bias. The objective is to maximize the margin $\frac{2}{\|w\|}$ between the classes, subject to the constraints:

$$w^T x_i + b) \geq 1 \text{ for all } i \quad \text{(IV.5)}$$

y_i(The effectiveness of SVM significantly depends on the selection of a kernel function, which allows the transformation of the input space into a higher-dimensional feature space, enabling linear discrimination even in complex datasets [136].

Kernel Trick can be represented as:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad \text{(IV.6)}$$

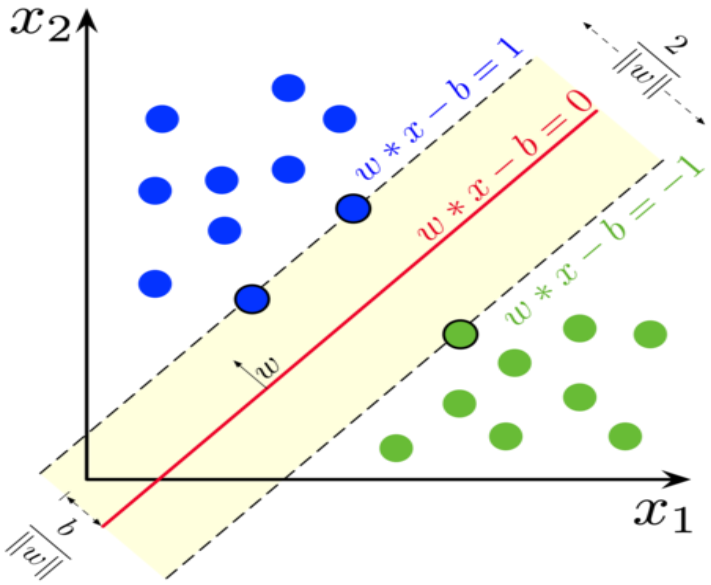


Figure IV.2 SVM algorithm principle

IV.2.1.3 K Nearest Neighbor (KNN)

The K Nearest Neighbor (KNN) algorithm is one of the simplest intelligent algorithms, characterized by its lack of a traditional learning stage. Its operation revolves around calculating distances between a query point and the points in the dataset to identify the nearest neighbors [137]. The class label assigned to the query point s is determined by the majority class among its k closest neighbors as in Figure IV.3, mathematically expressed as:

$$\hat{y} = \text{argmax}_c \sum_{j=1}^{k} I(y_j = c) \quad \text{(IV.7)}$$

where I is an indicator function that equals 1 if the class label y_j matches c and 0 otherwise. The distance between points is often measured using Euclidean distance:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \quad \text{(IV.8)}$$

This method enhances the classification process by extending the nearest neighbor concept, allowing the algorithm to leverage more information from multiple nearby points during the decision-making stage [138].

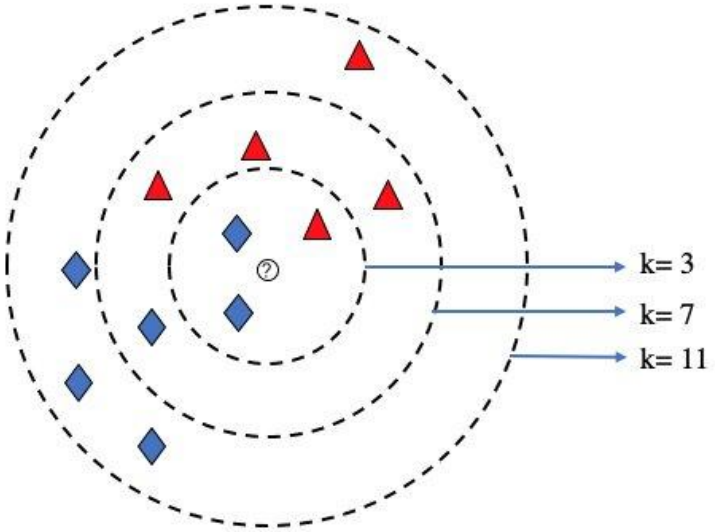


Figure IV.3 Principle of KNN algorithm with different k value

IV.2.1.4 Radial Basis Function Networks (RBFN)

Radial Basis Function Networks (RBFN) serve as an intelligent interpolation technique designed for modeling both linear and nonlinear relationships in multidimensional data, commonly applied in forecasting problems [139]. RBFNs are structured as multi-layer feed-forward neural networks, consisting of an input layer, a hidden layer, and an output layer, as shown in Figure IV.4. Each neuron in the hidden layer applies a radial basis function as an activation function, which can be defined as:

$$\phi_j(x) = \exp\left(-\frac{\|x - c_j\|^2}{2\sigma_j^2}\right) \qquad \text{(IV.9)}$$

where c_j is the center of the radial basis function and σ_j is the width. The network output $y(x)$ is given by:

$$y(x) = \sum_{j=1}^{N} w_j \phi_j(x) + b \qquad \text{(IV.10)}$$

where w are the weights connecting the hidden neurons to the output layer, and b is the bias term. The weights are iteratively adjusted to minimize prediction errors, commonly quantified using Mean Squared Error:

$$E = \frac{1}{2}\sum_{i=1}^{M}(y_i - \hat{y}_i)^2 \qquad \text{(IV.11)}$$

The construction and training of an RBFN focus on determining optimal centers, widths, and weights for the radial functions, which is a crucial aspect of its functionality [140].

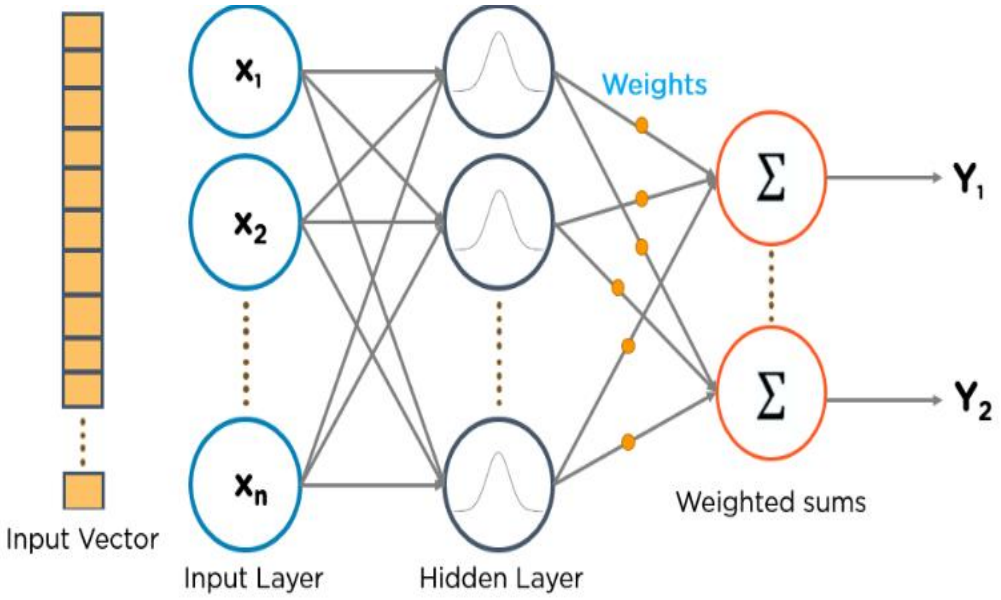


Figure IV.4 Structure of RBFN algorithm

IV.2.2 DL Techniques

IV.2.2.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory networks represent a sophisticated form of recurrent neural networks specifically designed for modeling complex patterns in sequential data [141]. By accounting for both past outputs and current inputs when predicting the next time step, LSTMs excel at capturing intricate dependencies over time. This capability is particularly valuable for generating synthetic datasets that mirror the statistical properties of real-world data while prioritizing privacy preservation [142]. A critical best practice is to shuffle the dataset before training the LSTM model to mitigate any potential bias stemming from the order of samples. Figure IV.5 illustrates the architecture of the LSTM algorithm.

In an LSTM, each cell at time t maintains a hidden state h_t and a cell state C_t that control the flow of information through the network. The core operations within an LSTM cell are as follows:

Forget Gate: Determines what fraction of the previous cell state C_{t-1} should be retained in the current cell state C_t . It is represented by:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \qquad \text{(IV.12)}$$

where W_f and b_f are the weight matrix and bias for the forget gate, h_{t-1} is the hidden state from the previous time step, x_t is the input at the current time step, and σ is the sigmoid activation function.

Input Gate: Decides the extent to which new information from the current input x_t will impact the cell state. This is done in two steps:

First, an input gate vector i controls the importance of the input:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad \text{(IV.13)}$$

Next, a candidate cell state $\tilde{C}_t$ is computed using a tanh function to add non-linearity:

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \quad \text{(IV.14)}$$

Cell State Update: The cell state C_t is updated by combining the previous cell state C_{t-1} (scaled by f_t and the candidate cell state $\tilde{C}_t$ (scaled by i_t):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad \text{(IV.15)}$$

Output Gate: Finally, the output of the LSTM cell, which also serves as the hidden state h_t for the next time step, is determined by:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \text{(IV.16)}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \quad \text{(IV.17)}$$

where o_t is the output gate vector, and h_t becomes both the output of the current time step and the hidden state for the next.

This structure allows LSTMs to maintain and update memory through long sequences, effectively capturing long-term dependencies that are essential for accurately generating synthetic time series data. By training on real-world data sequences, the LSTM model learns statistical patterns that can be used to synthesize new data points with similar characteristics, thus preserving privacy while ensuring statistical relevance [143].

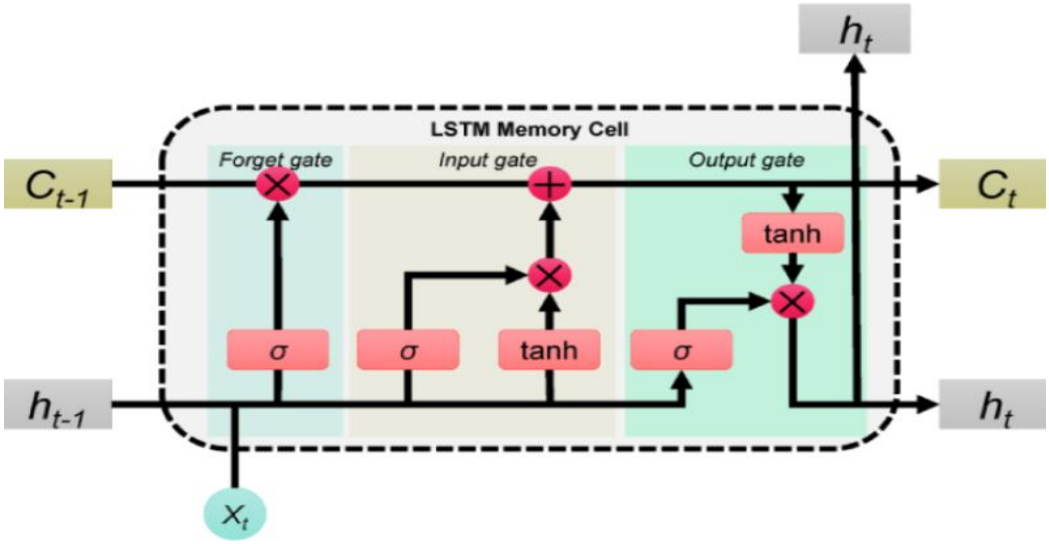


Figure IV.5 Architecture of LSTM algorithm

IV.2.3 EL Techniques

Ensemble learning techniques combine multiple models to enhance performance and robustness. Several key methods include:

IV.2.3.1 Bagging

Bootstrap aggregating, commonly referred to as bagging, is an ensemble strategy initially introduced by Breiman. Bagging stands out as one of the most well-known ensemble methods, aiming to enhance accuracy by amalgamating the outcomes of multiple learned classifiers into a single prediction, thereby creating an improved composite classifier [144]. The initial step in the bagging technique involves creating multiple forms, followed by generating models based on the actual method using random subsamples of the dataset. These subsamples are generated

from the original dataset randomly, employing the bootstrap sampling technique [145]. As shown in Figure IV.6.

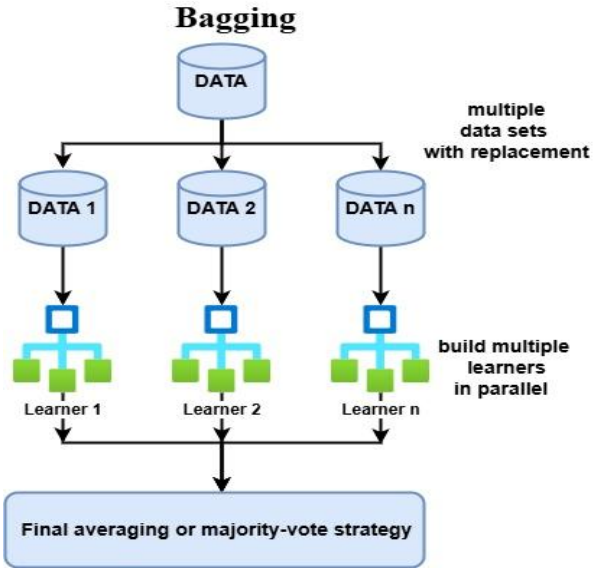


Figure IV.6 Structure of Bagging

IV.2.3.2 Decorate

The diverse ensemble creation by oppositional relabeling of artificial training examples method, known as Decorate, was introduced by Melville and Mooney. Differing from other ensemble methods, this meta-learning algorithm is distinctive in that it explicitly evaluates and utilizes variation in each iteration to generate ensemble classifiers [146]. In each iteration, the base learner is trained on the augmented training dataset, which includes artificially generated data points [147]. As shown in Figure IV.7.

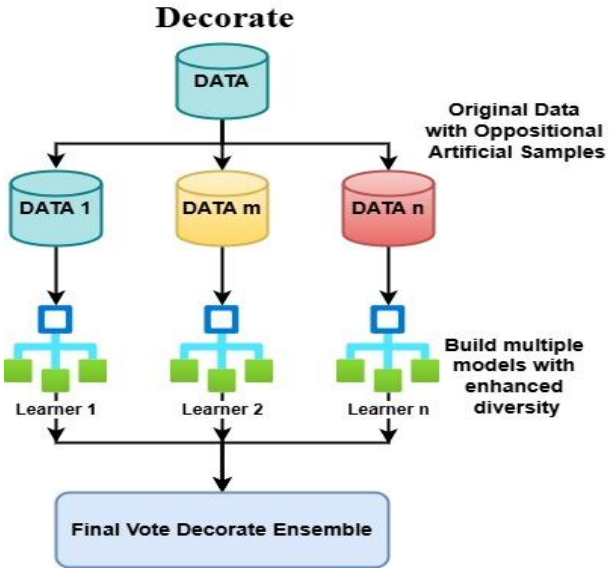


Figure IV.7 Structure of Decorate

IV.2.3.2 Boosting

The multi-boost learner ensemble learning, or boosting, was introduced to learners by Freund and Schapire. Boosting is a technique employed to combine weak learners, creating a robust classifier to enhance accuracy [148]. Operating as a form of sequential ensemble learning, boosting involves training weak learners on a given dataset, where the output of one weak learner becomes the input for the next, iteratively improving the overall accuracy of the final model. Boosting is notable for significantly reducing the bias and variance of test error, often resulting in more accurate classification outcomes compared to bagging and Decorate in many cases [149]. As shown in Figure IV.8.

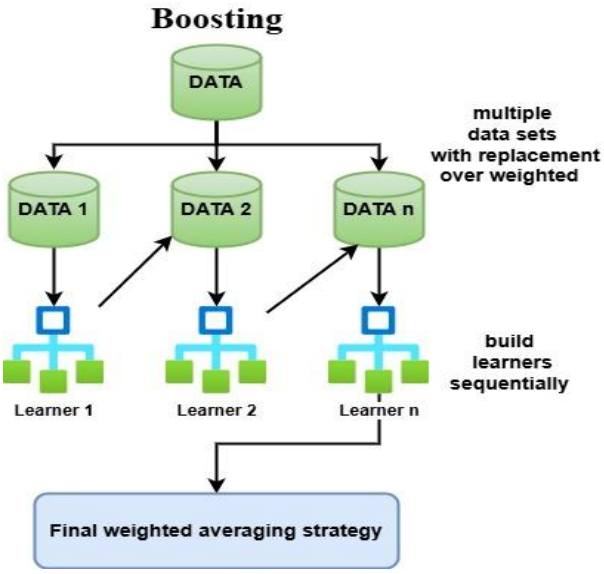


Figure IV.8 Structure of Boosting

IV.2.4 Principal Component Analysis (PCA)

PCA was originally introduced by Karl Pearson, and further developed by Hotelling. PCA, an unsupervised and non-parametric algorithm, finds significant application in machine learning for dimensionality reduction. Its primary purpose is to reduce the dimensionality of high-dimensional data by extracting the main feature components while preserving as much of the original data's variance as possible [150]. This analysis is geared towards identifying the optimal transformation that expresses the existing data with fewer variables. The transformed variables are referred to as principal components [151].

PCA first involves performing an eigenvalue decomposition on the covariance matrix of the standardized data to obtain the principal components. Here's a breakdown of the primary quantities involved:

Principal Component Coefficients: Let V represent the matrix of eigenvectors, also known as principal components or coefficients.

Principal Component Scores: These scores, denoted as Z, are projections of the standardized data onto the principal components:

$$Z = X'V \qquad \text{(IV.18)}$$

where X' is the standardized data, and Z is the matrix of scores.

Variance Explained by Each Component: The eigenvalues λ represent the amount of variance explained by each principal component. These are stored in the latent variable, while the explained variable gives the variance percentage for each component.

To retain a certain level of variance (e.g., 95%), we compute the cumulative variance explained by the principal components and select the minimum number k that meets this threshold:

$$\text{Cumulative Explained Variance} = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \times 100 \qquad \text{(IV.19)}$$

where p is the total number of features.

Choose k so that the cumulative explained variance meets or exceeds the desired threshold (e.g., 95%).

Once k, the number of principal components, is determined, we reduce the dimensionality by projecting the data onto these first k components. The reduced data, denoted as Z_k, is computed as:

$$Z_k = X'V_k \qquad \text{(IV.20)}$$

where V_k is the matrix of the top k eigenvectors.

PCA reduces the original data to a set of principal components that retain the essential variance, effectively simplifying the data while preserving critical information.

IV.3 Experimental setup for AI application for DGA

Researcher and engineer are exploring innovative methods with a focus on precision in diagnosis to ensure power equipment safety and extend electrical system lifespan, particularly power transformers. In this endeavor, the proposed methodology aims to introduce novel solutions for diagnosing power transformers. The initial phase involves meticulous data collection, followed by the conversion of this data into input vectors tailored for graphical methods. Subsequently, a normalization process is applied, accompanied by dimensionality reduction on the input vectors. In the final stage of data processing, a transformation from unbalanced to balanced data is achieved through the integration of synthetic data. Both the original and synthetic datasets are then employed to assess classification algorithms, which are further augmented by the incorporation of ensemble techniques to enhance overall accuracy. The sequential steps of this methodology are visually depicted in Figure IV.9, with comprehensive explanations provided in this section.

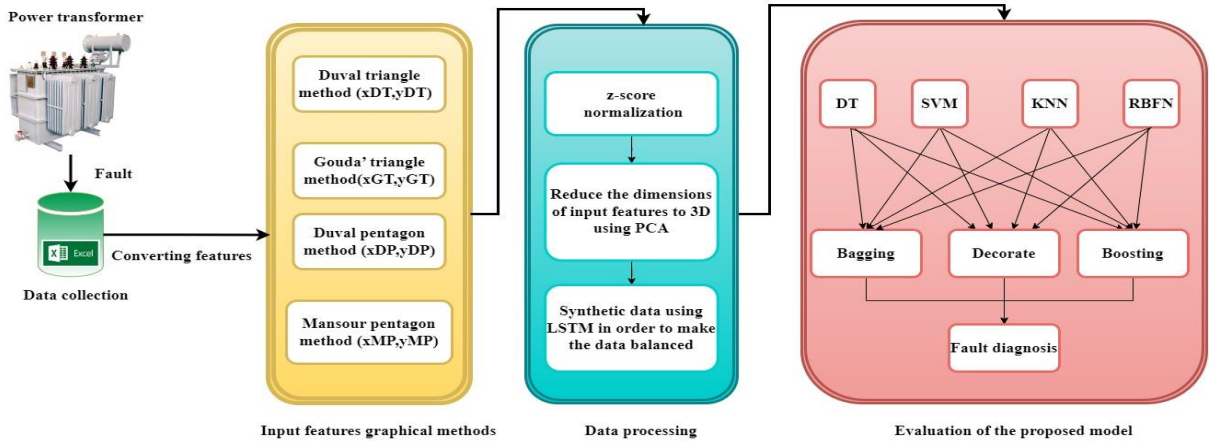


Figure IV.9 Proposed methodology

IV.3.1 DGA Data Collection

Data collection is an important and crucial step in order to evaluate the condition of power transformers, especially since they suffer from a lack of reliable data sources. 407 samples were collected in this chapter to evaluate the proposed method, used in Table II.2.

IV.3.2 Input vectors

The transformation of raw data into input vector features is a pivotal step for leveraging graphical methods such as the Duval Triangle, Duval Pentagon, Mansour Pentagon, and Gouda Triangle. This process is designed to effectively separate overlapping gases, ensuring precision in the analysis. The input features required for the graphical methods are consolidated into a single input vector, as illustrated in Table IV.1.

Table IV.1 Aggregation of features of graphical methods into the proposed input vector

Graphical methods	features equation
Duval Triangle Method	$F1 = x = 100 - \%C_2H_2 - \%CH_4 \cos\left(\frac{\pi}{6}\right) \cot\left(\frac{\pi}{3}\right)$ $F2 = y = \%CH_4 \cos\left(\frac{\pi}{6}\right)$
Gouda triangle method	$F3 = x = 100 - \%R_2 - \%R_3 \cos\left(\frac{\pi}{6}\right) \cot\left(\frac{\pi}{3}\right)$ $F4 = y = \%R_3 \cos\left(\frac{\pi}{6}\right)$
Duval Pentagon Method	$F5 = X(1) = \frac{1}{6}\frac{\sum_{i=0}^{4}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{\frac{1}{2}\sum_{i=0}^{4}(x_i y_{i+1} - x_{i+1} y_i)}$ $F6 = X(2) = \frac{1}{6}\frac{\sum_{i=0}^{4}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)}{\frac{1}{2}\sum_{i=0}^{4}(x_i y_{i+1} - x_{i+1} y_i)}$
Mansour Pentagon Method	$F7 = x_m = \frac{\sum_{i=1}^{n} m_i x_i}{100}$ $F8 = y_m = \frac{\sum_{i=1}^{n} m_i y_i}{100}$
Proposed vector	V=[F1, F2, F3, F4, F5, F6, F7, F8]

IV.3.3 Z-Score Normalization

Z-score normalization, also known as standardization, is a technique used in statistics to transform data into a standard normal distribution. This process is often applied to features or variables in a data set. The goal is to make the data comparable and suitable for analysis, especially in machine learning and statistical modelling.

The formula for calculating the Z-score for a data point x in a dataset with mean μ and standard deviation σ is given by:

$$Z = \frac{x - \mu}{\sigma} \qquad \text{(IV.21)}$$

Z is the Z-score.

x is the individual data point.

μ is the mean of the dataset.

σ is the standard deviation of the dataset.

Figure IV.10 illustrates a visual representation employing box plots to depict the dataset both before and after undergoing Z-score normalization. Visual inspection reveals that the data before normalization shows outliers, a characteristic mitigated in the normalized dataset where values are confined within the ±3 range. This highlights the efficacy of Z-score normalization in enhancing data quality, given its ability to center the data around a mean of 0 with a standard deviation of 1. Notably, this normalization step holds significance for subsequent dimensionality reduction using Principal Component Analysis (PCA) in the next step.

