

**Development of a Transformer Fault Identification Model via
Machine Learning: Integrating Gas Co-Relation and Feature
Extraction from DGA**

S M C D Samarakoon

208712C

Degree of Master of Science

In Electrical Installation

Department of Electrical Engineering

University of Moratuwa
Sri Lanka

September 2025

Development of a Transformer Fault Identification Model via Machine Learning: Integrating Gas Co-Relation and Feature Extraction from DGA

S M C D Samarakoon

208712C

Thesis was submitted to the Department of Electrical Engineering
of the University of Moratuwa in partial fulfillment of the requirements for the
Degree of Master of Science in Electrical Installation



September 2025

Supervisor: Prof. W.D.A.S. Rodrigo

1 DECLARATION OF THE CANDIDATE AND SUPERVISORS

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (Such as articles or books).

Signature:

.....

Date: 22/09/2025

S M C D Samarakoon

The above candidate has carried out research for the master's Dissertation under my supervision.

Signature of the supervisor:

.....

Date: 19/09/2025

Prof. W.D.A.S. Rodrigo

2 ABSTRACT

Dissolved Gas Analysis (DGA) is a widely established diagnostic tool for assessing the internal condition of power transformers by detecting gases generated from thermal and electrical stresses within the oil-paper insulation system. Conventional diagnostic methods such as the Duval Triangle, Rogers Ratio Method, and IEC 60599 rely on static gas ratio thresholds and predefined fault zones, which often result in ambiguous or inaccurate classifications when confronted with overlapping gas signatures, nonlinear gas interactions, and evolving transformer designs.

This research presents a transformer fault identification module that leverages machine learning (ML) integrated with an attention-based feature extraction mechanism. The proposed model enhances fault classification accuracy by learning complex, nonlinear correlations between dissolved gases, dynamically weighting their diagnostic significance, and overcoming the rigid boundaries of traditional methods. Trained and validated on real-world DGA datasets with confirmed fault types, the model successfully categorizes faults such as partial discharges, electrical arcing, and thermal faults (T1–T3).

By transforming raw DGA data into intelligent predictive insights, the system provides accurate diagnoses and actionable maintenance recommendations. The ultimate goal of this study is to advance transformer condition monitoring by offering a scalable, adaptive, and interpretable ML-driven module for fault recognition, contributing to more efficient asset management and predictive maintenance strategies.

Keywords: *Machine Learning, Artificial Interlligence, Dissolve Gas Analysis, Tranformer based ML model, Attention mechanism, Feature extraction*

3 ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my Prof. W.D.A.S. Rodrigo , for guiding me to develop this research concept and successful journey so far up to preparation of final thesis. And my heartfelt thanks for his motivation, enthusiasm and continuous support throughout my research. It is a great experience to have work with a supervisor who has good vision and understanding regarding research work.

Also, I would like to express my sincere gratitude to Mr. Nisal Amarasinghe, Chief Engineer, Transmission Asset Management & Condition Monitoring, Ceylon Electricity spending his valuable time sharing his knowledge and experience, all the electrical engineers of Transmission Asset Management & Condition Monitoring Unit, Ceylon Electricity Board for their support and valuable instructions for successful completion of my research.

Further, it is a great pleasure to remember my masters' coordinator Dr. Thushara Rathnayaka and all of my lecturers of University of Moratuwa for providing useful suggestions and valuable advice in progress reviews.

Finally, I would like to thank and express my heartfelt gratitude to my loving my parents for their unconditional support, sacrifice and understanding in supporting me always.

4 TABLE OF CONTENTS

1	DECLARATION OF THE CANDIDATE AND SUPERVISORS.....	i
2	ABSTRACT.....	ii
3	ACKNOWLEDGEMENT	iii
4	TABLE OF CONTENTS.....	iv
5	LIST OF FIGURES	vii
CHAPTER 1		1
1	INTRODCUTION	1
1.1	Importance of Power Transformer Monitoring.....	1
1.1.1	Role of Dissolved Gas Analysis (DGA)	1
1.1.2	Limitations of Traditional Diagnostic Techniques	1
1.1.3	Emergence of Machine Learning in Transformer Diagnostics	2
1.2	Problem Statement.....	2
1.3	Research Objectives.....	3
1.3.1	Main Objective	3
1.3.2	Specific Objectives	4
CHAPTER 2		5
2	LITERATURE REVIEW	5
2.1	Power Transformers and Important role and performance of the Dissolve Gas analysis 5	
2.2	The Critical Role of DGA in Transformer Diagnostics.....	6
2.2.1	DGA analyzing methods.....	7
2.2.2	Key Gas Method (Clause 5.3.1).....	7
2.2.3	Gas Ratio Method (Clause 5.3.2).....	8
2.2.4	Rate of Gas Generation Method (Clause 6.2 and Annex C)	9
2.3	Limitations of Traditional DGA Methods in Transformer Diagnostics.....	9
2.4	Limitations of Traditional DGA Methods and Need for Improvement	14
2.4.1	Inconsistent Diagnoses Across Methods.....	14
2.4.2	Overlapping Gas Signatures Between Fault Types.....	14
2.4.3	Insufficient Diagnostic Coverage.....	14
2.4.4	Lack of Adaptability and Learning Capability	14
2.4.5	No Severity or Predictive Insights	14