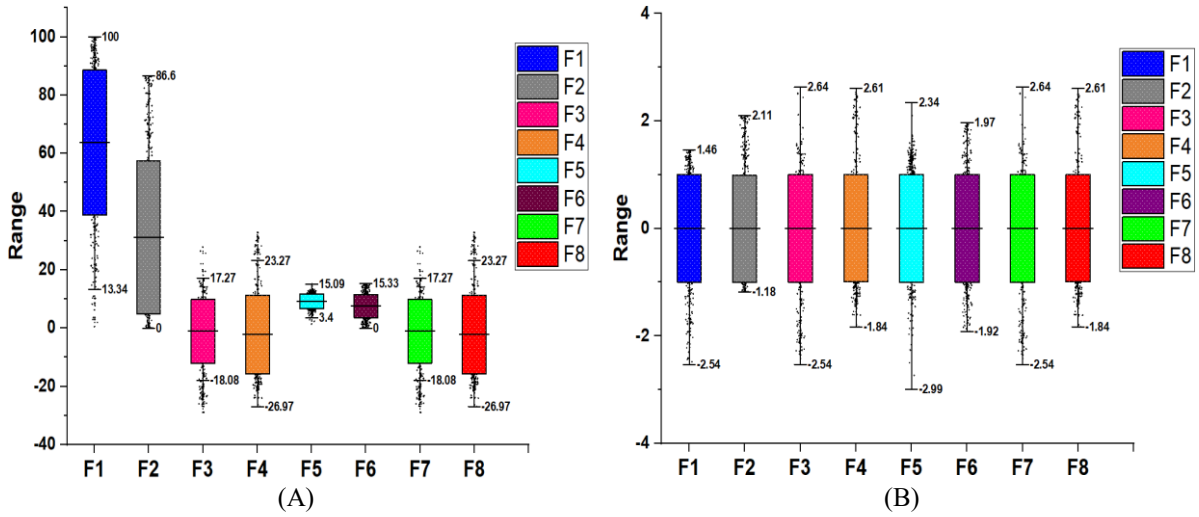


Figure IV.10 Examine outliers using Z-score normalization; (A) Before Z-score normalization; (B) After Z-score normalization

IV.3.4 Dimensional reduction of features

The dimensionality reduction step using PCA is crucial for distilling essential information from eight graphical method features, represented as the vector V= [F1, F2, F3, F4, F5, F6, F7, F8]. This process significantly impacts classification algorithms by simplifying feature spaces, improving computational efficiency, and enhancing accuracy. By reducing the number of features while retaining the most informative ones, PCA helps to alleviate the curse of

dimensionality, mitigating issues such as overfitting and improving the generalization capability of machine learning models. Additionally, PCA aids in creating synthetic data, addressing limitations in dataset size. After the 3D conversion, a new vector V'= [n1, n2, n3] is obtained, highlighting the transformation achieved by PCA. The reduced dimensionality not only facilitates visualization but also contributes to model interpretability. Figure IV.11 represents the feature conversion proposed for graphical methods applied to 3D features, providing a visual insight into the transformative effects of PCA on the datast.

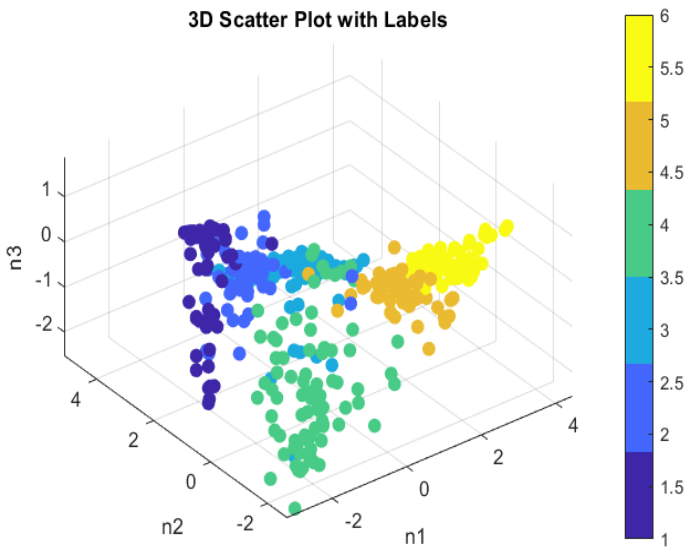


Figure IV.11 PCA Transformation of 3D Features

IV.3.5 Synthetic data creation

To enhance the diagnostic accuracy of classification algorithms, a strategy has been proposed to transform the imbalanced dataset into a balanced one. This involves generating synthetic data to ensure an equitable representation of faults, aligning with the principle of equal opportunity in fault evaluation. Examining Table 1 reveals that fault D1 stands out with the highest count of 99 samples, making it the benchmark. Synthetic data will be introduced to each of the faults, namely PD, D2, T1, T2, and T3, until their counts match that of D1.

The synthetic data generation process employs Long Short-Term Memory (LSTM) networks, chosen for their capability to create multi-modal tabular data. LSTMs are particularly suited to handling a mixture of categorical, numeric, time-series, and textual fields, as they are designed to capture long-range dependencies and maintain context over sequences of data.

Relevant references have been provided [37,38] to offer further explanation of the underlying principles of LSTM networks in this context, including how they effectively capture temporal dependencies and preserve statistical interdependencies. Figure IV.12 from Gretel's report illustrates a comparative analysis of correlation matrices, showcasing the relationships between variables in both the original training set and the synthetic dataset generated by the LSTM. The figure further quantifies the discrepancies between the two sets, providing a measure of the LSTM's ability to preserve the integrity of the data relationships during synthesis.

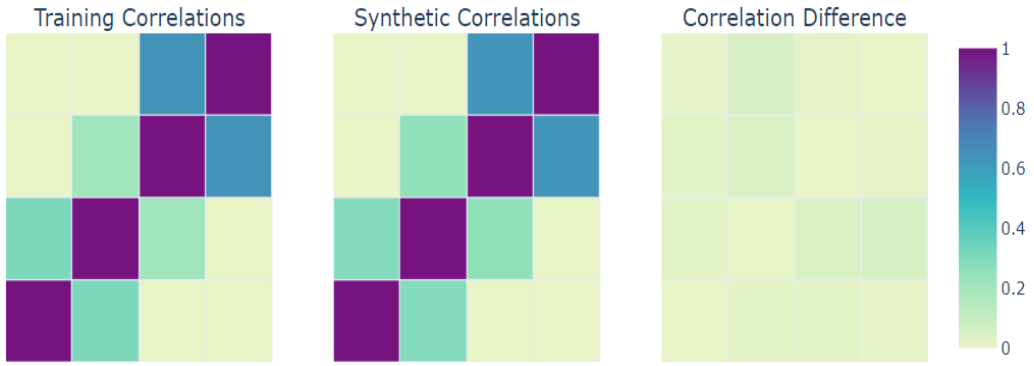


Figure IV.12 Synthetic data generation report

IV.3.6 Hyperparameters for classification algorithms

After generating synthetic data, employing hyperparameters becomes crucial to shape the behavior of machine learning models. The meticulous selection and fine-tuning of these hyperparameters play a pivotal role in bolstering the predictive capabilities of the proposed model. A model endowed with thoughtfully chosen hyperparameters is predisposed to consistently exhibit optimal performance and effectively mitigate the risk of overfitting. Notably, the hyperparameters outlined in Table IV.2 were utilized during the training of the proposed model.

Table IV.2 Hyperparameters of the AI techniques used

Algorithms	Parameters	Values
DT	Collapse Tree	True
	Confidence factor	0.25
	Num folds	3
	seedn	1
SVM	SVM Type	C-SVC
	Cache size	40
	Degree	3
	Kernel Type	radial basis function
KNN	k	2
	Search Algorithm	Linear NN Search
RBFN	Clutering seed	1
	maxIts	-1
	Num clusters	2
	ridage	1.0E-8
Ensemble Techniques		
Bagging	Bag Size Percent	100
	Num Execution Slots	1
	Num iterations	10
Decorate	Artificial Size	1
	Desired Size	15
	Num iterations	50
Boosting	Num Sub Cmtys	3
	Weight Threshold	100
	Num iterations	10

IV.4 Analysis of founded results

IV.4.1 Implementation and analysis of results

The proposed model was implemented using the KNIME analytics platform, an open-source tool tailored for developing data analyses and predictive models. KNIME leverages various programming languages and frameworks, including Python, Java, and Weka, to construct its analytical workflows. To assess the model's performance, a dataset comprising 407 real, unbalanced samples and 187 synthetic samples was employed, resulting in a hybrid balanced dataset of 594 samples. The dataset was split in a 70:30 ratio, with 70% used for training and 30% for testing. The effectiveness of the model was evaluated using the hybrid balanced dataset and compared against the real unbalanced data to provide deeper insights into the model's robustness. Figure IV.13 illustrates the sequential steps involved in implementing the ensemble classification algorithms within the proposed model.

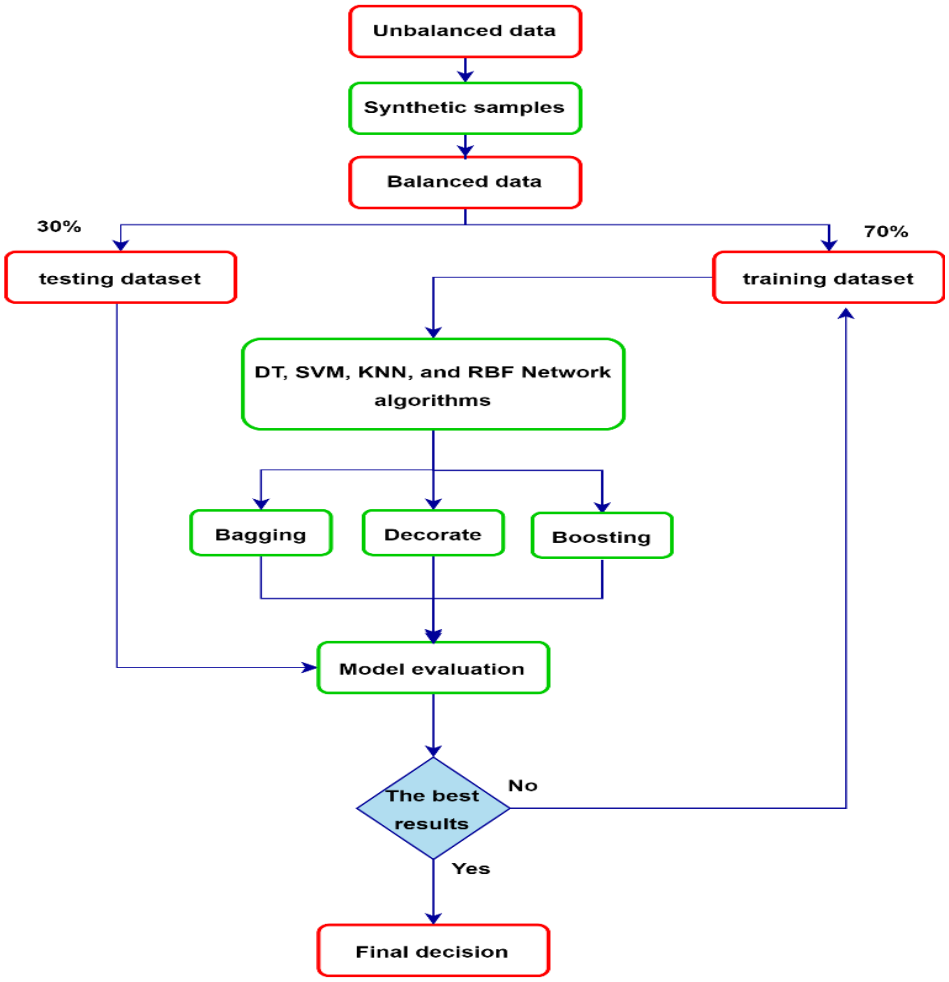


Figure IV.13 Flowchart for implementing the proposed method

Table IV.3 and 4 present key metrics evaluating the performance of the proposed ensemble classification algorithms. The metrics Accuracy, Recall, Precision, Specificity, and F1-Score provide a comprehensive assessment of the model's effectiveness across different algorithms and ensemble techniques, including Base, Bagging, Decorate, and Boosting. These results enable a comparison between real and hybrid datasets.

Boosting consistently emerges as the best-performing ensemble technique across all algorithms and datasets, particularly excelling in accuracy, recall, and F1-score. It shows clear advantages with more complex models like those using hybrid datasets. Bagging also delivers strong performance, frequently ranking as the second-best technique, especially in recall and specificity.

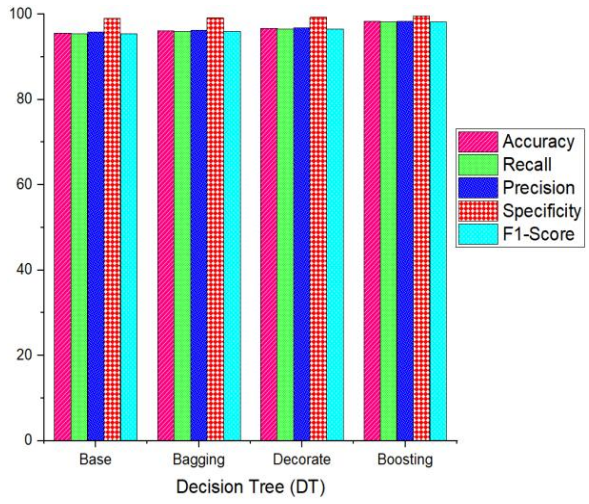
In contrast, Decorate tends to underperform relative to Bagging and Boosting, particularly with real datasets, though it demonstrates moderate improvements with hybrid datasets. Overall, this analysis highlights Boosting as the most robust technique across different algorithms, with Bagging being a strong competitor. While Decorate is less effective, particularly in real datasets, it performs better with hybrid datasets. This comparison provides clear insights into how different ensemble methods impact the performance of classification algorithms based on the dataset type.

Table IV.3 Metrics to evaluate the performance of the proposed ensemble classification algorithms for real datasets

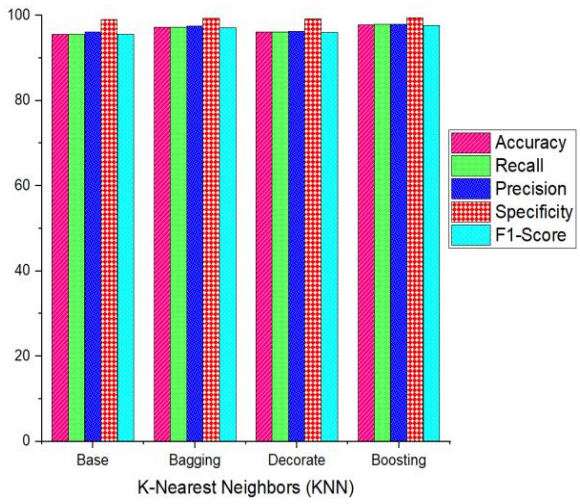
Algorithm	**Ensemble Techniques**	**Model Evaluation (%)**				
		Accuracy	Recall	Precision	Specificity	F1-Score
Decision Tree (DT)	**Base**	86.99	86.64	85.26	97.40	85.71
	Bagging	88.62	85.84	88.76	97.67	86.33
	Decorate	86.99	86.09	86.73	97.34	86.08
	Boosting	**89.43**	**88.22**	**89.66**	**97.81**	**88.44**
K-Nearest Neighbors (KNN)	**Base**	85.37	84.19	84.61	97.01	83.71
	Bagging	87.80	86.97	87.14	97.52	86.65
	Decorate	86.18	84.72	85.32	97.19	83.98
	Boosting	**88.62**	**86.93**	**88.87**	**97.64**	**87.14**
Radial Basis Function Network (RBFN)	**Base**	83.80	81.82	82.33	96.72	81.80
	Bagging	84.55	82.77	85.48	96.79	82.96
	Decorate	80.49	77.61	80.46	95.99	76.99
	Boosting	**87.80**	**84.91**	**88.10**	**97.50**	**85.52**
Support Vector Machine (SVM)	**Base**	76.42	72.32	84.87	94.98	68.95
	Bagging	78.05	74.36	85.69	95.33	72.04
	Decorate	73.98	68.80	83.37	94.48	63.27
	Boosting	**79.67**	**76.03**	**86.04**	**95.70**	**74.71**

Table IV.4 Metrics to evaluate the performance of the proposed ensemble classification algorithms for hybrid dataset

Algorithm	Ensemble Techniques	Model Evaluation (%)				
		Accuracy	Recall	Precision	Specificity	F1-Score
DT	**Base**	95.53	95.52	95.83	99.11	95.55
	Bagging	96.09	96.06	96.18	99.22	96.07
	Decorate	96.65	96.64	96.87	99.33	96.66
	Boosting	**98.32**	**98.31**	**98.38**	**99.67**	**98.33**
KNN	**Base**	95.53	95.50	96.07	99.11	95.59
	Bagging	97.21	97.18	97.50	99.44	97.23
	Decorate	96.09	96.07	96.32	99.22	96.12
	Boosting	**97.77**	**97.94**	**97.95**	**99.55**	**97.78**
RBFN	**Base**	94.41	94.43	94.45	98.89	94.44
	Bagging	95.53	95.54	96.70	99.11	95.59
	Decorate	94.41	94.43	94.46	98.89	94.42
	Boosting	**97.21**	**97.21**	**97.35**	**99.44**	**97.24**
(SVM	**Base**	96.09	96.09	96.23	99.22	96.10
	Bagging	95.53	95.54	95.61	99.11	95.55
	Decorate	94.41	94.43	94.53	98.89	94.40
	Boosting	**97.77**	**97.78**	**97.92**	**99.55**	**97.78**



(A)



(B)

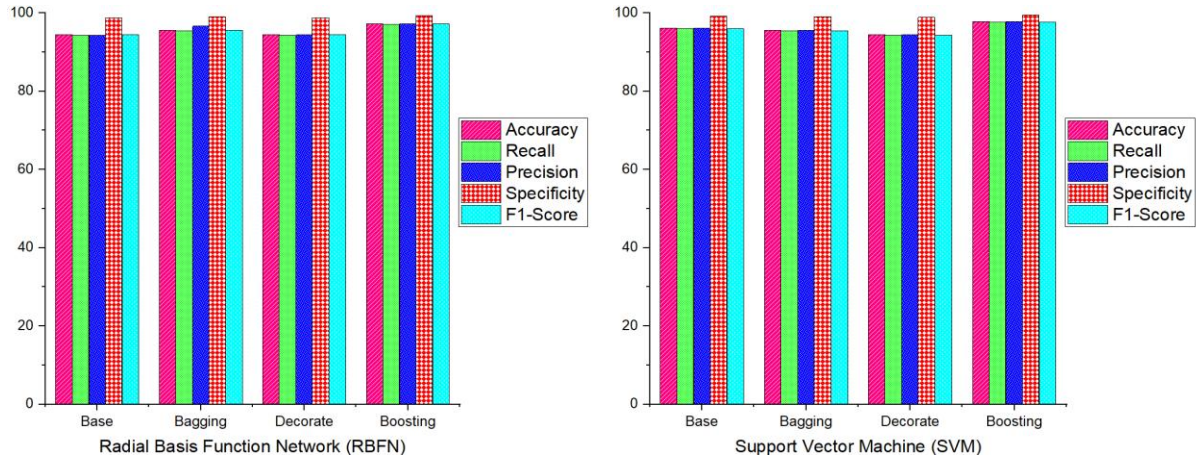


Figure IV.14. Performance chart of ensemble classification algorithms for different evaluation metrics for real datasets

Figure IV.14 provides a visual representation of the evaluation metrics for different algorithms. In Fig.10A, dedicated to the decision tree algorithm, boosting exhibits superior performance compared to Base and other ensemble techniques Bagging and Decorate across multiple metrics, including Accuracy, Recall, Precision, Specificity, and F1-Score, with respective values of 98.32%, 98.31%, 98.38%, 99.67%, and 98.33%.

Moving to Fig.10B, which focuses on the KNN algorithm, Boosting again outshines both the standalone algorithm and other ensemble techniques Bagging and Decorate in various metrics: Accuracy 97.77%, Recall 97.94%, Precision 97.95%, Specificity 99.55%, and F1-Score 97.78%.

Fig.10C presents the performance metrics for the RBFN algorithm, reaffirming Boosting's superiority over Base, Bagging, and Decorate in Accuracy 97.21%, Recall 97.21%, Precision 97.35%, Specificity, and F1-Score 97.24%.

Similarly, Fig.10D illustrates the performance evaluation for the SVM algorithm, with Boosting outperforming Base, Bagging, and Decorate across all metrics: Accuracy 97.77%, Recall 97.78%, Precision 97.92%, Specificity 99.55%, and F1-Score 97.78%.

The consistent outcomes emphasize Boosting's superiority over other ensemble classification algorithms, including Base, Bagging, and Decorate, across diverse performance metrics. For a deeper analysis and comparison of the results, Table IV.3 provides a comprehensive overview of all Boosting outcomes. This reinforces the inclination towards Boosting in ensemble classification algorithms, as indicated by the evaluated performance criteria.

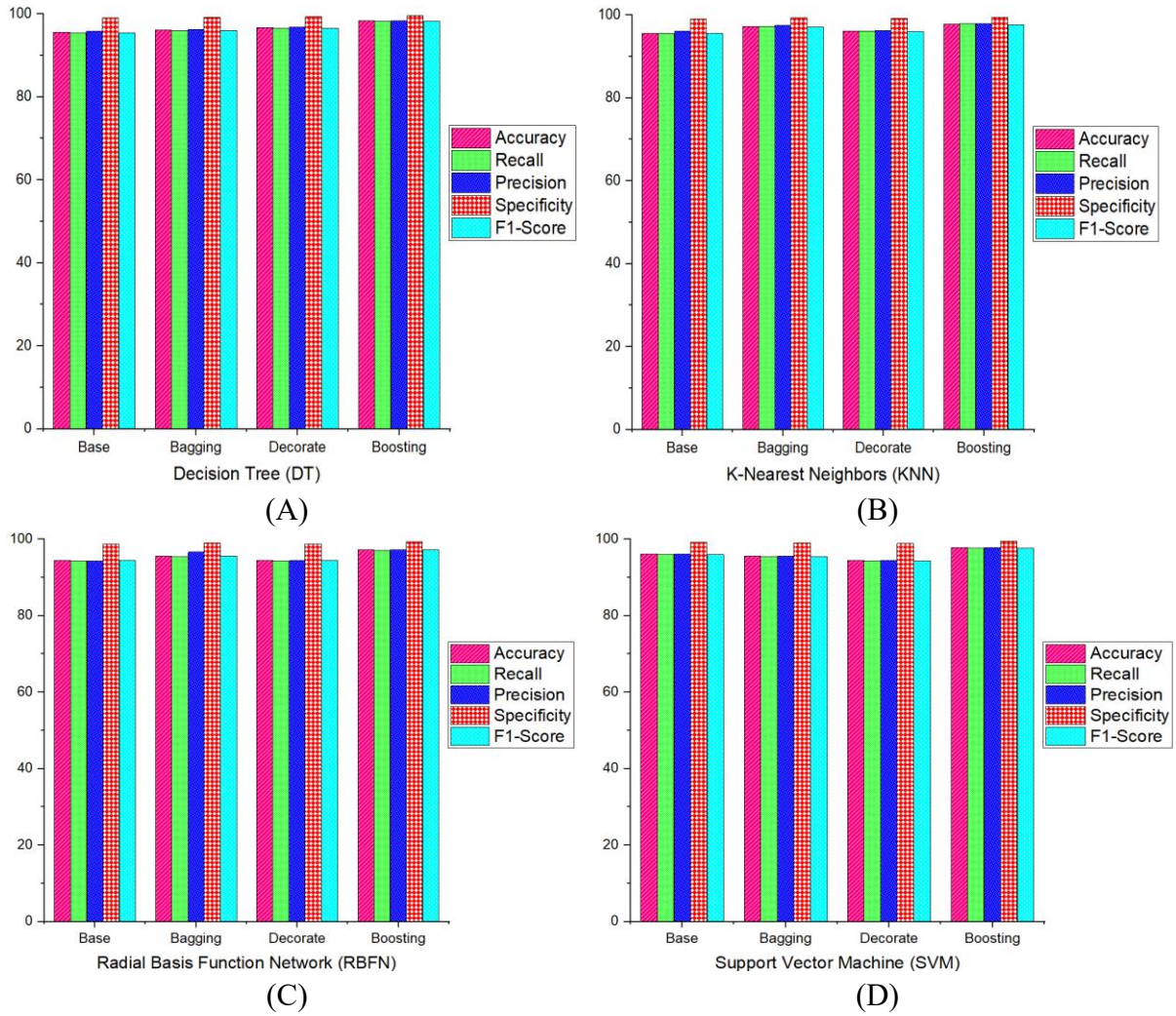


Figure IV.15 Performance chart of ensemble classification algorithms for different evaluation metrics for hybrid dataset

A thorough analysis of Table IV.5 and Figure IV.16 clearly demonstrates that the use of balanced hybrid data significantly outperforms real unbalanced data, with Boosting exhibiting the highest performance when paired with the Decision Tree (DT) algorithm in both cases. The algorithm achieved outstanding results with an accuracy of 98.32%, recall of 98.31%, precision of 98.38%, specificity of 99.67%, and an F1 score of 98.33% when applied to hybrid data. Notably, these values not only surpassed but substantially exceeded those of all other ensemble classification techniques, providing strong empirical evidence both visual and tabular that highlights the superior performance of Boosting when combined with the DT algorithm. This underscores its position as the optimal choice among ensemble classification strategies.

Table IV.5 A comprehensive overview of all boosting results

Boosting	Acc	Rec	Pre	Spe	F1-S
Unbalanced real data evaluation (%)					
DT	89.43	88.22	89.66	97.81	88.44
KNN	88.62	86.93	88.87	97.64	87.14
RBFN	87.80	84.91	88.10	97.50	85.52
SVM	79.67	76.03	86.04	95.70	74.71
Balanced hybrid data evaluation (%)					
DT	98.32	98.31	98.38	99.67	98.33
KNN	97.77	97.94	97.95	99.55	97.78
RBFN	97.21	97.21	97.35	99.44	97.24
SVM	97.77	97.78	97.92	99.55	97.78

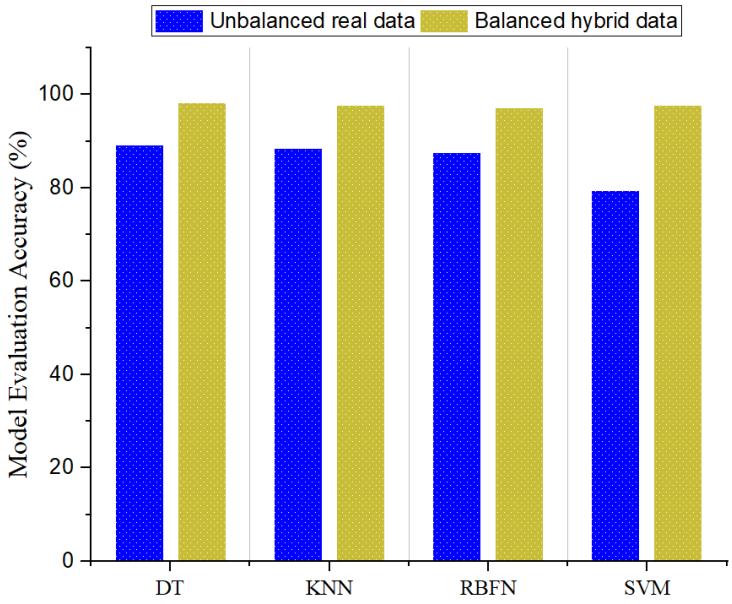


Figure IV.16 Comparison of the best results of boosting between balanced and unparallel data

Figure IV.17 presents the confusion matrices for the best results achieved by Boosting across four different algorithms DT, KNN, RBFN, and SVM labeled as (A), (B), (C), and (D), respectively. Each matrix compares predicted and actual classes for six distinct faults (D1, D2, PD, T1, T2, and T3), with the values representing correct and incorrect predictions, and percentages reflecting the classification accuracy for each fault.

When using hybrid data, which balances real and synthetic samples, Boosting significantly improves the classification accuracy of all algorithms. However, the Decision Tree (DT) algorithm (A) clearly outperforms the others. DT achieves near-perfect results, with 100% correct classification on four out of six faults, with two misclassifications in "D1" and one in "T1." This highlights the superior capability of the DT algorithm to effectively manage hybrid data and consistently maintain high performance across various fault types.

In comparison, KNN (B) also performs well with hybrid data, achieving 100% accuracy on three faults, but it experiences one misclassification in "D1," two in "PD," and one in "T3." Similarly, RBFN (C) and SVM (D) benefit from hybrid data but show more frequent misclassifications. Notably, SVM (D) demonstrates the lowest performance among the four algorithms, particularly in the classification of the "T3" fault.

Overall, the results confirm the effectiveness of Boosting in conjunction with hybrid data, significantly enhancing classification performance across all algorithms due to the balanced synthetic data generated using LSTM. Additionally, the benefits of dimensionality reduction through Principal Component Analysis (PCA) are evident in the improved interpretability and generalization of the results. By reducing the feature space while preserving essential information, PCA contributes to a clearer understanding of the performance metrics, offering insights into the underlying factors driving the superior performance of Boosting when combined with the Decision Tree (DT) algorithm. Consequently, dimensionality reduction enhances the robustness and reliability of the results, further reinforcing the validity of the conclusions drawn from this analysis.

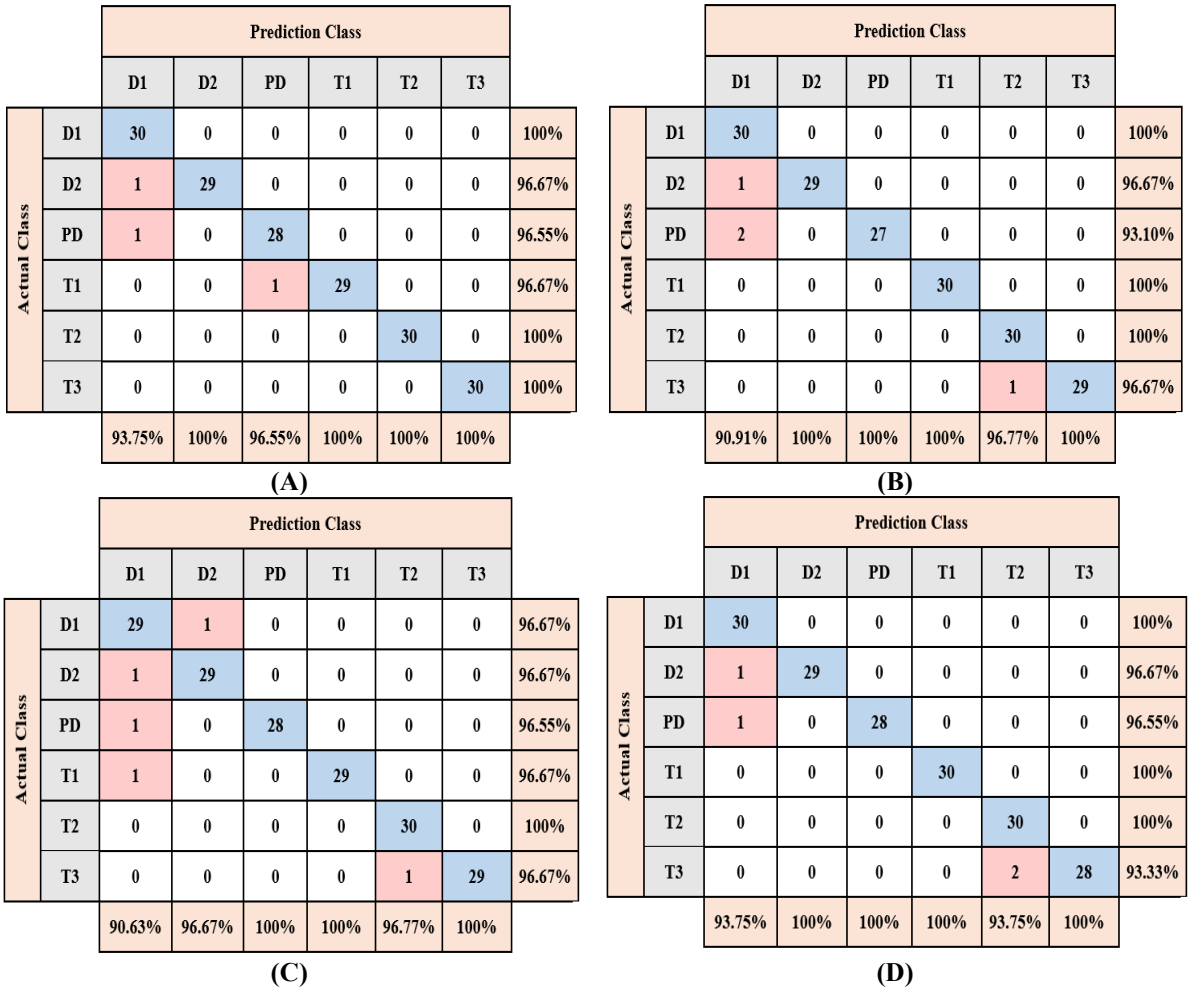


Figure IV.17 Confusion matrix for the best boosting results for DT(A), KNN(B), RBFN(C), and SVM(D)

IV.4.2 Validation and comparison with existing methods

For validation and comparison, a new dataset from the literature, comprising 89 samples, was employed to assess the efficiency of the proposed model and to compare it with current methods. These 89 samples, as detailed in [22], serve as a basis for assessing the performance of the proposed model and for comparative analysis with established methodologies.

Table IV.6 presents a comprehensive comparison between the results achieved by the proposed model and existing methodologies across various categories, namely DGA Ratios methods, DGA graphical methods, Intelligent methods, and hybrid methods. The assessment is based on the accuracy pertaining to specific faults (PD, D1, D2, T1, T2, and T3) as well as the overall accuracy. Notably, the Proposed Boosting Ensemble techniques exhibited superior performance in contrast to the Current methods. Specifically, for Decision Tree, K-Nearest Neighbors, Radial Basis Function Network, and Support Vector Machine, the Proposed Boosting Ensemble demonstrated overall accuracies of 95.51%, 94.38%, 92.13%, and 93.26%, respectively. This underscores the effectiveness of the proposed model across a spectrum of fault types and its ability to outperform existing methodologies in multiple scenarios.

Figure IV.18 illustrates a comprehensive comparison between the optimal outcomes obtained through the Proposed Boosting Ensemble techniques and the best results achieved by various categories of existing methods, including Ratios methods, graphical methods, Intelligent methods, and hybrid methods. Notably, the Boosting Decision Tree algorithm stands out by achieving the highest diagnostic accuracy at 95.51%, surpassing all other existing methods. Specifically, it outperforms the clustering method with 83.15%, the Gouda Triangle Method with 78.65%, the Trees Based Learning algorithm with 92.13%, and the Combined technique N°2 with 80.90%. The superiority of Boosting Decision Tree can be attributed to its utilization of graphical methods as vectors, which are then subjected to dimensionality reduction techniques such as PCA, thereby enhancing the discrimination power of the algorithm. Moreover, the boosting property of the algorithm, which iteratively improves the model's performance by focusing on misclassified instances, synergistically overlays with the decision tree's ability to capture complex relationships in the data, resulting in superior classification accuracy.

Table IV.6 Validation and comparison with existing methods

		Accuracy (%)						
	Existing methods	PD	D1	D2	T1	T2	T3	Total
DGA Ratios methods	Modified IEC 60599 [87]	87.50	61.54	84.21	84.62	57.14	89.66	80.90
	Three Ratios Technique [71]	75	46.15	89.47	69.23	71.43	93.10	78.65
	Clustering Method [83]	100	69.23	84.21	92.31	42.86	89.66	83.15
	Key Gases with Gas Ratios [85]	75	76.92	84.21	84.62	57.14	82.76	79.78
DGA Graphical methods	Duval triangle method [130]	75	53.85	78.95	53.85	57.14	89.66	73.03
	Duval Pentagon Method [77]	62.50	15.38	63.16	76.92	57.14	89.66	66.29
	Mansour Pentagon Method [79]	87.50	61.54	63.16	53.85	14.29	82.76	66.29
	Gouda Triangle Method [80]	87.50	69.23	84.21	53.85	57.14	93.10	78.65
Intelligent methods	CSUS ANN method [152]	75	38.46	68.42	92.31	14.29	62.07	61.80
	Trees Based Learning [132]	100	92.31	89.47	92.31	85.71	93.10	92.13
	Conditional probability[153]	87.50	53.85	89.47	61.54	57.14	93.10	78.65
	Hybrid DGA Method [154]	75	53.85	57.89	84.62	57.14	89.66	73.03

Hybrid methods	Combined technique N°2 [86]	87.50	69.23	89.47	61.54	71.43	89.66	80.90
	Combined technique N°3 [86]	87.50	69.23	89.47	61.54	71.43	82.76	78.65
	Novel CombinedTechniques [155]	87.50	53.85	84.21	76.92	00.00	86.21	73.03
Proposed Boosting Ensemble techniques	Decision Tree [156]	100	100	89.47	100	87.50	96.43	95.51
	K-Nearest Neighbors [156]	100	92.31	94.74	100	87.50	92.86	94.38
	Radial Basis Function Network [156]	100	92.31	78.95	92.31	100	96.43	92.13
	Support Vector Machine [156]	100	92.31	89.47	100	87.50	92.86	93.26

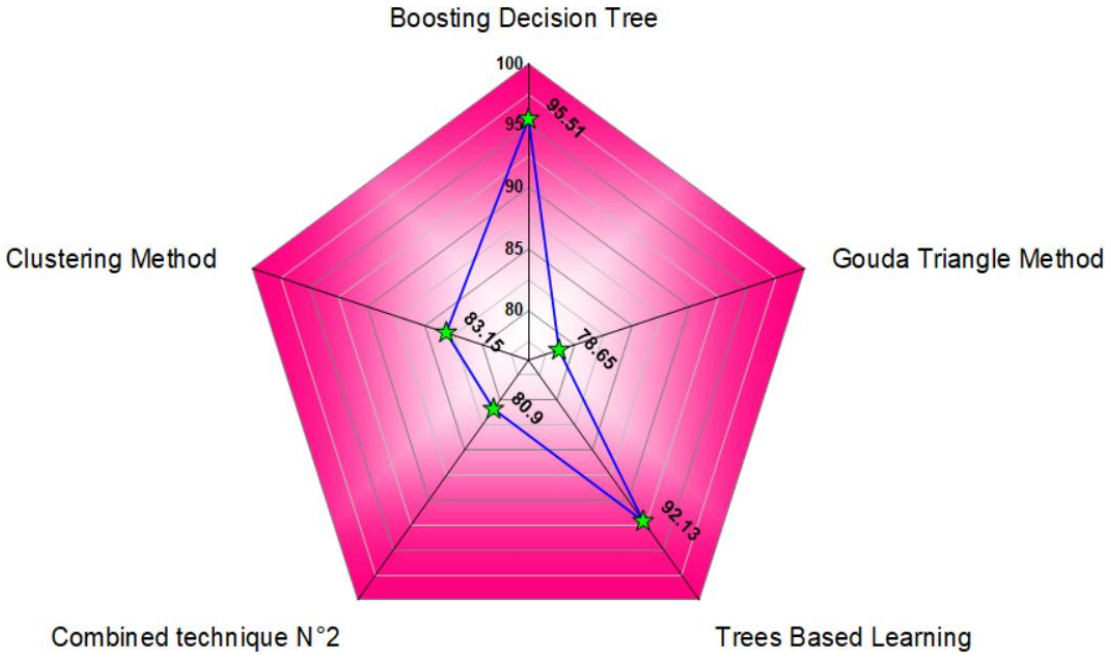


Figure IV.18 Compare the best results for each category of previous methods with the best proposed method

IV.5 Conclusion

This chapter presented a groundbreaking diagnostic approach for power transformers, integrating machine learning algorithms with ensemble techniques to address key challenges in fault analysis. By leveraging Principal Component Analysis (PCA) for dimensionality reduction and applying ensemble methods such as Bagging, Decorate, and Boosting in conjunction with algorithms like Decision Trees (DT), k-Nearest Neighbor (KNN), Radial Basis Function Networks (RBFN), and Support Vector Machines (SVM), the proposed method achieved significant improvements in diagnostic accuracy. The Boosting Ensemble technique, in particular, outperformed conventional ratio-based, graphical, intelligent, and hybrid methods, demonstrating its capacity to enhance data consistency and address the limitations of prior approaches.

The findings not only validate the proposed model's effectiveness but also offer a roadmap for future research. Areas for further exploration include advanced feature engineering techniques and their integration with machine learning, deep learning, and ensemble strategies. This chapter establishes a strong foundation for advancing DGA-based diagnostic models, driving continued progress in the field of power transformer fault analysis.

Chapter V. New technique based on dynamic axes for DGA interpretation

V.1 Introduction

Power transformers play a crucial role in power systems, and dissolved gas analysis (DGA) is essential for fault detection and maintenance. While traditional DGA Ratios methods provide valuable insights, they often struggle to accurately interpret faults outside predefined diagnostic codes, leading to unreliable results. Graphical methods have been developed to address this issue, but they rely on fixed axes and angles, limiting their adaptability to diverse fault scenarios and multi-source data.

Recently, artificial intelligence (AI) has been integrated into DGA techniques to improve fault diagnosis. However, the complexity and data-intensive nature of AI models can be a barrier to their widespread adoption, particularly among engineers.

This chapter proposes a new graphical technique for DGA interpretation that uses dynamic axes in a circular form. This method aims to offer more reliable and adaptable fault diagnosis, overcoming the limitations of existing approaches and improving the accuracy of DGA interpretation in power system monitoring.

V.2 Analysis of Optimization Algorithms Applied

V.2.1 Gray Wolf Optimization

GWO algorithm stands out as a pioneering swarm intelligence algorithm introduced by Mirjalili et al.[157]. It takes its inspiration from the intricate dynamics of grey wolves' predatory behavior and social interactions in the wild. In this algorithm, four distinct roles emulate the natural social hierarchy: alpha (α), beta (β), and delta (δ) represent the top three wolves in terms of fitness, while the remaining wolves are categorized as omega (ω).

With the GWO algorithm, the optimization process is steered by the α, β, and δ wolves, while the ω wolves trail behind them [158]. Figure V.1 illustrates the distribution of the different wolf pack levels.

The mathematical representation for the behavior of grey wolves as they encircle their prey is expressed as follows [159]:

$$\vec{D} = \left|C.\vec{X}_P(t) - \vec{X}(t)\right| \quad \text{(V.1)}$$

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A}.\vec{D} \quad \text{(V.2)}$$

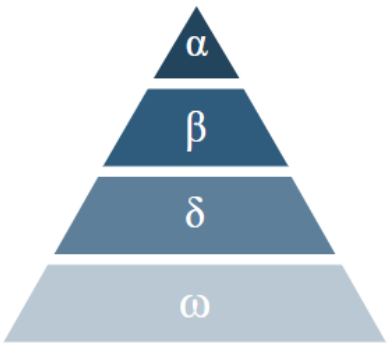


Figure V.1 Hierarchy of grey wolf

Here, t denotes the current iteration, and the coefficient vectors A and C are derived as:

$$\vec{A} = 2.\vec{a}.\vec{r}_1 - \vec{a}_1 \quad \text{(V.3)}$$

$$\vec{C} = 2.\vec{r}_2 \quad \text{(V.4)}$$

The vector a convergence factor linearly decreasing from 2 to 0 throughout the iterations, while $\vec{r_1}$ and $\vec{r_2}$ are random vectors in the range (0,1). The hunting process is mathematically governed by:

$$\begin{cases} \vec{D}_\alpha = \left|\vec{C}_1.\vec{X}_\alpha - \vec{X}\right| \\ \vec{D}_\beta = \left|\vec{C}_2.\vec{X}_\beta - \vec{X}\right| \\ \vec{D}_\delta = \left|\vec{C}_3.\vec{X}_\delta - \vec{X}\right| \end{cases} \quad \text{(V.5)}$$

$$\begin{cases} \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1.\vec{D}_\alpha \\ \vec{X}_2 = \vec{X}_\beta - \vec{A}_2.\vec{D}_\beta \\ \vec{X}_3 = \vec{X}_\delta - \vec{A}_3.\vec{D}_\delta \end{cases} \quad \text{(V.6)}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad \text{(V.7)}$$

V.2.2 Whale Optimization Algorithm

The WOA algorithm, introduced by Mirjalili and Lewis [160], is a recently developed metaheuristic algorithm that effectively balances exploration and exploitation, surpassing other optimization methods currently available. The WOA is derived from the observed foraging patterns of humpback whales in their natural habitat. Humpback whales can be nourished using bubble nets or by maneuvering along a 9-shaped trajectory, as depicted in Figure V.2 [161].

Initially, humpback whales encircle their prey, which can be mathematically described as:

$$\vec{D} = \left|C.\overrightarrow{X^*}(t) - \vec{X}(t)\right| \quad \text{(V.8)}$$

$$\vec{X}(t+1) = \overrightarrow{X^*}(t) - \vec{A}.\vec{D} \quad \text{(V.9)}$$

where t is the current iteration, $(X\hat{}*)\vec{}$ is the best solution so far, $X\vec{}$ is the position vector, and $A\vec{}$ and $C\vec{}$ are coefficient vectors given by:

$$\vec{A} = 2.\vec{a}.\vec{r}_1 - \vec{a} \quad \text{(V.10)}$$

$$\vec{C} = 2.\vec{r}_2 \qquad \text{(V.11)}$$

Here, $\vec{a}$ decreases linearly from 2 to 0, while $\vec{r}_1$ and $\vec{r}_2$ are random vectors in [0,1]. The shrinking encircling behavior is controlled by:

$$\vec{a} = 2\left(1 - \frac{t}{T}\right) \qquad \text{(V.12)}$$

where T is the maximum number of iterations.

In the exploitation phase, the algorithm simulates the whales' spiral-shaped movement around their prey using the following equations:

$$\vec{X}(t+1) = D'.e^{bl}.\cos(2\pi l) + \overrightarrow{X^*}(t) \qquad \text{(V.13)}$$

A probabilistic switch is applied:

$$\vec{X}(t+1) = \begin{cases} \overrightarrow{X^*}(t) - \vec{A}.\vec{D}, & p < 0 \\ D'.e^{bl}.\cos(2\pi l) + \overrightarrow{X^*}(t), & p \geq 0 \end{cases} \qquad \text{(V.14)}$$

where b is a constant, and l and p are random values between 0 and 1.

In the exploration phase, when $|\vec{A}|>1$, the WOA shifts to a random search strategy to avoid local optima, using a random position vector:

$$\vec{D} = \left|C.\vec{X}_{rand} - \vec{X}\right| \qquad \text{(V.15)}$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A}.\vec{D} \qquad \text{(V.16)}$$

The random position vector, denoted as $\vec{X}$rand, represents a random whale.

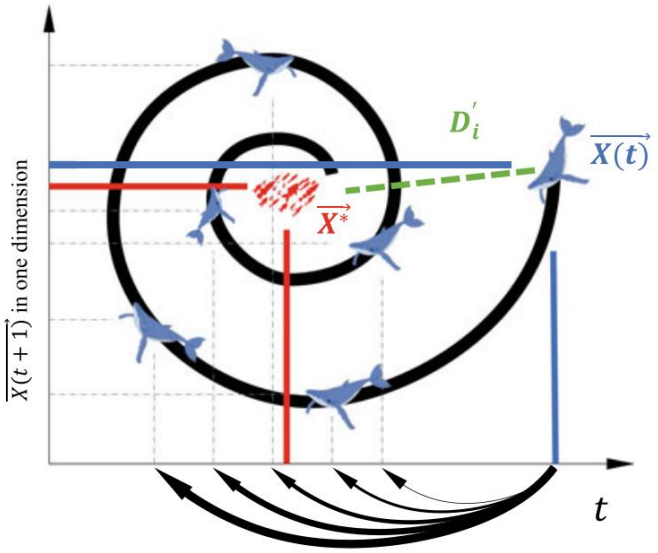


Figure V.2 Working mechanism considering exploration and exploitation of WOA

V.3 Proposed Circular Graphical Method

V.3.1 Addressing the Shortcomings of Conventional Approaches

Dissolved Gas Analysis (DGA) remains one of the most widely employed diagnostic tools for monitoring and detecting early-stage faults in power transformers [162]. To accurately identify the type of fault in power transformers, various interpretation methods for DGA results have been developed, including ratio-based and graphical techniques [163]. Over the years, advancements in DGA techniques have focused on improving the safety and reliability of transformer diagnostics, resulting in modern methods that address many of the limitations of traditional approaches.

Among the conventional DGA techniques are the Dornenburg method [112], the Roger method [64], and the IEC method [113]. These older methods, while foundational, exhibit significant shortcomings, particularly in diagnosing malfunctions that fall outside predefined codes, leading to unreliable interpretations. To overcome these limitations, the Duval method [130] was introduced. While effective and relatively simple, the Duval method is restricted to analyzing three gases (CH_4, C_2H_2, and C_2H_4). This narrow scope can result in diagnostic inaccuracies, especially in cases involving low-temperature faults or partial discharges where gases like C_2H_6 and H_2, which are not represented in the triangular graph, play a critical role. Although newer versions of the Duval triangle have been proposed to address these gaps [114], questions remain regarding their reliability, particularly for complex fault scenarios.