2.5	Need for Improved DGA Interpretation	15
2.6	Research Gap	16
2.6.1	Areas to Be Researched and Developed	16
3	Methodology.....	18
3.1	Data Collection	18
3.2	Traditional Methods Applied.....	19
3.2.1	Key Gas Method	19
3.2.2	Gas Ratio Methods.....	20
3.2.3	Rogers Ratio Method	21
3.2.4	Duval Triangle Method.....	22
3.3	Proposed Machine Learning Model.....	23
3.3.1	Rationale for ML-Based DGA Diagnosis.....	23
3.3.2	Model Architecture and Components	24
3.3.3	Data Processing and Training	25
3.3.4	Addressing Traditional Method Limitations.....	26
3.3.5	Summary	27
3.4	Model Evaluation and Performance Analysis.....	27
3.4.1	Evaluation of Metrics and Rationale.....	27
3.4.2	Validation Results.....	28
3.4.3	Confusion Matrix Analysis	31
3.4.4	Model Testing on Unseen Fault Cases.....	34
3.5	Traditional methods and ML model performance	41
3.5.1	IEC -Method overview and Diagnosis.....	42
3.5.2	Duval Triangle method diagnosis	45
3.5.3	Rogger method diagnosis.....	49
3.5.4	ML Model results.....	53
4	Discussion and results.....	57
4.1	Discussion: Overcoming Rule-Based Diagnostic Limitations Using a Transformer-Based ML Model	57
4.1.1	IEC Method Failure Cases	57
4.1.2	Dual triangle method failures.....	60
4.1.3	Roggers method failure cases	64
5	Conclusion and Recommendations.....	69

5.1	Limitations of the work.....	70
5.1.1	Limited and Imbalanced Dataset	70
5.1.2	Restricted Feature Space.....	71
5.1.3	Data Quality and Measurement Variability	71
5.1.4	Generalization and External Validity.....	72
5.1.5	Lack of User Interface and Deployment Framework.....	72
5.2	Future Work.....	72
5.2.1	Expansion and Balancing of the Dataset.....	72
5.2.2	Analysis of DGA Trends and Gas Generation Rates	73
5.2.3	Integration of Additional Diagnostic Indicators	74
5.2.4	Development of a User-Friendly Interface	74
5.2.5	Implementation of Online Learning and Periodic Updates.....	74

5 LIST OF FIGURES

Figure 1 DGA Interpretation chart form IEC60599.....	8
Figure 2 Proposed ML model - Transformer Based	25
Figure 3 Model Validation Results	28
Figure 4 Model training and validation loss	29
Figure 5 Model training accuracy and validation	30
Figure 6 Model validation confusion matrix.....	32
Figure 7 Classification report - Training and validation.....	33
Figure 8 DGA results for known fault types.....	37
Figure 9 classification report for 105 unseen data	37
Figure 10 confusion matrix for the 105 unseen data.....	40
Figure 11 ML Model classification report for 49 unseen data.....	53
Figure 12 Model Output for Case No 8	58
Figure 13 Model Output for Case No 7	59
Figure 14 Model Output for Case No 9	60
Figure 15 Model Output for Case No 2	61
Figure 16 Model Output for Case No 3	62
Figure 17 Model Output for Case No 9	62
Figure 18 Model Output for Case No 10	63
Figure 19 Model Output for Case No 23	64
Figure 20 Model Output for Case No 2	65
Figure 21 Model Output for Case No 3	66
Figure 22 Model Output for Case No 8	66
Figure 23 Model Output for Case No 10	67

6 LIST of TABLES

Table 1 - DGA analyze methods summery	22
Table 2 - Precision, Recall, F1 -Score summary.....	38
Table 3 Correct prediction summary for 105 unseen data	40
Table 4 Faut class used in ML model	42
Table 5 IEC ratio method interpretation	43
Table 6 DGA samples for IEC method analysis	43
Table 7 Caparison between D1 and D2 class.....	45
Table 8 Duval triangle method analysis summery	49
Table 9 Limitation and description of the Rogers method.....	51
Table 10 Rogers method analysis summery	52
Table 11 Summery of the output -49 unseen data	54
Table 12 Correct predations for 49 unseen data	55
Table 13 Fault classification notation in ML model	57
Table 14 ML model and traditional method Results summery table.....	68