Modern advancements in DGA include the pentagon methods developed by Duval and Mansour [77, 79]. These methods represent an evolution of the Duval Triangle, incorporating five combustible gases to improve diagnostic precision. However, these approaches are not without their challenges. One major limitation is their reliance on fixed axes and angles, which restrict their applicability across diverse fault scenarios. Additionally, these methods lack adaptability to varying datasets, as they tend to perform well only when applied to the specific datasets used during their development. Consequently, their reliability can diminish when applied to data from different sources [164].

The clustering method [83], which relies on gas concentration percentage limits, has demonstrated improved efficiency compared to older techniques. However, it is vulnerable to faults that fall outside its established limits, rendering it insufficient to meet modern safety standards for power transformers. Similarly, the three-ratio technique [71] offers useful insights but is constrained to only three ratios, limiting its effectiveness for diagnosing specific faults. Deviations from predefined ratio limits can lead to inaccurate conclusions.

The Gouda triangle method, introduced in [80], was designed as an extension of the Duval triangle. Despite its advancements, this method exhibits significant overlap in thermal fault zones, which compromises its ability to provide precise fault differentiation. Additionally, the Cartesian graphical method [82], recognized for its superiority over prior techniques, introduces new gas ratios plotted on a Cartesian (x, y) axis. While this method offers enhanced diagnostic accuracy, it shares the limitations of the pentagon and triangle methods, including complexity in selecting gas ratios and a dependence on fixed axes and angles that fail to accommodate all potential fault cases.

V.3.2 A Novel Solution for Enhanced Fault Separation and Diagnosis

In describing the graphical Circle method, reliance is placed on the percentages of the five combustible gases [%H_2, %CH_4, %C_2H_2, %C_2H_4, %C_2H_6] that are analyzed by DGA. These gases, in turn, represent a pentagonal parameter, which is depicted by a binary parameter (x, y), as shown in Figure V.3.

Figure IV.3(A) represents the five combustible gases in the parameter (x, y) with fixed angles for each (θ1, θ2, θ3, θ4, θ5). This representation is very similar to the representation found in the Duval Pentagon method [77], while Figure V.3(B) represents the five combustible gases on the moving axis (x, y) with variable angles (θ1, θ2, θ3, θ4, θ5). The variable and random selection of angles allows for identifying different spreads of samples, which allows for choosing the best spread that achieves separation between samples for the proposed method.

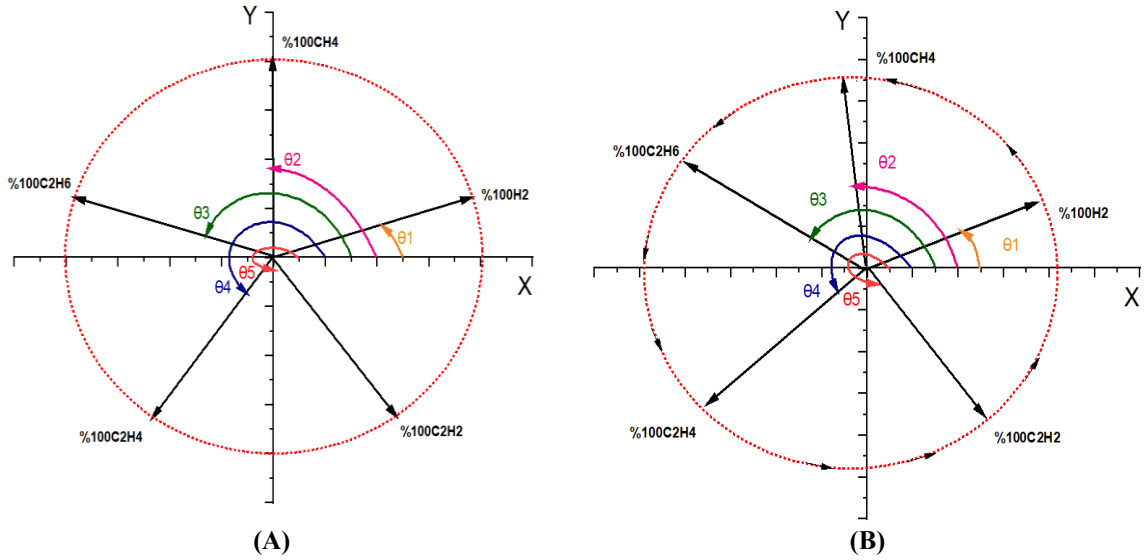


Figure V.3 Representation of the five combustible gases on the parameter coordinates (x, y) with fixed and moving axes

The parameters of the five gases are projected into binary parameters, resulting in the conversion of pentagonal features into binary features represented by x and y. The two equations illustrating this process are as follows:

For x_i, the individual components are defined as follows:

$$\begin{aligned} x_{i1} &= \%H_2 \cos(\theta 1) \\ x_{i2} &= \%CH_4 \cos(\theta 2) \\ x_{i3} &= \%C_2H_6 \cos(\theta 3) \\ x_{i4} &= \%C_2H_4 \cos(\theta 4) \\ x_{i5} &= \%C_2H_2 \cos(\theta 5) \end{aligned} \qquad \text{(V.17)}$$

For y_i, the individual components are defined as follows:

$$\begin{aligned} y_{i1} &= \%H_2 \sin(\theta 1) \\ y_{i2} &= \%CH_4 \sin(\theta 2) \end{aligned}$$

$$y_{i3} = \%C_2H_6 \sin(\theta 3) \quad \text{(V.18)}$$
$$y_{i4} = \%C_2H_4 \sin(\theta 4)$$
$$y_{i5} = \%C_2H_2 \sin(\theta 5)$$

The combined equations for x_i and y_i are:

$$x_i = x_{i1} + x_{i2} + x_{i3} + x_{i4} + x_{i5} \quad \text{(V.19)}$$
$$y_i = y_{i1} + y_{i2} + y_{i3} + y_{i4} + y_{i5} \quad \text{(V.20)}$$

These equations effectively project the gas parameters into binary features x and y, allowing for further analysis and interpretation within this transformed coordinate system.

Figure V.4 illustrates the distribution of all samples on the x-y axis using random and variable angles. Various colors and shapes are employed to represent different power transformer faults. The figure demonstrates how altering angles impacts the distribution of fault samples. Notably, whenever the angles are are modified: (θ1, θ2, θ3, θ4, θ5) A, (θ1, θ2, θ3, θ4, θ5) B, (θ1, θ2, θ3, θ4, θ5) C and (θ1, θ2, θ3, θ4, θ5) D are different, the arrangement of samples also changes accordingly.

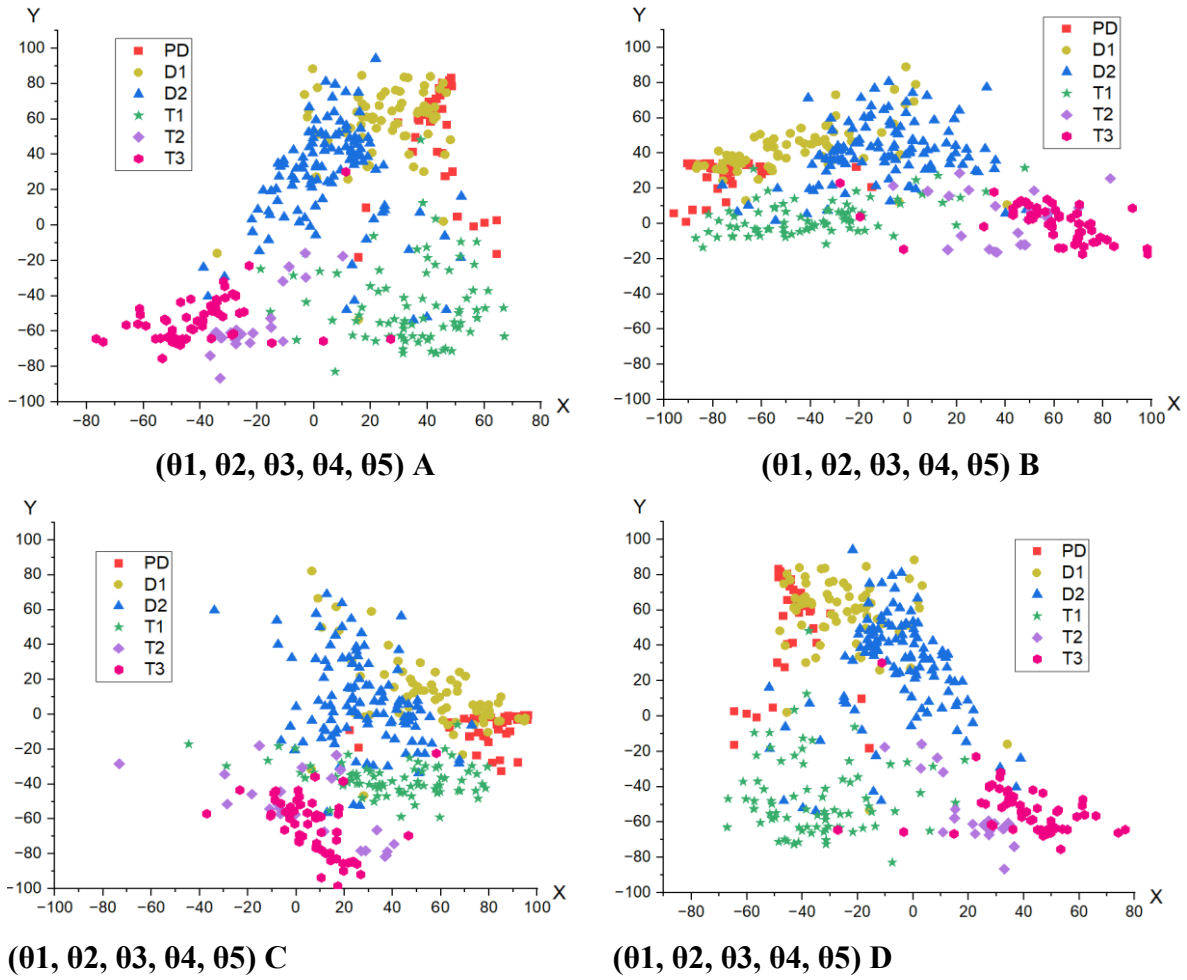


Figure V.4 Distribution of samples in the axis (x, y)

V.4 Analysis of Sample Divergence and Convergence Behavior

Understanding the behavior of divergence and convergence is crucial in various fields. Divergence indicates that samples are distinct and well-separated, while convergence indicates that sample points within the same group are tightly clustered together. This divergence and convergence of samples are formulated in terms of the center point.

The coordinated center is the point that represents the average position of a set of points. It can be calculated by adding up all the x-coordinates and y-coordinates of the points in the set and then dividing by the total number of points. The coordinated center is represented by the (xc, yc) coordinates, where xc is the average of all the x-coordinates and yc is the average of all the y-coordinates.

V.4.1 Center Coordinates Points

The equations for calculating the centroid or coordinated center of a set of points for each fault j are:

$$xc_j = \frac{\sum x_{ij}}{N_j} \quad \text{(V.21)}$$

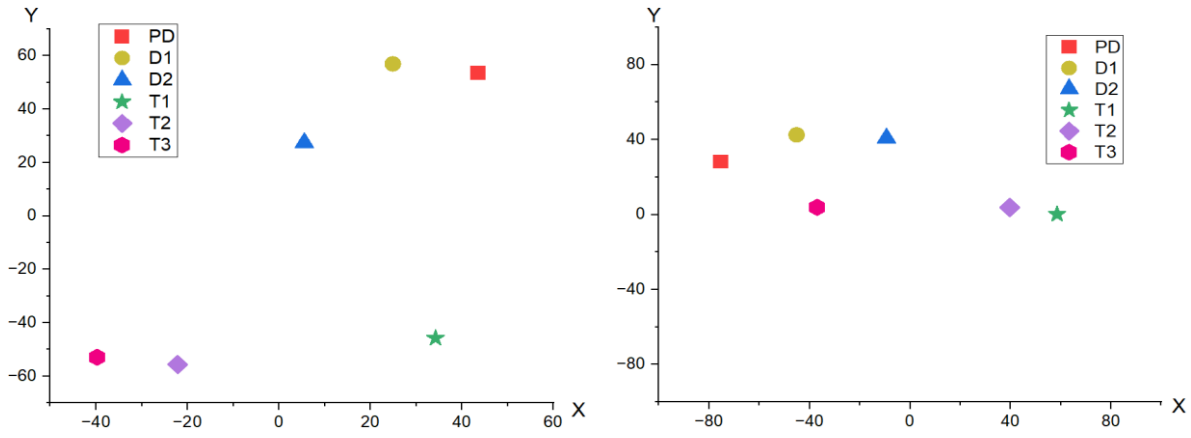
$$yc_j = \frac{\sum y_{ij}}{N_j} \quad \text{(V.22)}$$

xc_j and yc_j represent the coordinates of the centroid for each fault j.

Nj is the total number of points for each fault j.

The summations $\sum xij$ and $\sum yij$ are taken over all the points associated with fault j.

j varies from 1 to 6 for different types of faults (PD=1, D1=2, D2=3, T1=4, T2=5, T3=6), i denotes the total number of faults considered. Figure V.5 shows the center coordinates of each fault with different angles.



(θ1, θ2, θ3, θ4, θ5) A **(θ1, θ2, θ3, θ4, θ5) B**

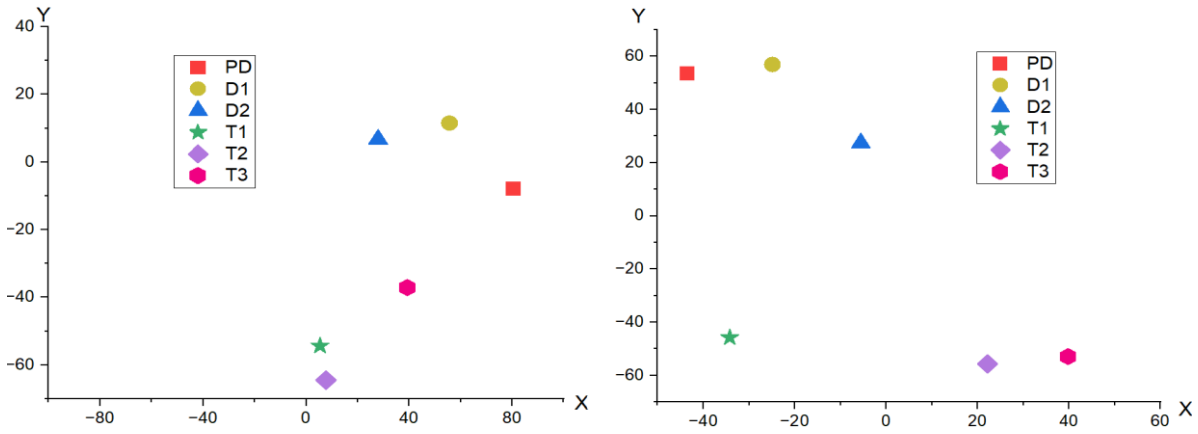


Figure V.5 Distribution of the Coordinating Center of each fault

V.4.2 Convergence Coefficient of The Coordinate Center

The coefficient of convergence from the coordinate center is a powerful metric for assessing the distribution of a set of points in relation to its center. A smaller value of convergence indicates that the points are more closely clustered around their center, while a larger value indicates that the points are more widely spread out.

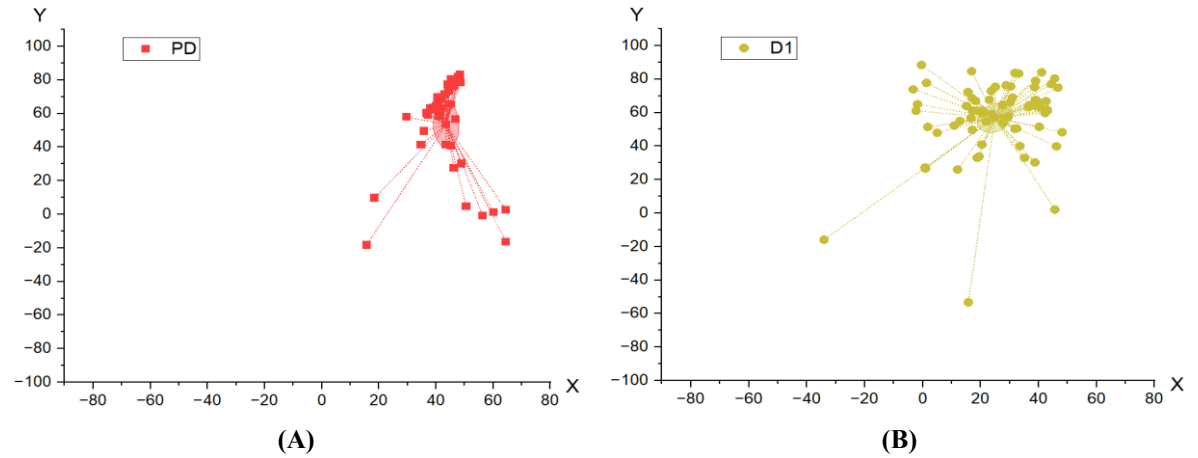
The equation for calculating the coefficient of convergence from the centroid coordinates for fault j is:

$$F_{JConvergence} = \frac{\sum d_j}{N_j} \qquad \text{(V.23)}$$

Where N_j is the total number of points for fault j, and d_j is the distance between each sample i and the centroid coordinates (xc_j, yc_j) for fault j, calculated using the equation:

$$d_{ij} = \sqrt{(x_{ij} - xc_j)^2 + (y_{ij} - yc_j)^2} \qquad \text{(V.24)}$$

The summation $\sum$ is taken over all the points associated with fault j. Figure V.6 shows the steps for calculating the convergence coefficient for each fault type sample from its center of coordinates, where the intersecting lines represent the distance of each type of fault coordinates from its center.



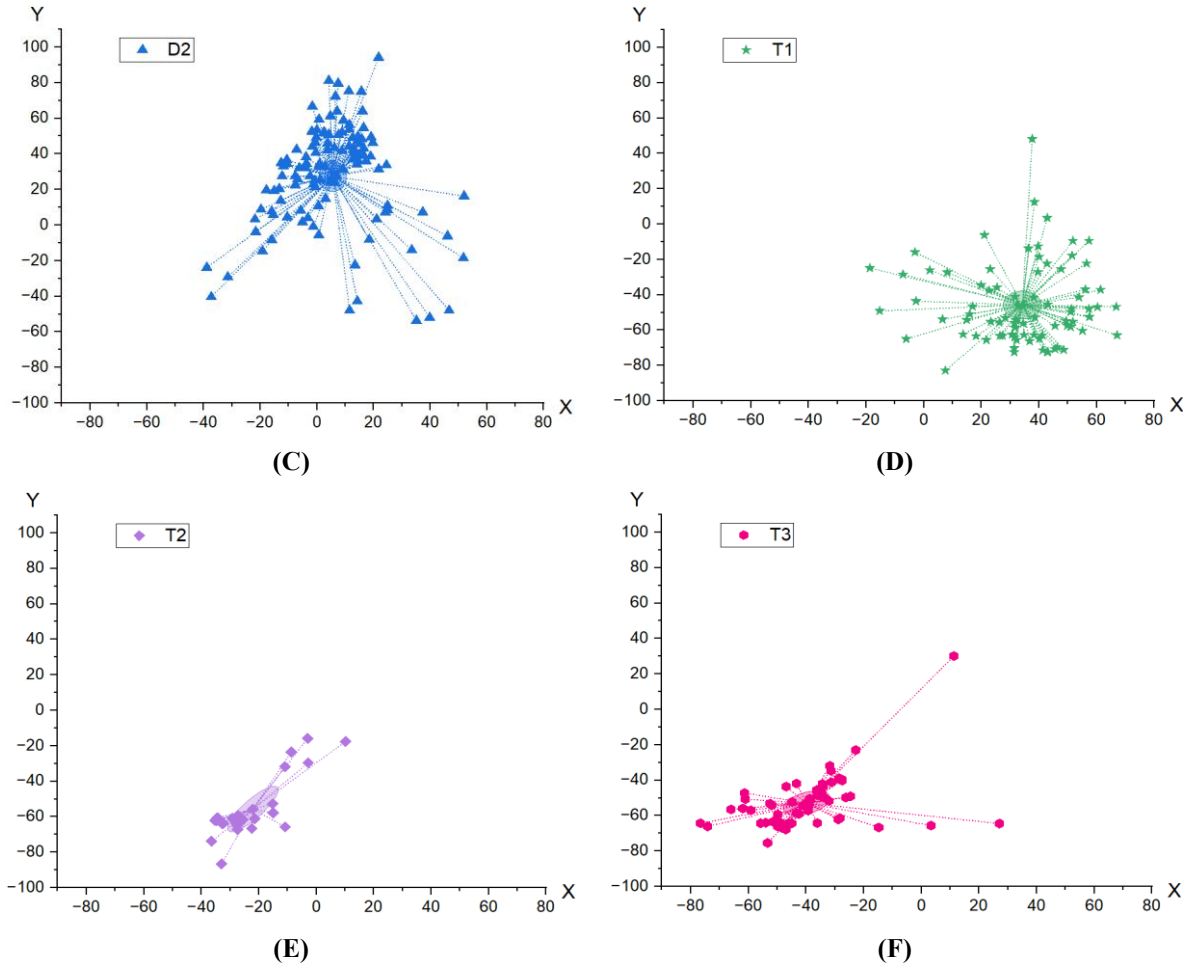


Figure V.6 Coefficient of convergence of each fault type point from its center coordinates

The global convergence coefficient is determined by aggregating the convergence coefficients of all faults and then dividing the cumulative value by the total number of faults, which is six in this context.

$$F_{GConvergence} = \frac{\sum F_{JConvergence}}{N_n} \qquad (V.25)$$

Where:

$\sum F_{JConvergence}$ represents the sum of the convergence coefficients for all six faults.

N_n is the total number of faultsand equals 6.

Figure V.7 shows the collection of all the convergence coefficients for the faults. The figure shows the convergence of all the samples towards their centers in order to achieve contrast and separation between the faults. However, convergence was achieved between the samples and their centers, but with the presence of overlap between the fault samples. To solve this problem, we must address the study of the spacing between the fault centers to achieve the desired separation of fault samples.

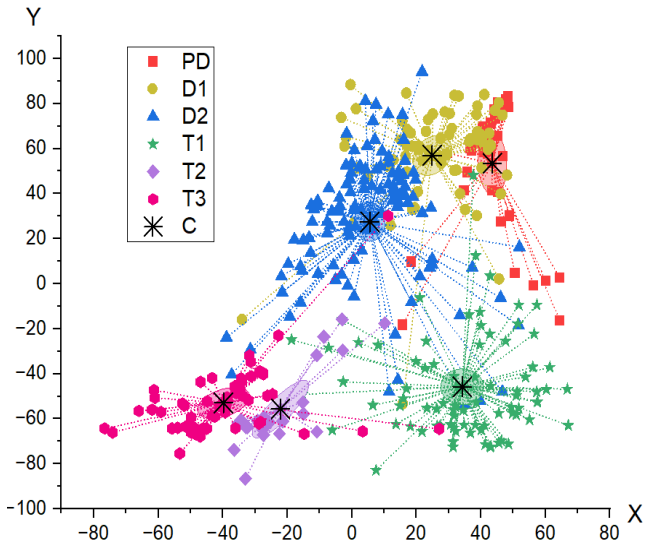


Figure V.7 Aggregation of all convergence coefficients for each fault type sample

V.4.3 Coefficient of Divergence Between the Coordinate Center

The coefficient of divergence between the coordinate center is a measure of how much the center points of a set of cases are dispersed or spread out from each other. The global divergence coefficient between the coordinate centers can be calculated using the following equation:

$$F_{G\ divergence} = \frac{\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} d_{jk}}{M} \qquad \text{(V.26)}$$

Where M is the total number of cases, which is given by M = (n * (n-1)) / 2, and n is the total number of center points.

The distance between the jth and kth center points, represented by d_{jk}, is given by:

$$d_{jk} = \sqrt{\left(xc_j - xc_k\right)^2 + \left(yc_j - yc_k\right)^2} \qquad \text{(V.27)}$$

Here, xc_j and yc_jrepresent the centroid coordinates for the jth fault, while xc_k and yc_k represent the centroid coordinates for the kth fault.

The coefficient of divergence is a measure of the average distance between all pairs of center points, normalized by the total number of cases. Figure V.8 shows the coefficient of divergence between the coordinate center. The intersecting points represent the possible cases for calculating each distance.

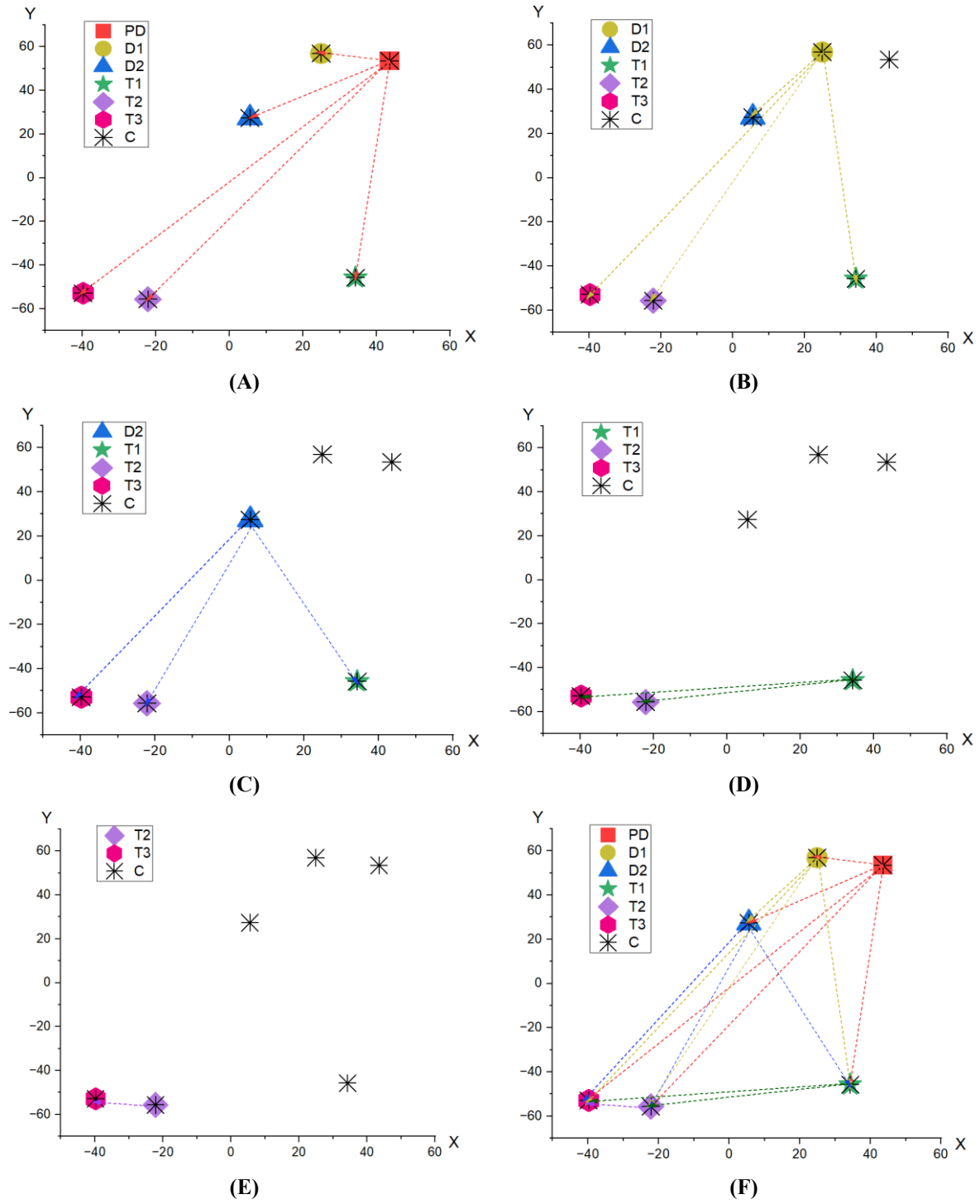


Figure V.8 The coefficient of divergence between center coordinates for each fault type

V.4.4 Objective Function

The objective function aims to identify and separate overlapping samples between faults by achieving convergence between the coordinates of the fault points, their centers, and the divergence between the fault centers. This process enables the determination of fault areas and angles for each region within the proposed circle shape. The equation is as follows:

$$Objective\ Function\ = (c1F_{G\ Convergence} + c2(100 - F_{G\ divergence}))/(c1\ +\ c2) \quad \text{(V.28)}$$

This equation calculates an objective function that takes into account both convergence and divergence measures between center points. The weights $c1$ and $c2$ represent the preference coefficients for both measures, where a larger value of $c1$ indicates a higher preference for convergence, and a larger value of $c2$ indicates a higher preference for divergence.

This objective function value will be a measure of the trade-off between the preference for convergence and the preference for divergence. A larger value of the objective function indicates a better overall performance in terms of both convergence and divergence.

Since there are likely to be multiple solutions that achieve the same objective function value, an optimization algorithm must be employed to identify the optimal solution that ensures the best angles for the Circle while also achieving convergence towards its center and divergence between the centers of each fault's coordinates. The optimization algorithm searches the space of all possible center point coordinates, subject to any constraints, evaluates the objective function for each candidate solution, and selects the solution that maximizes the objective function while satisfying the constraints. The choice of optimization algorithm will depend on the specific problem constraints and the complexity of the objective function.

V.5 Experimental Implementation and Discussion

V.5.1 Comparison of The Best Result Between Optimization Algorithms

GWO is renowned for its speed and effectiveness in finding optimal solutions. To substantiate this claim, a comparison was proposed with a modern advanced algorithm, the Whale Optimization Algorithm (WOA). Table V.1 shows the control parameters of GWO and WOA. Through the comparative analysis of Figure V.9, it is evident that the GWO exhibits slightly faster and more efficient convergence in obtaining the best result during the initial iterations compared to the WOA. Despite this, both algorithms achieve similar best final results. Therefore, the GWO demonstrates superior performance in terms of convergence speed and early stability. In the long term, while both algorithms show a high degree of convergence, the GWO maintains a slight edge. To further underscore the superiority of the GWO, a comparative analysis between Circle 1 and Circle 2 using both algorithms will be presented in the subsequent sections of this paper.

Table V.1 Control Parameter of GWO and WOA

Parameter	GWO	WOA
Search Agent No.	40	30
Maximum iterations	1000	1000
Dimension	4	4
Random number (r)	[0, 1]	[0, 1]
Best_score	Alpha_score	Leader_score

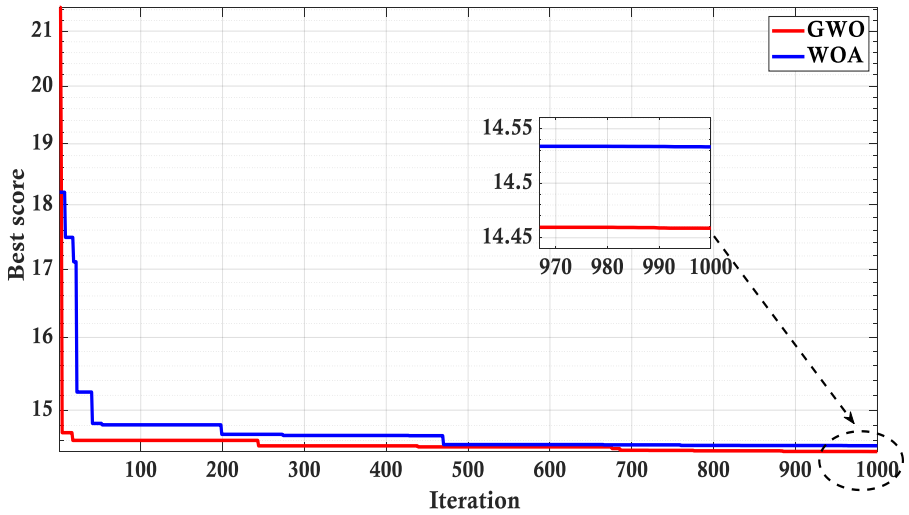


Figure V.9 Best score of GWO and WOA comparative analysis

By analyzing the comparison results, the initial focus will be on delineating the process for determining the circular shape using the proposed method. This involves explicating the steps required to ascertain the final form of Circle 1, derived from the GWO algorithm. Based on the practical principles of the proposed method, the final configuration of Circle 2, obtained through the WOA algorithm, will subsequently be deduced.

V.5.2 Proposed Circle 1 Configurations:

In order to form the proposed circular shape through the objective function that achieves the convergence of the fault samples around its center and the divergence of the fault centers between them, infinite solutions are reached that fulfill this condition, but based on the GWO algorithm, the optimal solution was reached. The main objective is that this solution achieves the convergence and divergence of the samples, which resulted in choosing the optimal angles that determine the optimal distribution that achieves the separation of overlapping gases. Table V.2 shows the optimum angles with their corresponding gases.

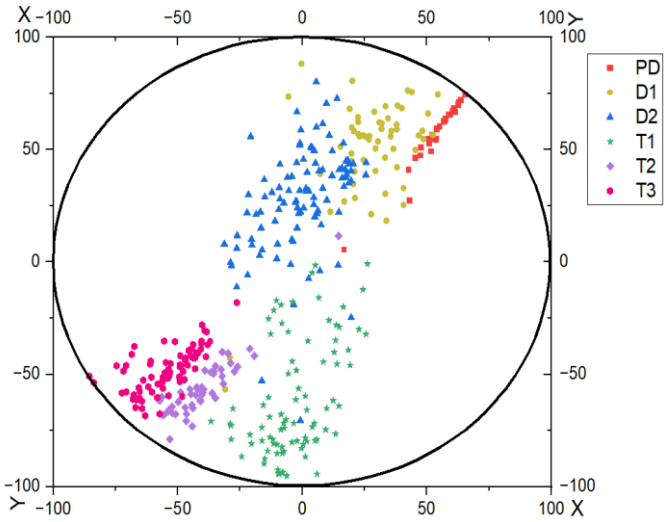


Figure V.10 The best results of the GWO algorithm for choosing the optimal angles and distribution

Figure V.10 shows the best results of the GWO algorithm for the desired distribution. This figure represents the sampling distribution achieved by the optimal angles through the sum of the key ratios of the five combustible gases (H_2, CH_4, C_2H_2, C_2H_4, C_2H_6), which equals 100%. This distribution represents the zone that the circle occupies, ranging from 0% to 100%, and indicates its center and radius, respectively, within the proposed circle. The circular shape was chosen because it provides a comprehensive coverage of all possible cases in the moving axes, ensuring that the optimal angles can effectively distribute and separate the types of overlapping faults. This geometric flexibility and symmetry make the circular shape the most suitable for this purpose.

Table V.2 Optimal Angles Obtained from The GWO Algorithm

Gases	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Angles	319.0371°	147.7378°	186.1282°	121.5745°	0°

V.5.2.1 Distinguish between electrical and thermal faults:

After the attainment of the optimal distribution facilitating the segregation of all power transformer faults, the initial phase in the development of the proposed circle methodology was initiated. This initial phase entails the delineation of boundaries distinguishing between electrical and thermal faults. With reference to Figure V.11, the proposed circle is partitioned into two regions, demarcated by lines, AB and BC, intersecting at the circle's central point, denoted as B. The upper section of the circle is indicative of electrical faults, visually represented in blue, while the lower portion denotes thermal faults, depicted in red.

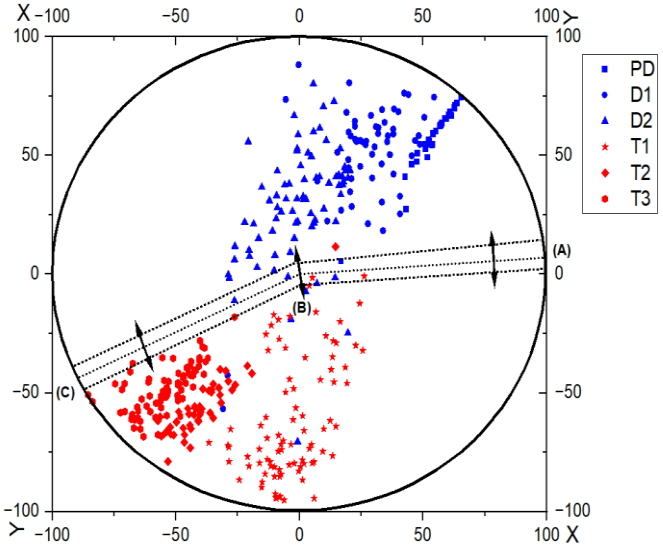


Figure V.11 Distinction between thermal and electrical faults of circle 1

In order to precisely demarcate the boundary of the electrical fault, the primary objective is to minimize any potential overlap between electrical and thermal samples, or conversely. This boundary is defined by an intersecting line that can be vertically adjusted, enabling the attainment of maximal precision in the minimization of overlap between electrical and thermal fault manifestations.

V.5.2.2 Determining zones of electrical faults:

The next step in the process involves the delineation of regions associated with electrical faults. These regions are categorized into three distinct zones: low-energy discharge (D1), high-energy discharge (D2), and partial discharge (PD). The classification begins at the centroid of each electrical fault, with fault zones confined within potential boundaries represented by dashed lines. These lines are oriented horizontally (right to left) and vertically (top to bottom) to define the extent of each fault region. The aim is to minimize interference between the different electrical fault zones by carefully considering the spatial distribution of fault samples.

As illustrated in Figure V.12(A), the distribution of electrical fault samples is optimized to reduce overlap and interference. Meanwhile, Figure V.12(B) highlights the precise boundaries necessary to achieve effective separation between the electrical faults. Using a counterclockwise orientation of angle θ, the PD fault is represented in red and confined to the region defined by points (I, H, G, F). The D1 fault is illustrated in yellow, covering the zone enclosed by points (A, I, F, G, H, D, E, B). Lastly, the D2 fault is depicted in blue and limited to the dotted zone bounded by points (D, C, B, E). This systematic delineation ensures clear differentiation and facilitates accurate identification of electrical faults.

V.5.2.3 Determining zones of thermal faults:

In the final step, thermal fault zones are determined based on temperature variations, enabling their classification into three distinct temperature ranges: low temperature (T1), medium temperature (T2), and high temperature (T3). Each range is centered around its respective midpoint to provide a clear reference for fault identification. The boundaries of these zones are delineated using dashed lines, which are adjusted horizontally to minimize potential overlaps between thermal fault regions. This adjustment ensures a more accurate differentiation, as demonstrated in Figure V.13(A).

Figure V.13(B) illustrates the final boundaries that effectively separate the thermal fault zones. The low-temperature fault (T1) is represented in green and confined to the area enclosed by points (K, A, B). The medium-temperature fault (T2) is shown in violet, occupying the zone defined by points (J, K, B). Lastly, the high-temperature fault (T3) is depicted in pink and limited to the region bounded by points (C, J, B). This structured approach ensures precise categorization and enhances the reliability of thermal fault identification within the system.

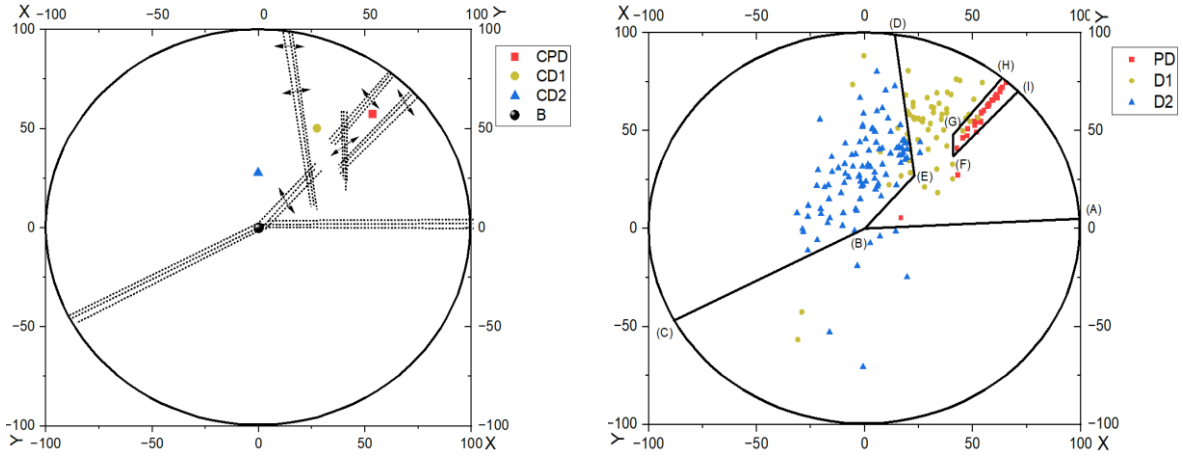
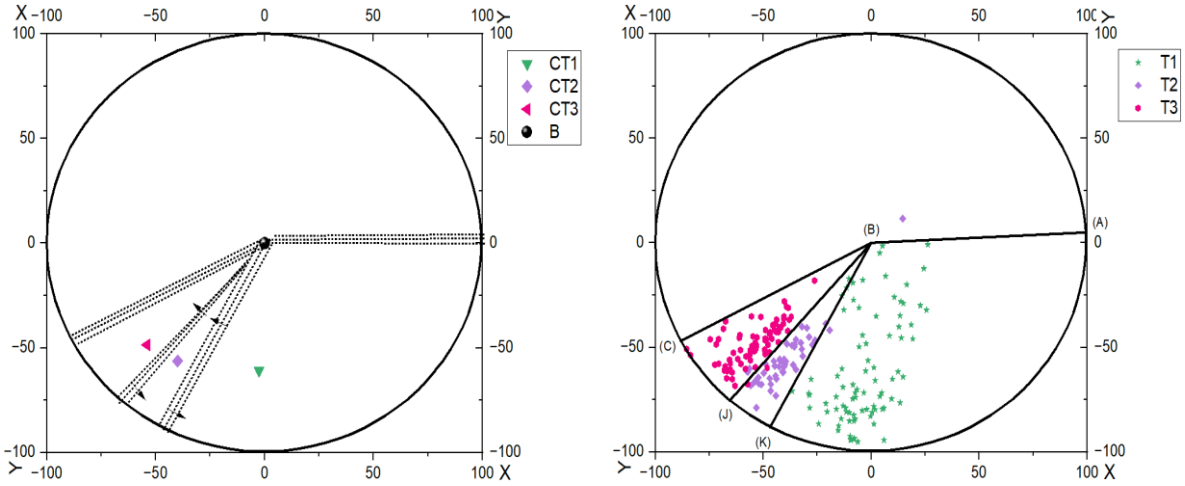


Figure V.12 Determining zones of electrical faults of circle 1



(A) **(B)**

Figure V.13 Determining zones of thermal faults of circle 1

V.5.3 Proposed Circle 2 Configurations:

Building on the framework established in Circle Method 1, the proposed Circle Shape 2 leverages the same objective function to ensure fault sample convergence near their centers and maximize the divergence among different fault centers. This process inherently allows for an infinite set of potential solutions. The Whale Optimization Algorithm (WOA) was employed to derive the optimal solution by determining the most effective angles for sample distribution, thereby facilitating the separation of overlapping gases. Table V.3 outlines these optimal angles and their corresponding gas types.

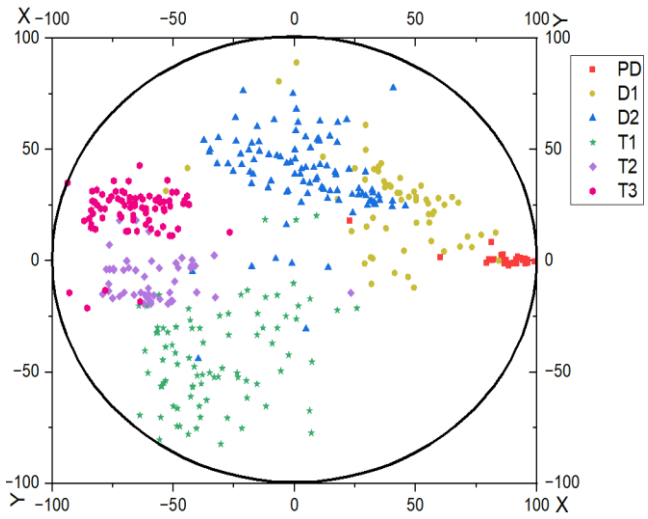


Figure V.14 The best results of the WOA algorithm for choosing the optimal angles and distribution

Figure V.14 displays the optimal sample distribution achieved using the Whale Optimization Algorithm (WOA) algorithm. The figure illustrates how the chosen angles enable the cumulative distribution of the five major combustible gases H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6 across the area defined by Circle Shape 2. This circle, with a center located at 0% and a radius extending to 100%, provides comprehensive coverage of all potential cases along the moving

axes. The results demonstrate the capability of Circle Shape 2 to effectively distribute samples and separate overlapping fault types by leveraging the optimal angles.

Table V.3 Optimal Angles Obtained from The WOA Algorithm

Gases	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2
Angles	269.7509°	55.0585°	154.1765°	69.4878°	0°

V.5.3.1 Distinguish between electrical and thermal faults:

The first step in identifying fault areas using the Circle 2 method involves distinguishing between electrical and thermal faults. According to the Pentagon Mansour method, electrical faults are characterized by the presence of gases such as H_2 and C_2H_2, with some traces of C_2H_6, which are shared between electrical and thermal fault conditions. Consequently, electrical fault cases are shifted toward the upper region of the circuit. In contrast, thermal faults are primarily associated with the presence of CH_4, C_2H_4, and certain traces of C_2H_6, leading to the placement of thermal fault cases in the lower region of the circuit. This classification divides the circuit into two distinct regions, separated by the line AB, as depicted in Figure V15.

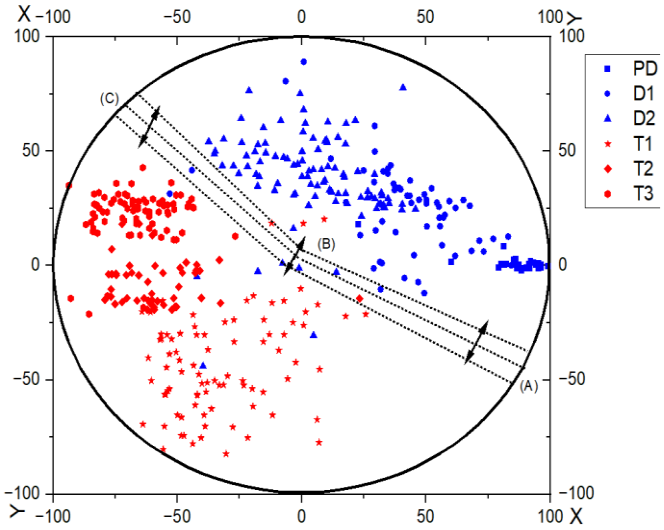


Figure V.15 Distinction between thermal and electrical faults of circle 2

Once the expected boundary between electrical and thermal faults is established, Dissolved Gas Analysis (DGA) data is integrated into the circuit to refine the precise boundary using the Circle 2 method. Electrical faults are represented in blue, while thermal faults are depicted in red. The exact boundary, denoted as AB, is determined by minimizing the interference between electrical and thermal fault samples. As shown in Figure V.15, this boundary is represented by a dashed line, which is iteratively adjusted vertically until the overlap between electrical and thermal fault samples is minimized.

V.5.3.2 Determining zones of electrical faults:

The process of identifying electrical fault zones involves determining regions characterized by varying discharge levels, including low, high, and partial discharge. The analysis begins by establishing the centroid of each electrical fault and subsequently delineating the fault zones. These zones are represented by dashed lines that extend horizontally (to the right and left) and

vertically (up and down) to encapsulate the discharge patterns. The boundaries are defined using Dissolved Gas Analysis (DGA) data specific to electrical faults, as shown in Figure V.16(A). This systematic approach ensures a well-structured distribution of electrical fault samples and their midpoints, thereby minimizing interference and overlap among the zones.

In Figure V.16(B), the specific boundaries for separating fault zones are depicted. The Partial Discharge (PD) fault is shown in red and confined within the region defined by points (I, H, G, F). The D1 fault, represented in yellow, occupies the zone enclosed by points (A, I, H, D, E, B). Meanwhile, the D2 fault, shown in blue, is delimited by a dotted boundary formed by points (B, E, D, C). Each fault type is distinctly categorized, ensuring minimal interference between the zones. This precise delineation facilitates enhanced fault diagnosis and classification, contributing to more reliable electrical fault detection.

V.5.3.3 Determining zones of thermal faults:

The accurate identification of thermal fault types poses a critical challenge in power transformer systems. According to the IEC standard, thermal fault zones are categorized into three distinct levels based on temperature: low-temperature (T1), medium-temperature (T2), and high-temperature (T3). These faults are primarily associated with the presence of methane and ethane gases, with each fault zone defined by its respective midpoint. To ensure minimal interference, the boundaries for each fault are delineated by dashed lines that are adjusted horizontally. This adjustment is performed using Dissolved Gas Analysis (DGA) data specific to thermal faults, as demonstrated in Figure V.17(A) This step enables for precise identification of the fault zone to reduce interference and improve diagnostic accuracy.

In Figure V.17(B), the finalized boundaries for the separation of thermal fault zones are presented. The low-temperature fault (T1) is represented in green and enclosed within the region defined by points (A, B, K). The medium-temperature fault (T2) is illustrated in violet and confined to the area bounded by points (K, B, J). Finally, the high-temperature fault (T3) is depicted in pink and limited to the region marked by points (J, B, C). This systematic delineation of fault zones ensures a clear distinction between different thermal fault types, facilitating improved fault identification and effective system management.

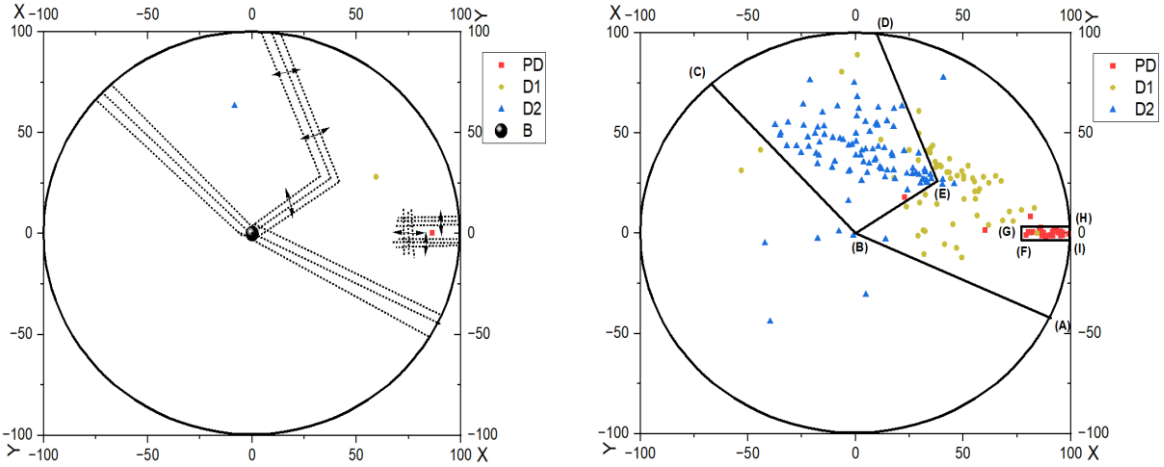
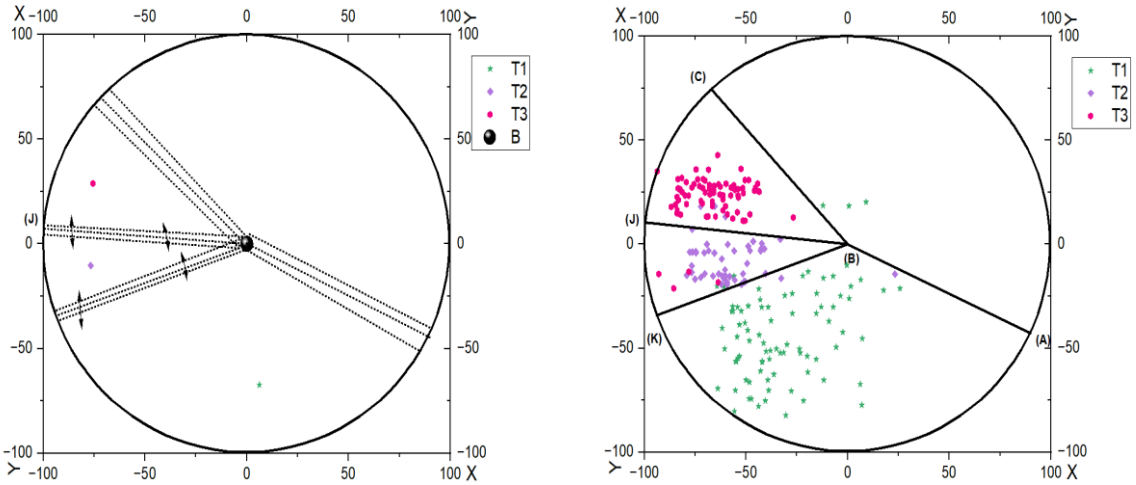


Figure V.16 Determining zones of electrical faults of circle 2



(A) (B)

Figure V.17 Determining zones of thermal faults of circle 2

V.5.4 Final proposed Circle 1 shape

Based on the information presented earlier, Figure V.18(A) depicts the final circle method, incorporating a dataset of 407 DGA samples sourced from the literature [119], [120],[121]. This dataset encompasses six distinct electrical and thermal faults, forming the basis for the implement of the new circle method.

Utilizing the DGA data, boundaries were established for each of the six fault zones, with a focus on identifying the minimum boundaries that yield the highest accuracy and enable effective differentiation between the various fault types. The new circle method, along with its fault zone delineations, is succinctly summarized in Figure V.18(B).

Furthermore, it is essential to underscore the importance of elucidating the specific (x, y) coordinates of all points that played a pivotal role in defining the boundaries of the fault regions within the proposed circle 1. These coordinates are detailed as follows:

PD: I (70, 71), H (66, 75), G (41, 48), F (41, 37).

D1: A (99.75, 6.97), I (70, 71), F (41, 37), G (41, 48), H (66, 75), D (13.9, 99), E (23, 27), B (0, 0).

D2: D (13.9, 99), C (-88.29, -46.94), B (0, 0), E (23, 27).

T1 : K (-46.94, -88.29), A (99.75, 6.97), B (0, 0).

T2 : J (−65.6, -75.46), K (-46.94, -88.29), B (0, 0).

T3 : C (-88.29,-46.94), J (−65.6,−75.46), B (0, 0).

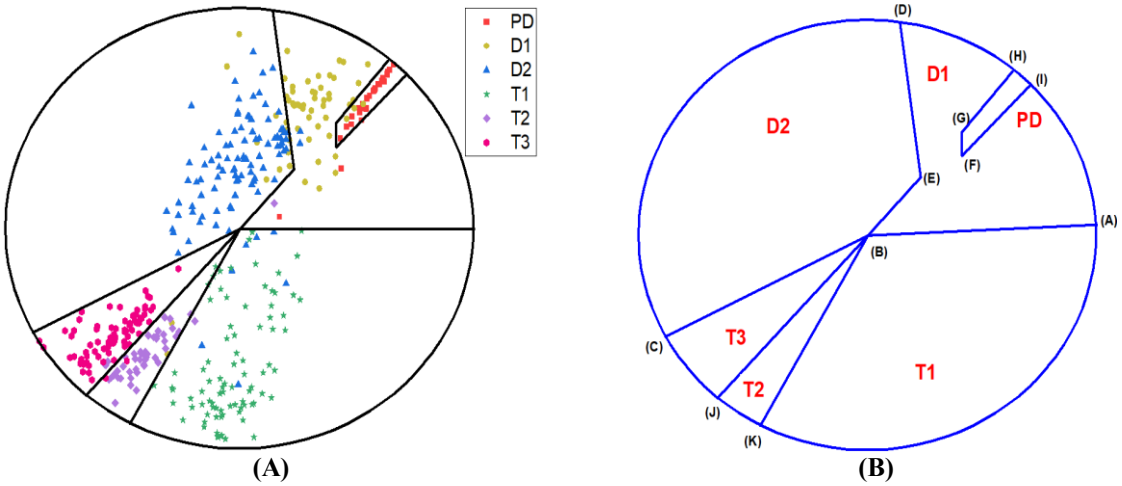


Figure V.18 Proposed final Circle 1 shape

V.5.5 Final proposed Circle 2 shape

To obtain the final configuration of the proposed Circle 2 method, an objective function is employed, following the procedures akin to the Gray Wolf algorithm. This process identifies the solution that achieves optimal distribution and angles, replicating the methodology for determining the zone boundaries for each fault in Circle 1. Figure V.19(A) illustrates the optimal distribution achieved using the Whale Optimization Algorithm, while Figure V.19 (B) depicts the final form of the proposed Circle 2 method.

In addition, the boundaries of the fault zones within the proposed Circle 2 have been identified through the following coordinates:

PD: I (99.94, -3.49), H (99.94, 3.49), G (77, 3.49), F (77, -3.49).

D1: A (90.63, -42.26), I (99.94, -3.49), F (77, -3.49), G (77, 3.49), H (99.94, 3.49), D (10, 99), E (38, 26), B (0, 0).

D2: D (10, 99), C (-66.91, 74.31), B (0, 0), E (38, 26).

T1 : K (-93.97, -34.2), A (90.63, -42.26), B (0, 0).

T2 : J (−99.45, 10.45), K (-93.97, -34.2), B (0, 0).

T3 : C (-66.91, 74.31), J (−99.45, 10.45), B (0, 0).

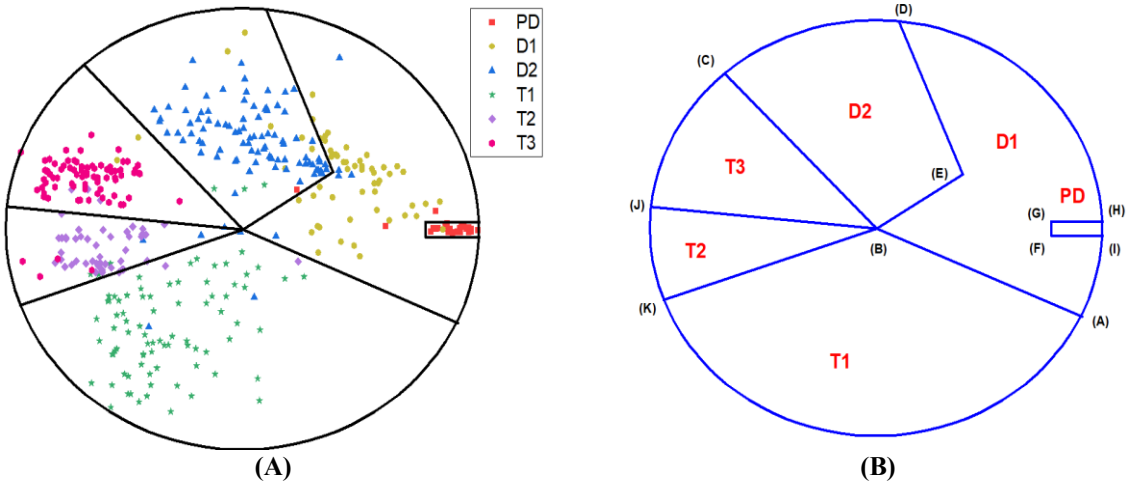


Figure V.19 Proposed final Circle 2 shape

V.5.6 Implementation of Test Examples on The Proposed Circle Methods 1 And 2

Figure V.20 and Table V.4 presents a comparative analysis of the new DGA techniques, Circle 1 and Circle 2, based on the provided dataset. The results indicate that the Circle 1 method, derived using the Gray Wolf algorithm, demonstrated superior performance compared to the Circle 2 method, which was obtained using the Whale Optimization algorithm. Specifically, the Circle 1 method achieved a diagnostic accuracy of 91.15%, surpassing the 90.66% accuracy of the Circle 2 method.

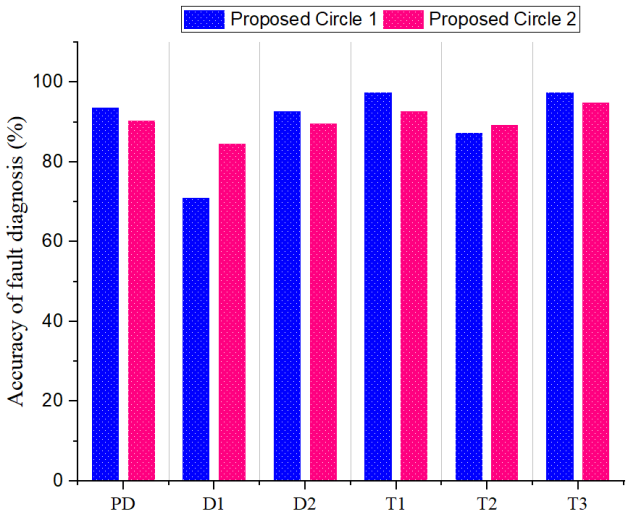


Figure V.20 Comparative analysis between the accuracy of the proposed Circle 1 and Circle 2 methods

Table V.4 Comparison between the proposed Circle 1 and 2 methods

Accuracy (%)	Proposed Circle 1	Proposed Circle 2
PD	93.75	90.62
D1	71.17	84.75
D2	92.93	89.90
T1	97.67	93.02
T2	87.5	89.58
T3	97.59	95.18
Total	91.15	90.66

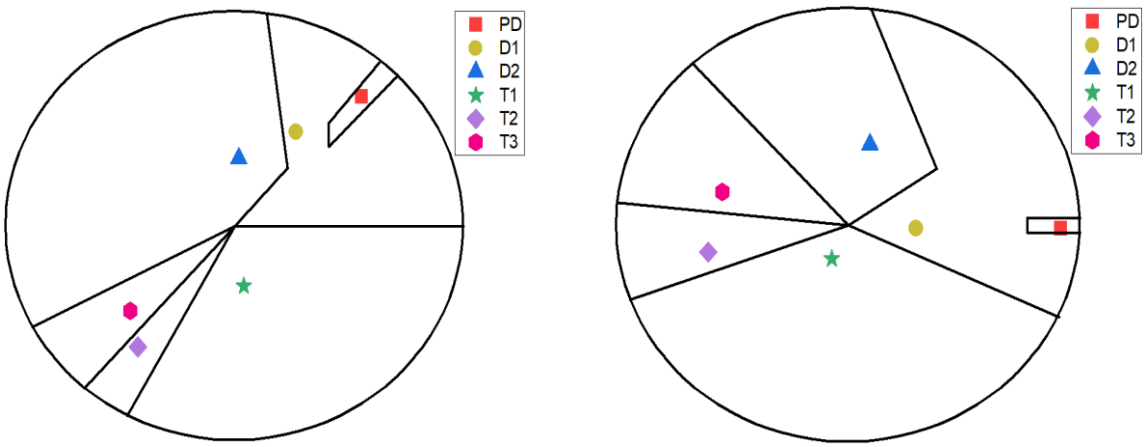
Table V.5 presents six sample test examples, each with their respective actual conditions derived from the collected database [121]. These samples were meticulously chosen to evaluate the performance limits of the proposed circles. By utilizing the previously established equations and applying them to these six samples, the coordinates listed in Table V.6 were determined. Consequently, the fault zone boundaries illustrated in Figure V.21(A) Proposed circle 1, (B) Proposed circle 2 were validated.

Table V.5 Different samples of test examples with their actual faults [121]

	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	Faults
S1	625	49	9	7	0.6	PD
S2	101.72	27.65	7.13	16.92	53.87	D1
S3	755	229	32	404	460	D2
S4	33	29	9	12	0.001	T1
S5	23.51	61.33	45.21	98.03	1.01	T2
S6	165.62	240.95	61.32	514.53	13.53	T3

Table V.6 Proposed methods coordinates for testing examples

Samples	Xcircle 1	Ycircle 1	Xcircle 2	Ycircle 2
S1	55.1932	60.2574	91.3264	-1.3267
S2	26.5606	43.9657	28.9933	-1.2354
S3	1.7271	31.5503	9.3058	37.2679
S4	3.90198	-27.8105	-7.14723	-15.3265
S5	-42.2906	-56.1778	-60.3574	-12.3169
S6	-45.6603	-39.4003	-54.2714	15.3265



(A) **(B)**

Figure V.21 Different samples for test examples

V.5.7 Validation and comparison with current techniques

To validate the results obtained from the proposed Circle 1 and 2 methods, 70 different samples were selected from the IEC TC 10 database [120] for comparison with the best newly-established conventional methods with the aim of verifying the reliability of the proposed approach and carefully analyzing its strengths and weaknesses.

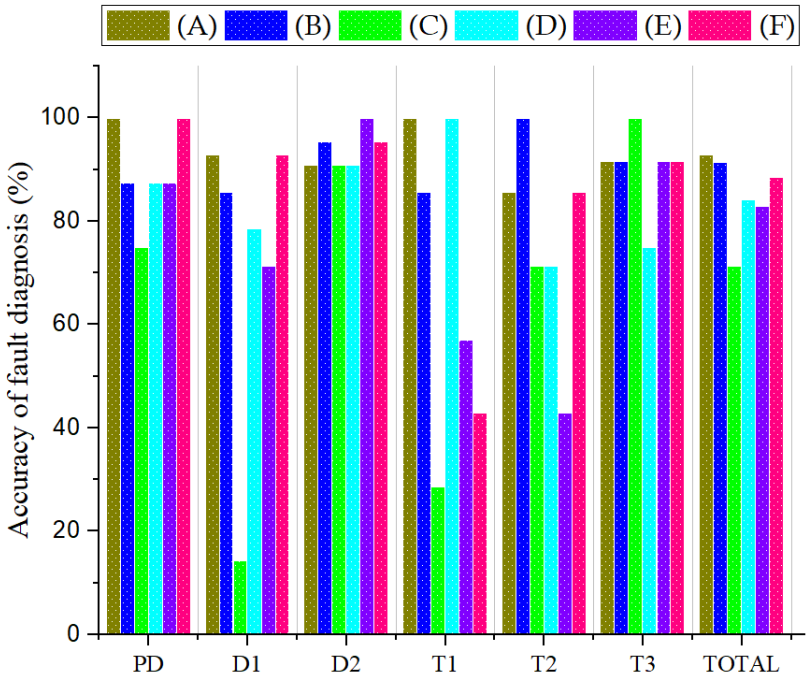


Figure V.22 Various samples for testing examples

According to Table V.7, the conventional methods used for comparison and validation include the Duval Pentagon, the Mansour Pentagon, the Cartesian Graphical, and the Gouda Triangle. Table V.8 and Figure V.22 reveal that the overall accuracy of the new circles methods 1 and 2 reached 92.86% and 91.43%, respectively. In contrast, the accuracy percentages of the conventional techniques were 71.43% for the Duval Pentagon, 84.29% for the Mansour Pentagon, 82.86% for the Cartesian Graphical, and 88.57% for the Gouda Triangle. This marked superiority of the new circles methods underscores the high reliability of the novel DGA approach in diagnosing power transformer faults.

Figure V.23 shows a visual representation of the distribution of transformer faults in order to visually illustrate the superiority of the proposed circles methods using DGA of the IEC TC 10 database for: (A) Proposed circle1, (B) Proposed circle2, (C) Duval pentagon, (D) Mansour Pentagon (E) Cartesian Graphical, and (F) Gouda triangle and the fault zones for each method.

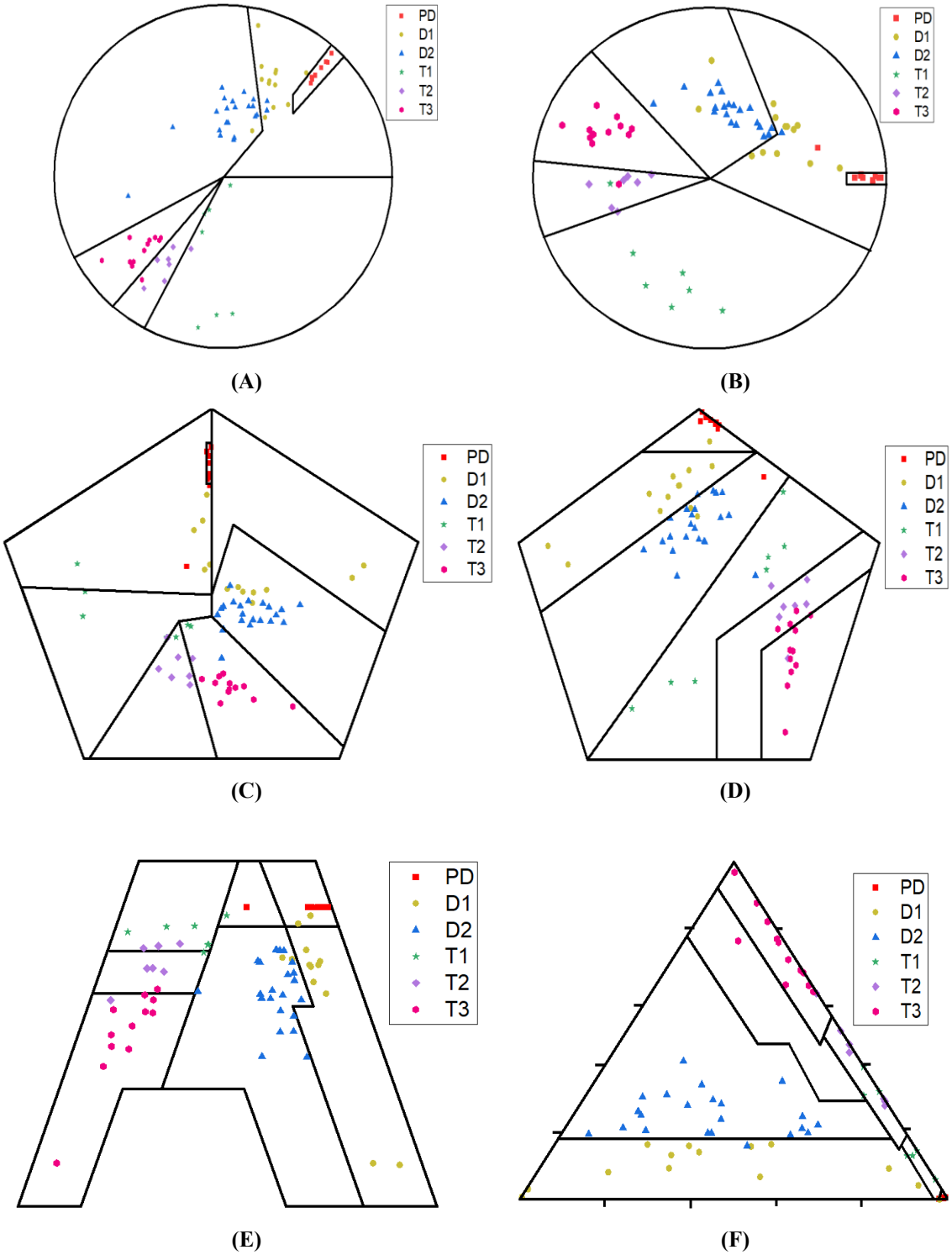


Figure V.23 IEC TC 10 database DGA cases for: (A) Proposed circle1, (B) Proposed circle2, (C) Duval pentagon, (D) Mansour Pentagon (E) Cartesian Graphical, and (F) Gouda triangle and the fault zones for each method

Table V.7 Description of the IEC tc10 Database set

	PD	D1	D2	T1	T2	T3	Total
No of Samples	9	13	22	7	7	12	70

Table V.8 Comparing the proposed circles methods with conventional methods

Diagnostic techniques	Accuracy of types of faults (%)						Accuracy overall (%)
DGA	PD	D1	D2	T1	T2	T3	
Proposed Circle 1 [165]	100	92.86	90.91	100	85.71	91.67	92.86
Proposed Circle 2 [165]	87.50	85.71	95.45	85.71	100	91.67	91.43
Duval Pentagon [77]	75.00	14.29	90.91	28.57	71.43	100	71.43
Mansour Pentagon [79]	87.50	78.57	90.91	100	71.43	75.00	84.29
Cartesian Graphical [82]	87.50	71.29	100	57.14	42.86	91.67	82.86
Gouda Triangle [80]	100	92.86	95.45	42.86	85.71	91.67	88.57

V.6 Conclusion

In this chapter, two new graphical DGA analysis methods, Circle Method 1 and Circle Method 2, are presented to interpret DGA in power transformers. These new circuit methods rely on the convergence and divergence of samples to distinguish between different faults by finding the optimal angles that achieve the best distribution. This is accomplished using advanced metaheuristic optimization algorithms. The proposed methods are validated and their performance is compared with those in the literature using the IEC TC10 database. From the details obtained, the proposed Circle Methods 1 and 2 have higher overall accuracy in diagnosing fault conditions compared to other conventional DGA techniques.

Chapter VI. Conclusions and Future Work

VI.1 Conclusions

The power transformer is a critical component of the electrical energy transmission and distribution network. To ensure its reliable and safe operation, an effective and timely assessment of its condition is essential. In practice, various techniques are available for evaluating the transformer's condition, among which oil analysis stands out as the simplest and most effective method. This technique is considered simple because it only requires a straightforward sampling and analysis of a small quantity of oil without necessitating a service interruption. It is also effective because the oil serves as a vital information medium, retaining records of internal processes within the transformer and providing insight into any thermal, electrical, or combined malfunctions.

The analysis of dissolved gases in oil is a cutting-edge tool in this field. By precisely determining the types and quantities of gases dissolved in the oil, it enables the identification of their sources. This method is regarded as a diagnostic tool for detecting faults and ensuring continuous monitoring of transformers.

Accurate diagnosis through Dissolved Gas Analysis (DGA) is contingent upon the measured gases reliably reflecting the actual condition of the equipment, coupled with an accurate interpretation of the data. This thesis contributes to this field by introducing an intelligent information system aimed at enhancing the diagnostic and predictive capabilities for power transformer faults. To contextualize this work, **Chapter II** critically examines the limitations of traditional DGA methods, laying the groundwork for the proposed innovations.

Addressing the challenges posed by the limitations and insufficient documentation of traditional DGA methods, an intelligent model was developed in **Chapter II** through the application of tree-based ensemble learning algorithms. These algorithms were integrated with input vectors derived from traditional DGA techniques, resulting in a model with superior diagnostic accuracy compared to conventional methods. **Chapter IV** expanded on this work by constructing a refined model that utilized only the core components of traditional DGA input vectors. The application of ensemble learning strategies such as Bagging, Decorate, and Boosting demonstrated enhanced predictive performance by outperforming single-decision models. Furthermore, balancing the dataset through the generation of synthetic data and its combination with real data significantly contributed to improving model accuracy, as evidenced by the comparative analysis of real versus hybrid datasets.

Expanding on the intelligent models for fault classification, **Chapter V** introduced a groundbreaking graphical DGA technique that overcame the persistent limitations of traditional methods. This approach integrated mathematical equations, physical principles, and optimization algorithms to develop a robust diagnostic framework. Central to this innovation were the Circle 1 and Circle 2 methodologies, which employed five dynamic axes to analyze the behavior of sample data. The method systematically investigated the convergence of samples around their respective centers and the divergence between fault centers, enabling clear differentiation between fault types. While this method initially offered infinite solutions, the use of metaheuristic optimization algorithms, specifically GWO and WOA, refined the process by identifying optimal angles with a distinct sample distribution. This refinement allowed for

an accurate and effective representation of fault zones for Circle 1 and Circle 2, achieving superior diagnostic performance compared to all widely used traditional DGA techniques.

VI.2 Future Work

The research presented in this thesis offers several avenues for further investigation:

VI.2.1 Development of a hybrid intelligent information system:

Future work could explore the integration of DGA methods with other diagnostic techniques, such as partial discharge analysis, breakdown voltage testing, and thermography. This would enable the creation of a comprehensive, multi-source intelligent system for accurately detecting the condition of power transformers.

VI.2.2 Incorporation of optimization algorithms:

When using intelligent systems that combine DGA input vectors and classification algorithms, optimization algorithms should also be employed to enhance model accuracy. These optimization methods can be further integrated with dimensionality reduction techniques to improve computational efficiency and predictive performance.

VI.2.3 Discovery of novel DGA input vectors:

Investigating and identifying new DGA input vectors not previously utilized could pave the way for advancements in DGA methods, ensuring that they remain aligned with evolving diagnostic techniques and transformer technologies.

VI.2.4 Advancement of DGA interpretation algorithms through deep learning:

To enhance the accuracy and reliability of DGA-based diagnostics, more advanced deep learning approaches could be employed. This would allow for improved fault detection and prediction capabilities.

VI.2.5 Enhancement of the graphical DGA method:

Building on the dynamic-axes graphical DGA method introduced in this thesis, future research could focus on incorporating additional parameters, such as CO and CO_2 gases, to develop a new graphical technique based on seven dynamic axes. This would provide more robust coverage of various fault scenarios and offer greater efficiency in analyzing the impact of insulation performance.

VI.2.6 Integration of graphical DGA methods with artificial intelligence:

Combining the proposed graphical DGA technique with advanced artificial intelligence algorithms and exploring the use of alternative optimization algorithms could significantly improve the reliability and precision of the diagnostic model.

References

[1] R. R. Singh, G. Bhatti, and D. Saravanan, "New-age condition monitoring of on-load tap changing transformers in distributed energy systems for Industry 4.0," *e-Prime-Advances in Electrical Engineering, Electronics and Energy,* vol. 2, p. 100087, 2022, doi: 10.1016/j.prime.2022.100087.

[2] P. Gill, *Electrical power equipment maintenance and testing*. CRC press, 2016.

[3] L. I. U. Raanaa, "Condition monitoring of power transformers in digital substations," NTNU, 2020.

[4] M. E. A. Senoussaoui, "Contribution des techniques intelligentes au diagnostic industriel des transformateurs de puissance," Université de Sidi Bel Abbès-Djillali Liabes, 2019.

[5] A. Peimankar, "Intelligent condition assessment of power transformers," 2017.

[6] M. A. Mahmoud, N. R. Md Nasir, M. Gurunathan, P. Raj, and S. A. Mostafa, "The current state of the art in research on predictive maintenance in smart grid distribution network: Fault's types, causes, and prediction methods—a systematic review," *Energies,* vol. 14, no. 16, p. 5078, 2021, doi: 10.3390/en14165078.

[7] A. Hechifa, A. Lakehal, L. Saidi, A. Nanfak, R. Kelaiaia, and M. E. A. Senoussaoui, "Improvement of transformer fault diagnosis using fuzzy rule and decision tree based on dissolved gas analysis," in *2023 1st International Conference on Renewable Solutions for Ecosystems: Towards a Sustainable Energy Transition (ICRSEtoSET)*, 2023: IEEE, pp. 1-6, doi: 10.1109/ICRSEtoSET56772.2023.10525429.

[8] S. Boudraa, "Maintenance des transformateurs de puissance par l'analyse de l'huile-Apport de l'Intelligence artificielle," Université de Batna 2, 2017.

[9] J. Sanchez, "Aide au diagnostic de défauts des transformateurs de puissance," Université de Grenoble, 2011.

[10] B. Ahmed and M. Hocine, "Application de méthodes intelligentes pour le diagnostic de l'huile de transformateur par analyse des gaz dissous."

[11] R. Minkner and J. Schmid, "The Technology of Instrument Transformers."

[12] S. A. Wani, "Transformer fault diagnosis Based on Dissolved Gas Analysis DGA."

[13] K. F. Thang, *An improved approach to data analysis & interpretation in transformer condition assessment based on unsupervised neutral network*. University of Bath (United Kingdom), 2002.

[14] S. Li, "Study of dissolved gas analysis under electrical and thermal stresses for natural esters used in power transformers," The University of Manchester (United Kingdom), 2012.

[15] Y. Li, "The State of the Art in transformer fault diagnosis with artificial intelligence and Dissolved Gas Analysis: A Review of the Literature," *arXiv preprint arXiv:2304.11880,* 2023.

[16] A. Ahmed, "Power Transformer Condition Monitoring and Diagnosis," *United Kingdom: IET, The Institution of Engineering and Technology,* 2018.

[17] J. Ni, Z. Zhao, S. Tan, Y. Chen, C. Yao, and C. Tang, "The actual measurement and analysis of transformer winding deformation fault degrees by FRA using mathematical indicators," *Electric Power Systems Research,* vol. 184, p. 106324, 2020, doi: 10.1016/j.epsr.2020.106324.

[18] S. X. Qian, Q. Du, X. J. Gu, and J. Y. Song, "Transformer Winding Deformation Fault Diagnose Based on FRA and Energy Feature Extraction by Wavelet Packet," *Advanced Materials Research,* vol. 490, pp. 1486-1490, 2012, doi: 10.4028/www.scientific.net/AMR.490-495.1486.

[19] T.-D. Do, V.-N. Tuyet-Doan, Y.-S. Cho, J.-H. Sun, and Y.-H. Kim, "Convolutional-neural-network-based partial discharge diagnosis for power transformer using UHF sensor," *IEEE access,* vol. 8, pp. 207377-207388, 2020, doi: 10.1109/ACCESS.2020.3038386.

[20] H. Karami, H. Tabarsa, G. B. Gharehpetian, Y. Norouzi, and M. A. Hejazi, "Feasibility study on simultaneous detection of partial discharge and axial displacement of HV transformer

winding using electromagnetic waves," *IEEE Transactions on Industrial Informatics,* vol. 16, no. 1, pp. 67-76, 2019, doi: 10.1109/TII.2019.2915685.

[21] A. Indarto, A. W. Murti, F. Husnayain, I. Garniwa, A. Rahardjo, and C. Hudaya, "Influence of different adhesives on partial discharge in power transformer winding cylinder insulation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 27, no. 3, pp. 964-970, 2020, doi: 10.1109/TDEI.2020.008533.

[22] Y. Shi, S. Ji, F. Zhang, Y. Dang, and L. Zhu, "Application of operating deflection shapes to the vibration-based mechanical condition monitoring of power transformer windings," *IEEE Transactions on Power Delivery,* vol. 36, no. 4, pp. 2164-2173, 2020, doi: 10.1109/TPWRD.2020.3021909.

[23] H. Zhou, K. Hong, H. Huang, and J. Zhou, "Transformer winding fault detection by vibration analysis methods," *Applied Acoustics,* vol. 114, pp. 136-146, 2016, doi: 10.1016/j.apacoust.2016.07.024.

[24] Y. Zhou *et al.*, "Vibration signal-based fusion residual attention model for power transformer fault diagnosis," *IEEE Sensors Journal,* vol. 24, no. 10, pp. 17231-17242, 2024, doi: 10.1109/JSEN.2024.3382811.

[25] L. Jin, D. Kim, A. Abu-Siada, and S. Kumar, "Oil-immersed power transformer condition monitoring methodologies: A review," *Energies,* vol. 15, no. 9, p. 3379, 2022, doi: 10.3390/en15093379.

[26] N. A. Baka, A. Abu-Siada, S. Islam, and M. El-Naggar, "A new technique to measure interfacial tension of transformer oil using UV-Vis spectroscopy," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 22, no. 2, pp. 1275-1282, 2015, doi: 10.1109/TDEI.2015.7076831.

[27] H. Torkaman and F. Karimi, "Measurement variations of insulation resistance/polarization index during utilizing time in HV electrical machines–A survey," *Measurement,* vol. 59, pp. 21-29, 2015, doi: 10.1016/j.measurement.2014.09.034.

[28] T. Mariprasath and V. Kirubakaran, "A real time study on condition monitoring of distribution transformer using thermal imager," *Infrared Physics & Technology,* vol. 90, pp. 78-86, 2018, doi: 10.1016/j.infrared.2018.02.009.

[29] F. Segovia *et al.*, "Connected system for monitoring electrical power transformers using thermal imaging," *Integrated Computer-Aided Engineering,* vol. 30, no. 4, pp. 353-368, 2023, doi: 10.3233/ICA-230712.

[30] A. Abdali, A. Abedi, H. Masoumkhani, K. Mazlumi, A. Rabiee, and J. M. Guerrero, "Magnetic-thermal analysis of distribution transformer: Validation via optical fiber sensors and thermography," *International Journal of Electrical Power & Energy Systems,* vol. 153, p. 109346, 2023, doi: 10.1016/j.ijepes.2023.109346.

[31] J. Lukic *et al.*, "Changes of new unused insulating kraft paper properties during drying-Impact on degree of polymerization," *Electra (Paris),* no. 312, pp. 76-81, 2020.

[32] S. Gao, L. Yang, and T. Ke, "Failure mechanism of transformer oil-immersed cellulosic insulation induced by sulfur corrosion," *Cellulose,* vol. 27, no. 12, pp. 7157-7174, 2020, doi: 10.1007/s10570-020-03271-x.

[33] A. Kumar, D. Mishra, and A. Baral, "Importance of depolarization current in the diagnosis of oil-paper insulation of power transformer," *IEEE Access,* vol. 11, pp. 56858-56864, 2023, doi: 10.1109/ACCESS.2023.3283914.

[34] D. Mishra, A. Baral, and S. Chakravorti, "Reliable assessment of oil-paper insulation used in power transformer using concise dielectric response measurement," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 30, no. 3, pp. 1255-1264, 2023, doi: 10.1109/TDEI.2023.3261824.

[35] B. Standard, "Insulating liquids-determination of breakdown voltage at power frequency-test method," *BSEN60156,* 1996.

[36] M. R. Mahmoudi, M. H. Heydari, and R. Roohi, "A new method to compare the spectral densities of two independent periodically correlated time series," *Mathematics and Computers in Simulation,* vol. 160, pp. 103-110, 2019, doi: 10.1016/j.matcom.2018.12.008.

[37] I. Fofana, A. Bouaicha, and M. Farzaneh, "Characterization of aging transformer oil–pressboard insulation using some modern diagnostic techniques," *European Transactions on Electrical Power,* vol. 21, no. 1, pp. 1110-1127, 2011, doi: 10.1002/etep.499.

[38] J. S. 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, 2016, doi: 10.3390/en9050367.

[39] M. R. Mahmoudi, M. Mahmoudi, and E. Nahavandi, "Testing the difference between two independent regression models," *Communications in Statistics-Theory and Methods,* vol. 45, no. 21, pp. 6284-6289, 2016, doi: 10.1080/03610926.2014.960584.

[40] A. R. Abbasi, "Fault detection and diagnosis in power transformers: a comprehensive review and classification of publications and methods," *Electric Power Systems Research,* vol. 209, p. 107990, 2022, doi: 10.1016/j.epsr.2022.107990.

[41] S. Pramanik, V. C. Duvvury, and S. Sahoo, "Tank current measurement of three-phase transformer: Its resonance behavior and sensitivity to detect mechanical faults," *IEEE Transactions on Power Delivery,* vol. 34, no. 6, pp. 2211-2218, 2019, doi: 10.1109/TPWRD.2019.2914249.

[42] Z. Zhao, C. Yao, C. Li, and S. Islam, "Detection of power transformer winding deformation using improved FRA based on binary morphology and extreme point variation," *IEEE Transactions on Industrial Electronics,* vol. 65, no. 4, pp. 3509-3519, 2017, doi: 10.1109/TIE.2017.2752135.

[43] S. Gajbhiye, "Detection of Incipient Faults in Transformer by Dissolved Gas Analysis Using Artificial Intelligence Techniques," *Available at SSRN 4344852,* 2023, doi: 10.2139/ssrn.4344852.

[44] S. Singh and M. Bandyopadhyay, "Dissolved gas analysis technique for incipient fault diagnosis in power transformers: A bibliographic survey," *IEEE Electrical insulation magazine,* vol. 26, no. 6, pp. 41-46, 2010, doi: 10.1109/MEI.2010.5599978.

[45] J. Faiz and M. Soleimani, "Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 25, no. 5, pp. 1798-1806, 2018, doi: 10.1109/TDEI.2018.007191.

[46] H. G. Zaini, "Diagnosis of the Power Transformer Faults Based on DGA Using Intelligent Classifier," *International Journal of Engineering Research and Technology,* vol. 12, no. 11, pp. 1964-1971, 2019.

[47] J. Rodríguez, J. Contreras, and C. Gaytán, "Evaluation and interpretation of dissolved gas analysis of soybean-based natural ester insulating liquid," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 28, no. 4, pp. 1343-1348, 2021, doi: 10.1109/TDEI.2021.009467.

[48] N. S. N. Klairuang, "Condition Assessment of Oil-Impregnated Paper Insulation of Transformers by Using CFA Index and Degree of Polymerization."

[49] M. Buchholz, "The Buchholz protection system and its application in practice," *ETZ,* vol. 49, no. 34, pp. 1257-1262, 1928.

[50] S. D. Myers, J. J. Kelly, and R. H. Parrish, *A guide to transformer maintenance*. Transformer Maintenance Institute, 1981.

[51] J. J. Kelly, "Transformer fault diagnosis by dissolved-gas analysis," *IEEE Transactions on Industry Applications,* no. 6, pp. 777-782, 2008, doi: 10.1109/TIA.1980.4503871.

[52] O.-F. E. Equipment, "Sampling of Gases and Analysis of Free and Dissolved Gases—Guidance," *Standard IEC,* vol. 60567, 2011.

[53] B. Vahidi and A. Teymouri, *Quality confirmation tests for power transformer insulation systems*. Springer, 2019.

[54] C. Riedmann, U. Schichler, W. Häusler, and W. Neuhold, "Online dissolved gas analysis used for transformers–possibilities, experiences, and limitations," *e & i Elektrotechnik und Informationstechnik,* vol. 139, no. 1, pp. 88-97, 2022, doi: 10.1007/s00502-022-00992-8.

[55] D. Pugh, "Advances in fault diagnosis by combustible gas analysis," in *Minutes of Forty-First International Conference of Doble Clients*, 1974.

[56] N. A. Bakar, A. Abu-Siada, and S. Islam, "A review of dissolved gas analysis measurement and interpretation techniques," *IEEE Electrical Insulation Magazine,* vol. 30, no. 3, pp. 39-49, 2014, doi: 10.1109/MEI.2014.6804740.

[57] A. Krontiris, "Fuzzy systems for condition assessment of equipment in electric power systems," 2012.

[58] E. Hubacher, "Analysis of dissolved gas in transformer oil to evaluate equipment condition," in *PCEA Engineering and Operation Section Spring Conference*, 1976.

[59] A. Mollmann and B. Pahlavanpour, "New guidelines for interpretation of dissolved gas analysis in oil-filled transformers," *Electra,* vol. 186, pp. 31-51, 1999.

[60] B. Fallou, "Detection of and research for the characteristics of an incipient fault from analysis of dissolved gases in the oil of an insulation," *Electra,* vol. 42, no. 3, pp. 31-52, 1975.

[61] I. Davidenko and K. Ovchinnikov, "Identification of transformer defects via analyzing gases dissolved in oil," *Russian Electrical Engineering,* vol. 90, no. 4, pp. 338-343, 2019, doi: 10.3103/S1068371219040035.

[62] E. Dornenburg and W. Strittmatter, "Monitoring oil-cooled transformers by gas-analysis," *Brown Boveri Review,* vol. 61, no. 5, pp. 238-247, 1974.

[63] S. Chakravorti, D. Dey, and B. Chatterjee, "Recent trends in the condition monitoring of transformers," *Power Systems,* 2013.

[64] R. R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis," *IEEE transactions on electrical insulation,* no. 5, pp. 349-354, 1978.

[65] S. Babukutty and S. Khule, "Combined Dissolved Gas Analysis: A Prescient Methodology for Recognizing Faults in Transformer by MATLAB GUI," *International Research Journal of Engineering and Technology (IRJET),* 2019.

[66] I. IEC, "599-Interpretation of the Analysis of Gases in Transformer and Other Oil-Filled Electrical Equipment in Service," *International Electro-technical Commission, Geneva, Switzerland,* 1978.

[67] L. Cheng and T. Yu, "Dissolved gas analysis principle-based intelligent approaches to fault diagnosis and decision making for large oil-immersed power transformers: A survey," *Energies,* vol. 11, no. 4, p. 913, 2018, doi: 10.3390/en11040913.

[68] S. EE, "46.501: Diagnosis oil-filled transformer equipment based on the results of chromatographic analysis of free gas with gas relay selected, i gases dissolved in insulating oil," *Kyiv, Ukrainian,* 2007.

[69] O. Shutenko and O. Kulyk, "Comparative analysis of new methods for defect type recognition by dissolved gas analysis," in *2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, 2022: IEEE, pp. 1-6, doi: 10.1109/KhPIWeek57572.2022.9916319.

[70] S.-w. Kim, S.-j. Kim, H.-d. Seo, J.-r. Jung, H.-j. Yang, and M. Duval, "New methods of DGA diagnosis using IEC TC 10 and related databases Part 1: application of gas-ratio combinations," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 20, no. 2, pp. 685-690, 2013, doi: 10.1109/TDEI.2013.6508773.

[71] O. E. Gouda, S. H. El-Hoshy, and H. H. EL-Tamaly, "Proposed three ratios technique for the interpretation of mineral oil transformers based dissolved gas analysis," *IET Generation, Transmission & Distribution,* vol. 12, no. 11, pp. 2650-2661, 2018, doi: 10.1049/iet-gtd.2017.1927.

[72] A. Nanfak, C. H. Kom, and S. Eke, "Hybrid Method for Power Transformers Faults Diagnosis Based on Ensemble Bagged Tree Classification and Training Subsets Using Rogers and Gouda Ratios," *International Journal of Intelligent Engineering & Systems,* vol. 15, no. 5, 2022, doi: 10.22266/ijies2022.1031.02.

[73] M. Duval, "Fault gases formed in oil-filled breathing EHV power transformers-Interpretation of gas-analysis data," in *IEEE Transactions on Power Apparatus and Systems*, 1974, no. 6: IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC 345 E 47TH ST, NEW YORK, NY …, pp. 1745-1746.

[74] О. С. Кулик, "Analysis of the diagnostic criteria used to defect type recognition based on the results of analysis of gases dissolved in oil," *Вісник Національного технічного університету «ХПІ». Серія: Енергетика: надійність та енергоефективність,* no. 1, pp. 15-25, 2020, doi: 10.20998/2224-0349.2020.01.03.

[75] A. Gupta, K. Jain, Y. R. Sood, and N. K. Sharma, "Comparative study of duval triangle with the new DGA interpretation scheme," in *Advances in System Optimization and Control: Select Proceedings of ICAEDC 2017*: Springer, 2018, pp. 261-268.

[76] M. S. Ali, A. Omar, A. S. A. Jaafar, and S. H. Mohamed, "Conventional methods of dissolved gas analysis using oil-immersed power transformer for fault diagnosis: A review," *Electric Power Systems Research,* vol. 216, p. 109064, 2023, doi: 10.1016/j.epsr.2022.109064.

[77] M. Duval and L. Lamarre, "The duval pentagon-a new complementary tool for the interpretation of dissolved gas analysis in transformers," *IEEE Electrical Insulation Magazine,* vol. 30, no. 6, pp. 9-12, 2014, doi: 10.1109/MEI.2014.6943428.

[78] D.-E. A. Mansour, "A new graphical technique for the interpretation of dissolved gas analysis in power transformers," in *2012 annual report conference on electrical insulation and dielectric phenomena*, 2012: IEEE, pp. 195-198, doi: 10.1109/CEIDP.2012.6378754.

[79] D.-E. A. Mansour, "Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 22, no. 5, pp. 2507-2512, 2015, doi: 10.1109/TDEI.2015.004999.

[80] O. E. Gouda, S. H. El-Hoshy, and H. H. EL-Tamaly, "Condition assessment of power transformers based on dissolved gas analysis," *IET Generation, Transmission & Distribution,* vol. 13, no. 12, pp. 2299-2310, 2019.

[81] O. E. Gouda, S. H. El-Hoshy, and H. H. El-Tamaly, "Proposed heptagon graph for DGA interpretation of oil transformers," *IET Generation, Transmission & Distribution,* vol. 12, no. 2, pp. 490-498, 2018, doi: 10.1049/iet-gtd.2017.0826.

[82] F. F. Selim, M. A. Hassan, A. M. Azmy, and E. G. Atiya, "Graphical shape in Cartesian plane based on dissolved gas analysis for power transformer incipient faults discrimination," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 32, no. 1, pp. 589-597, 2024, doi: 10.1109/TDEI.2024.3404365.

[83] S. S. Ghoneim and I. B. 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, doi: 10.1016/j.ijepes.2016.02.018.

[84] M. M. Emara, G. D. Peppas, and I. F. Gonos, "Two graphical shapes based on DGA for power transformer fault types discrimination," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 28, no. 3, pp. 981-987, 2021, doi: 10.1109/TDEI.2021.009415.

[85] A. Nanfak, S. Eke, C. H. Kom, R. Mouangue, and I. Fofana, "Interpreting dissolved gases in transformer oil: A new method based on the analysis of labelled fault data," *IET Generation, Transmission & Distribution,* vol. 15, no. 21, pp. 3032-3047, 2021, doi: 10.1049/gtd2.12239.

[86] S. A. Ward *et al.*, "Towards precise interpretation of oil transformers via novel combined techniques based on DGA and partial discharge sensors," *Sensors,* vol. 21, no. 6, p. 2223, 2021, doi: 10.3390/s21062223.

[87] I. B. Taha, A. Hoballah, and S. S. Ghoneim, "Optimal ratio limits of rogers' four-ratios and IEC 60599 code methods using particle swarm optimization fuzzy-logic approach," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 27, no. 1, pp. 222-230, 2020, doi: 10.1109/TDEI.2019.008395.

[88] F. R. Barbosa, O. M. Almeida, A. P. Braga, M. A. Amora, and S. J. Cartaxo, "Application of an artificial neural network in the use of physicochemical properties as a low cost proxy of power transformers DGA data," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 19, no. 1, pp. 239-246, 2012, doi: 10.1109/TDEI.2012.6148524.

[89] A. Hechifa, A. Lakehal, A. Nanfak, L. Saidi, and C. Labiod, "Machine learning algorithms fusion based on DGA data for improving fault diagnosis of electrical power transformer," *The Scientific Bulletin of Electrical Engineering Faculty,* vol. 23, no. 2, pp. 5-11, 2023, doi: 10.2478/sbeef-2023-0014.

[90] C. Wei, W. Tang, and Q. Wu, "Dissolved gas analysis method based on novel feature prioritisation and support vector machine," *IET Electric Power Applications,* vol. 8, no. 8, pp. 320-328, 2014, doi: 10.1049/iet-epa.2014.0085.

[91] D. Tanfilyeva, O. Tanfyev, and Y. Kazantsev, "K-nearest neighbor method for power transformers condition assessment," in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 643, no. 1: IOP Publishing, p. 012016, doi: 10.1088/1757-899X/643/1/012016.

[92] S. D. Patil, M. Dharme, A. J. Patil, G. C. AK, R. Jarial, and A. Singh, "DGA based ensemble learning and random forest models for condition assessment of transformers," in *2022*

International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), 2022: IEEE, pp. 1-6, doi: 10.1109/SMARTGENCON56628.2022.10083672.

[93] I. B. Taha, S. Ibrahim, and D.-E. A. 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, doi: 10.1109/ACCESS.2021.3102415.

[94] B. Qi, Y. Wang, P. Zhang, C. Li, and H. Wang, "A novel deep recurrent belief network model for trend prediction of transformer DGA data," *IEEE access,* vol. 7, pp. 80069-80078, 2019, doi: 10.1109/ACCESS.2019.2923063.

[95] M. S. Katooli and A. Koochaki, "Detection and classification of incipient faults in three-phase power transformer using DGA information and rule-based machine learning method," *Journal of Control, Automation and Electrical Systems,* vol. 31, no. 5, pp. 1251-1266, 2020, doi: 10.1007/s40313-020-00625-5.

[96] S. Gopakumar and T. Sree Renga Raja, "RETRACTED ARTICLE: Determination of Power Transformer Fault's Severity Based on Fuzzy Logic Model with GR, Level and DGA Interpretation," *Journal of Electrical Engineering & Technology,* vol. 19, no. 4, pp. 2527-2548, 2024, doi: 10.1007/s42835-023-01691-w.

[97] X. D. Hoang and X. H. Vu, "An improved model for detecting DGA botnets using random forest algorithm," *Information Security Journal: A Global Perspective,* vol. 31, no. 4, pp. 441-450, 2022, doi: 10.1080/19393555.2021.1934198.

[98] M. Zhang *et al.*, "Fault diagnosis of oil-immersed power transformer based on difference-mutation brain storm optimized catboost model," *Ieee Access,* vol. 9, pp. 168767-168782, 2021, doi: 10.1109/ACCESS.2021.3135283.

[99] L. Breiman, "Random forests," *Machine learning,* vol. 45, no. 1, pp. 5-32, 2001, doi: 10.1023/a:1010933404324.

[100] S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities," *Information Fusion,* vol. 64, pp. 205-237, 2020, doi: 10.1016/j.inffus.2020.07.007.

[101] N. Haque, A. Jamshed, K. Chatterjee, and S. Chatterjee, "Accurate sensing of power transformer faults from dissolved gas data using random forest classifier aided by data clustering method," *IEEE Sensors Journal,* vol. 22, no. 6, pp. 5902-5910, 2022, doi: 10.1109/JSEN.2022.3149409.

[102] B. Williams *et al.*, "Data-driven model development for cardiomyocyte production experimental failure prediction," in *Computer aided chemical engineering*, vol. 48: Elsevier, 2020, pp. 1639-1644.

[103] Z. Wahid, A. Z. Satter, A. Al Imran, and T. Bhuiyan, "Predicting absenteeism at work using tree-based learners," in *Proceedings of the 3rd international conference on machine learning and soft computing*, 2019, pp. 7-11, doi: 10.1145/3310986.331099.

[104] F. Ahmed *et al.*, "SperoPredictor: an integrated machine learning and molecular docking-based drug repurposing framework with use case of COVID-19," *Frontiers in public health,* vol. 10, p. 902123, 2022, doi: 10.3389/fpubh.2022.902123.

[105] M. R. I. Rifat, A. Al Imran, and A. Badrudduza, "Educational performance analytics of undergraduate business students," *International Journal of Modern Education and Computer Science,* vol. 10, no. 7, p. 44, 2019, doi: 10.5815/ijmecs.2019.07.05.

[106] M. Kropf *et al.*, "Cardiac anomaly detection based on time and frequency domain features using tree-based classifiers," *Physiological measurement,* vol. 39, no. 11, p. 114001, 2018, doi: 10.1088/1361-6579/aae13e.

[107] Z. Xiao, Y. Wang, K. Fu, and F. Wu, "Identifying different transportation modes from trajectory data using tree-based ensemble classifiers," *ISPRS International Journal of Geo-Information,* vol. 6, no. 2, p. 57, 2017, doi: 10.3390/ijgi6020057.

[108] L. Zhang and C. Zhan, "Machine learning in rock facies classification: An application of XGBoost," in *International Geophysical Conference, Qingdao, China, 17-20 April 2017*, 2017: Society of Exploration Geophysicists and Chinese Petroleum Society, pp. 1371-1374, doi: 10.1190/IGC2017-351.

[109] X. Zhu, H. Guo, J. J. Huang, S. Tian, W. Xu, and Y. Mai, "An ensemble machine learning model for water quality estimation in coastal area based on remote sensing imagery," *Journal of Environmental Management,* vol. 323, p. 116187, 2022, doi: 10.1016/j.jenvman.2022.116187.

[110] H. Nasiri and S. A. Alavi, "A Novel Framework Based on Deep Learning and ANOVA Feature Selection Method for Diagnosis of COVID-19 Cases from Chest X-Ray Images," *Computational intelligence and neuroscience,* vol. 2022, no. 1, p. 4694567, 2022, doi: 10.1155/2022/4694567.

[111] ETRA, "Conservation and Control of Oil-insulated Components by Diagnosis of Gas in Oil " *Electr. Coop. Res. Assoc.,* vol. 36, no. 1, 1980.

[112] E. Dornerburg and W. Strittmatter, "Monitoring oil cooled transformers by gas analysis," *Brown Boveri Review,* vol. 61, no. 5, pp. 238-247, 1974.

[113] "Mineral oil impregnated electrical equipment in service. Guide to the interpretation of dissolved free gas analysis," *2nd ed. CEI, IEC 60599,* 2007.

[114] M. Duval, "The duval triangle for load tap changers, non-mineral oils and low temperature faults in transformers," *IEEE Electrical Insulation Magazine,* vol. 24, no. 6, pp. 22-29, 2008, doi: 10.1109/MEI.2008.4665347.

[115] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the american statistical association,* vol. 32, no. 200, pp. 675-701, 1937.

[116] F. Wilcoxon, "Individual comparisons by ranking methods," in *Breakthroughs in statistics: Methodology and distribution*: Springer, 1992, pp. 196-202.

[117] S. A. Wani, D. Gupta, M. U. Farooque, and S. A. 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, doi: 10.1049/iet-smt.2018.5135.

[118] A. S. Malarvizhi, Q. Liu, D. Sha, H. Lan, and C. Yang, "An open-source workflow for spatiotemporal studies with COVID-19 as an example," *ISPRS International Journal of Geo-Information,* vol. 11, no. 1, p. 13, 2022, doi: 10.3390/ijgi11010013.

[119] "Egyptian Electricity Holding Company (EEHC) Reports, Ministry Electr.," 2016.

[120] M. Duval and A. DePabla, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electrical Insulation Magazine,* vol. 17, no. 2, pp. 31-41, 2001, doi: 10.1109/57.917529.

[121] E. Li, L. Wang, and B. Song, "Fault diagnosis of power transformers with membership degree," *IEEE Access,* vol. 7, pp. 28791-28798, 2019, doi: 10.1109/ACCESS.2019.2902299.

[122] *Egyptian Electricity Holding Company (EEHC) Reports, Ministry Electr.," 2016*.

[123] J. Li, Q. Zhang, K. Wang, J. Wang, T. Zhou, and Y. Zhang, "Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 23, no. 2, pp. 1198-1206, 2016, doi: 10.1109/TDEI.2015.005277.

[124] S. Seifeddine, B. Khmais, and C. Abdelkader, "Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network," in *2012 first international conference on renewable energies and vehicular technology*, 2012: IEEE, pp. 230-236, doi: 10.1109/REVET.2012.6195276.

[125] O. E. Gouda, S. M. Saleh, and S. H. El-Hoshy, "Power transformer incipient faults diagnosis based on dissolved gas analysis," *TELKOMNIKA Indonesian J. Electr. Eng,* vol. 17, no. 1, pp. 10-16, 2016, doi: 10.11591/telkomnika.v17i1.9405.

[126] G. Zhang, K. Yasuoka, S. Ishii, L. Yang, and Z. Yan, "Application of fuzzy equivalent matrix for fault diagnosis of oil-immersed insulation," in *Proceedings of 1999 IEEE 13th International Conference on Dielectric Liquids (ICDL'99)(Cat. No. 99CH36213)*, 1999: IEEE, pp. 400-403.

[127] J.-t. Hu, L.-x. Zhou, and M.-l. Song, "Transformer fault diagnosis method of gas hromatographic analysis using computer image analysis," in *2012 second international conference on intelligent system design and engineering application*, 2012: IEEE, pp. 1169-1172, doi: 10.1109/ISdea.2012.599.

[128] M. Rajabimendi and E. P. Dadios, "A hybrid algorithm based on neural-fuzzy system for interpretation of dissolved gas analysis in power transformers," in *TENCON 2012 IEEE region 10 conference*, 2012: IEEE, pp. 1-6, doi: 10.1109/TENCON.2012.6412171.

[129] D. S. Sarma and G. Kalyani, "ANN approach for condition monitoring of power transformers using DGA," in *2004 IEEE Region 10 Conference TENCON 2004.*, 2004, vol. 100: IEEE, pp. 444-447, doi: 10.1109/TENCON.2004.1414803.

[130] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE electrical Insulation magazine,* vol. 18, no. 3, pp. 8-17, 2002, doi: 10.1109/MEI.2002.1014963.

[131] R. Soni and K. Chaudhari, "A Novel Proposed Model to Diagnose Incipient Faults of Power Transformer Using Dissolved Gas Analysis by Ratio methods," in *Int. Conf. on Comp. of Power, Energy, Information and Communication*, 2015.

[132] A. Hechifa *et al.*, "Improved intelligent methods for power transformer fault diagnosis based on tree ensemble learning and multiple feature vector analysis," *Electrical Engineering,* vol. 106, no. 3, pp. 2575-2594, 2024, doi: 10.1007/s00202-023-02084-y.

[133] L. Lessa, C. Grilo, A. Moraes, D. Coury, and R. Fernandes, "A travelling wave-based fault locator for radial distribution systems using decision trees to mitigate multiple estimations," *Electric Power Systems Research,* vol. 223, p. 109646, 2023, doi: 10.1016/j.epsr.2023.109646.

[134] J. Liang, Z. Qin, S. Xiao, L. Ou, and X. Lin, "Efficient and secure decision tree classification for cloud-assisted online diagnosis services," *IEEE Transactions on Dependable and Secure Computing,* vol. 18, no. 4, pp. 1632-1644, 2019, doi: 10.1109/TDSC.2019.2922958.

[135] A. Joshi, R. Khathoon, and A. P. PV, "A very fast and easily implementable support vector machine based relay algorithm to classify the fault and non fault disturbance in VSC HVDC terminals," *Electric Power Systems Research,* vol. 220, p. 109380, 2023, doi: 10.1016/j.epsr.2023.109380.

[136] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 13, pp. 6308-6325, 2020, doi: 10.1109/JSTARS.2020.3026724.

[137] O. Kherif, Y. Benmahamed, M. Teguar, A. Boubakeur, and S. S. Ghoneim, "Accuracy improvement of power transformer faults diagnostic using KNN classifier with decision tree principle," *IEEE Access,* vol. 9, pp. 81693-81701, 2021, doi: 10.1109/ACCESS.2021.3086135.

[138] W. Xing and Y. Bei, "Medical health big data classification based on KNN classification algorithm," *Ieee Access,* vol. 8, pp. 28808-28819, 2019, doi: 10.1109/ACCESS.2019.2955754.

[139] M. Madhiarasan, "Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network," *Protection and Control of Modern Power Systems,* vol. 5, no. 3, pp. 1-9, 2020, doi: 10.1186/s41601-020-00166-8.

[140] A. H. Fath, F. Madanifar, and M. Abbasi, "Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems," *Petroleum,* vol. 6, no. 1, pp. 80-91, 2020, doi: 10.1016/j.petlm.2018.12.002.

[141] F. Liu, T. Dong, Q. Liu, Y. Liu, and S. Li, "Combining fuzzy clustering and improved long short-term memory neural networks for short-term load forecasting," *Electric Power Systems Research,* vol. 226, p. 109967, 2024, doi: 10.1016/j.epsr.2023.109967.

[142] C. Wang, C. Li, S. Yong, X. a. Wang, and C. Yang, "Time Series and Non-Time Series Models of Earthquake Prediction Based on AETA Data: 16-Week Real Case Study," *Applied Sciences,* vol. 12, no. 17, p. 8536, 2022, doi: 10.3390/app12178536.

[143] M. Elsaraiti and A. Merabet, "A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed," *Energies,* vol. 14, no. 20, p. 6782, 2021, doi: 10.3390/en14206782.

[144] S. K. Gunturi and D. Sarkar, "Ensemble machine learning models for the detection of energy theft," *Electric Power Systems Research,* vol. 192, p. 106904, 2021, doi: 10.1016/j.epsr.2020.106904.

[145] K. Dhibi, M. Mansouri, K. Bouzrara, H. Nounou, and M. Nounou, "Reduced neural network based ensemble approach for fault detection and diagnosis of wind energy converter systems," *Renewable Energy,* vol. 194, pp. 778-787, 2022, doi: 10.1016/j.renene.2022.05.082.

[146] Y. Tounsi, L. Hassouni, and H. Anoun, "An enhanced comparative assessment of ensemble learning for credit scoring," *Journal of Intelligent Computing Volume,* vol. 10, no. 1, p. 15, 2019, doi: 10.6025/jic/2019/10/1/15-33.

[147] S. S. Rathore and S. Kumar, "An empirical study of ensemble techniques for software fault prediction," *Applied Intelligence,* vol. 51, no. 6, pp. 3615-3644, 2021, doi: 10.1007/s10489-020-01935-6.

[148] N. Sapountzoglou, J. Lago, and B. Raison, "Fault diagnosis in low voltage smart distribution grids using gradient boosting trees," *Electric Power Systems Research,* vol. 182, p. 106254, 2020, doi: 10.1016/j.epsr.2020.106254.

[149] H. Su, Y. Yu, Q. Du, and P. Du, "Ensemble learning for hyperspectral image classification using tangent collaborative representation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 58, no. 6, pp. 3778-3790, 2020, doi: 10.1109/TGRS.2019.2957135.

[150] B. Zhao, X. Dong, Y. Guo, X. Jia, and Y. Huang, "PCA dimensionality reduction method for image classification," *Neural processing letters,* vol. 54, no. 1, pp. 347-368, 2022, doi: 10.1007/s11063-021-10632-5.

[151] E. Bergil, M. R. Bozkurt, and C. Oral, "An evaluation of the channel effect on detecting the preictal stage in patients with epilepsy," *Clinical EEG and Neuroscience,* vol. 52, no. 5, pp. 376-385, 2021, doi: 10.1177/1550059420966436.

[152] S. S. Ghoneim, I. B. Taha, and N. I. 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, doi: 10.1109/TDEI.2016.005301.

[153] I. B. Taha, D. E. A. Mansour, S. S. Ghoneim, and N. I. Elkalashy, "Conditional probability-based interpretation of dissolved gas analysis for transformer incipient faults," *IET Generation, Transmission & Distribution,* vol. 11, no. 4, pp. 943-951, 2017, doi: 10.1049/iet-gtd.2016.0886.

[154] A. Nanfak, S. Eke, F. Meghnefi, I. Fofana, G. M. Ngaleu, and C. H. Kom, "Hybrid DGA method for power transformer faults diagnosis based on evolutionary k-means clustering and dissolved gas subsets analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 30, no. 5, pp. 2421-2428, 2023, doi: 10.1109/TDEI.2023.3275119.

[155] M. Badawi *et al.*, "Reliable estimation for health index of transformer oil based on novel combined predictive maintenance techniques," *IEEE access,* vol. 10, pp. 25954-25972, 2022, doi: 10.1109/ACCESS.2022.3156102.

[156] A. Hechifa, S. Dutta, A. Lakehal, H. A. Illias, A. Nanfak, and C. Labiod, "Enhancing power transformer health assessment through dimensional reduction and ensemble approaches in Dissolved Gas Analysis," *IET Nanodielectrics,* vol. 7, no. 4, pp. 321-333, 2024, doi: 10.1049/nde2.12092.

[157] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in engineering software,* vol. 69, pp. 46-61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[158] A. Mahmoudi, I. Jlassi, A. J. M. Cardoso, and K. Yahia, "Model free predictive current control based on a grey wolf optimizer for synchronous reluctance motors," *Electronics,* vol. 11, no. 24, p. 4166, 2022, doi: 10.3390/electronics11244166.

[159] K. Dudani and A. Chudasama, "Partial discharge detection in transformer using adaptive grey wolf optimizer based acoustic emission technique," *Cogent Engineering,* vol. 3, no. 1, p. 1256083, 2016, doi: 10.1080/23311916.2016.1256083.

[160] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in engineering software,* vol. 95, pp. 51-67, 2016, doi: 10.1016/j.advengsoft.2016.01.008.

[161] S. Mirjalili, S. M. Mirjalili, S. Saremi, and S. Mirjalili, "Whale optimization algorithm: theory, literature review, and application in designing photonic crystal filters," *Nature-inspired optimizers: theories, literature reviews and applications,* pp. 219-238, 2019, doi: 10.1007/978-3-030-12127-3_13.

[162] J. A. Ardila-Rey, M. P. Cerda-Luna, C. B. Munoz, B. A. De Castro, and S. Govindarajan, "A novel E-nose system for the characterization of dissolved gases in dielectric oils," *IEEE Transactions on Instrumentation and Measurement,* vol. 72, pp. 1-16, 2023, doi: 10.1109/TIM.2023.3307177.

[163] J. Faiz and M. Soleimani, "Dissolved gas analysis evaluation in electric power transformers using conventional methods a review," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 24, no. 2, pp. 1239-1248, 2017, doi: 10.1109/TDEI.2017.005959.

[164] A. Hechifa, A. Lakehal, C. Labiod, A. Nanfak, D.-E. A. Mansour, and D. Said, "The effect of source data on graphical pentagons DGA methods for detecting incipient faults in power transformers," in *2023 International Conference on Decision Aid Sciences and Applications (DASA)*, 2023: IEEE, pp. 152-157, doi: 10.1109/DASA59624.2023.10286665.

[165] A. Hechifa, C. Labiod, A. Lakehal, A. Nanfak, and D.-E. A. Mansour, "A novel graphical method for interpretating dissolved gases and fault diagnosis in power transformer based on dynamique axes in circular form," *IEEE Transactions on Power Delivery,* 2024, doi: 10.1109/TPWRD.2024.3454230.

Appendices

Appendix A

A. *DGA samples used in the thesis and Input vectors for graphical DGA methods*

https://github.com/hechifa97/DGA-samples-used-in-the-thesis

Appendix B

B. Optimization algorithm codes for circle 1 and 2 methods

B.1. Grey Wolf Optimizer (GWO)

```
% Grey Wolf Optimizer
clear all
clc
format long
Feuille = 2;


n1 = xlsread('Data_max.xlsm',Feuille,'A2:A408'); % ('Concentration de H2')
n2 = xlsread('Data_max.xlsm',Feuille,'B2:B408'); % ('Concentration de CH4')
n3 = xlsread('Data_max.xlsm',Feuille,'C2:C408'); % ('Concentration de C2H6')
n4 = xlsread('Data_max.xlsm',Feuille,'D2:D408'); % ('Concentration de C2H4')
n5 = xlsread('Data_max.xlsm',Feuille,'E2:E408'); % ('Concentration de C2H2')
Act = xlsread('Data_max.xlsm',Feuille,'F2:F408');

earchAgents_no=40; % Number of search agents
Function_name='F3'; % Name of the test function
Max_iteration=1000; % Maximum numbef of iterations
% Load details of the selected benchmark function
[lb,ub,dim,fobj]=Get_Functions_details(Function_name);
[Best_score,Best_pos,GWO_cg_curve,Positions,BestPosition]=GWO(n1,n2,n3,n4,n5,Act,Sea
rchAgents_no,Max_iteration,lb,ub,dim,fobj);
figure('Position',[500 500 660 290])
figure(3);

%Draw objective space
%subplot(1,2,2);
semilogy(GWO_cg_curve,'Color','r')
title('Objective space')
xlabel('Iteration');
ylabel('Best score obtained so far');
axis tight
grid on
box on
legend('GWO')
figure(4);
plot(BestPosition);

display(['The best solution obtained by GWO is : ', num2str(Best_pos)]);
display(['The best optimal value of the objective funciton found by GWO is : ',
num2str(Best_score)]);
```

B.2. Whale Optimization Algorithm (WOA)

```
% Whale Optimization Algorithm (WOA)
clear all
clc
format long
Feuille = 2;


n1 = xlsread('Data_max.xlsm',Feuille,'A2:A408'); % ('Concentration de H2')
n2 = xlsread('Data_max.xlsm',Feuille,'B2:B408'); % ('Concentration de CH4')
n3 = xlsread('Data_max.xlsm',Feuille,'C2:C408'); % ('Concentration de C2H6')
n4 = xlsread('Data_max.xlsm',Feuille,'D2:D408'); % ('Concentration de C2H4')
n5 = xlsread('Data_max.xlsm',Feuille,'E2:E408'); % ('Concentration de C2H2')
Act = xlsread('Data_max.xlsm',Feuille,'F2:F408');

SearchAgents_no=30; % Number of search agents
Function_name='F3'; % Name of the test function
Max_iteration=1000; % Maximum numbef of iterations
% Load details of the selected benchmark function
[lb,ub,dim,fobj]=Get_Functions_details(Function_name);
[Best_score,Best_pos,WOA_cg_curve,Positions,BestPosition]=WOA(n1,n2,n3,n4,n5,Act,Sea
rchAgents_no,Max_iteration,lb,ub,dim,fobj);
figure('Position',[269   240   660   290])
%Draw search space
% subplot(1,2,1);
% func_plot(Function_name);
% title('Parameter space')
% xlabel('x_1');
% ylabel('x_2');
% zlabel([Function_name,'( x_1 , x_2 )'])
figure(3);

%Draw objective space
% subplot(1,2,2);
semilogy(WOA_cg_curve,'Color','r')
title('Objective space')
xlabel('Iteration');
ylabel('Best score obtained so far');
axis tight
grid on
box on
legend('WOA')
figure(4);

plot(BestPosition);
display(['The best solution obtained by WOA is : ', num2str(Best_pos)]);
display(['The best optimal value of the objective funciton found by WOA is : ',
num2str(Best_score)]);
```