electronics

MDPI

Article

Application of Logistic Regression Algorithm in the Interpretation of Dissolved Gas Analysis for Power Transformers

Yousuf D. Almoallem [1], Ibrahim B. M. Taha [2], Mohamed I. Mosaad [1], Lara Nahma [3] and Ahmed Abu-Siada [3,*]

[1] Electrical & Electronics Engineering Technology Department, Royal Commission Yanbu Colleges & Institutes, Yanbu Industrial City 46452, Saudi Arabia; ymoallem@gmail.com (Y.D.A.); m_i_mosaad@hotmail.com (M.I.M.)
[2] Electrical Engineering Department, College of Engineering, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; i.taha@tu.edu.sa
[3] School of Electrical Engineering and Computing, Curtin University, Perth 6845, Australia; lara.al-ibreesm@curtin.edu.au
* Correspondence: a.abusiada@curtin.edu.au

Abstract: Dissolved gas analysis (DGA) is one of the regular routine tests accepted by worldwide utilities to detect power transformer incipient faults. While the DGA measurement has fully matured since the development of offline and online sensors, interpretation of the DGA results still calls for advanced approaches to automate and standardize the process. Current industry practice relies on various interpretation techniques that are reported to be inconsistent and, in some cases, unreliable. This paper presents a new application for the advanced logistic regression algorithm to improve the reliability of the DGA interpretation process. In this regard, regularized logistic regression is used to improve the accuracy of the DGA interpretation process. Results reveal the superior features of the proposed logistic regression approach over the conventional and artificial intelligence techniques presented in the literature.

Keywords: power transformers; dissolved gas analysis; fault diagnosis; logistic regression algorithm

Citation: Almoallem, Y.D.; Taha, I.B.M.; Mosaad, M.I.; Nahma, L.; Abu-Siada, A. Application of Logistic Regression Algorithm in the Interpretation of Dissolved Gas Analysis for Power Transformers. *Electronics* **2021**, *10*, 1206. https://doi.org/10.3390/electronics10101206

Academic Editor: Giampaolo Buticchi

Received: 20 April 2021
Accepted: 13 May 2021
Published: 19 May 2021

1. Introduction

In electricity chains, power transformers represent vital links that ensure the reliability of the entire grid. Due to the significant electrical and thermal stress that an operating power transformer insulation system experiences, global utilities have adopted various condition monitoring and fault diagnosis schemes to detect incipient faults and take a corrective action to avoid any potential consequences. Dissolved gas analysis (DGA) is one of the most reliable measurements conducted on transformer oil samples either offline or online [1]. The technique was developed based on the fact that under high thermal stress, insulating oil and paper decompose and release some chemical by-products that dissolve in the oil and reduce its dielectric stress [2]. These by-products include hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), ethane (C_2H_6), carbon monoxide (CO), and carbon dioxide (CO_2) [3]. Moreover, furan compounds are produced due to cellulous degradation [4,5]. As these by-products are generated at particular temperature ranges, the amount and type of the dissolved gases in the transformer oil can be used to identify the health condition of the transformer [6]. For example, increased amount of H_2 is an indication of partial discharge activity while thermal faults can be identified by the amount of C_2H_4 [7]. Arcing fault generates all gases, including traceable amounts of C_2H_2 [8]. Measurement of such gases can be conducted in a laboratory-based environment using gas chromatography–mass spectroscopy, or through a diverse range of online sensors [9]. While such measurement techniques are fully matured and being improved with the advancement in electronic sensors technology, analyses of the results are still calling for

more research in order to standardize and automate the entire DGA process. Current industry practice relies on some common DGA interpretation techniques that are briefly presented below.

Key gas method: This method employs the absolute values of the individual gases along with the total combustible gas concentration (TCGC) to identify the risk level within the power transformer that is categorized into four conditions as shown in Table 1 [10]. While the application of this method is straightforward, it is not widely accepted due to its conservative nature, as the gas evolution is not considered. A transformer may be reported healthy if all gases are less than the threshold limits specified by the normal condition while a particular gas is evolving rapidly. On the other hand, a transformer may be classified as being at risk because one individual gas or more exceed the normal limit; however, the transformer can still be considered as not being at risk as long as this gas is not continuously increasing.

Table 1. Dissolved key gas concentration limits (ppm).

Status	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	CO	CO_2	TDCG
Normal	100	120	35	50	65	350	2500	720
Modest Concern	101–700	121–400	36–50	51–100	66–100	351–570	2500–4000	721–1920
Major Concern	701–1800	401–1000	51–80	101–200	101–150	571–1400	4001–10,000	1921–4630
Imminent Risk	>1800	>1000	>80	>200	>150	>1400	>10,000	>4630

Doernenburg ratio method: This method utilizes four ratios to identify the thermal partial discharge (PD) and arcing faults within the transformer as per the code in Table 2 [7]. This method cannot be employed unless the concentration of at least one of the gases used in the ratios (H_2, C_2H_4, CH_4, C_2H_6, and C_2H_2) exceeds twice the corresponding limit L1 shown in Table 3 [7].

Table 2. Doernenburg-C ratio method.

Suggested Fault Diagnosis	CH_4/H_2		C_2H_2/C_2H_4		C_2H_2/CH_4		C_2H_6/C_2H_2	
	Oil	Gas Space	Oil	Gas Space	Oil	Gas Space	Oil	Gas Space
Thermal fault	>1	>0.1	<0.75	<1	<0.3	<0.1	>0.4	>0.2
PD	<0.1	<0.01	Not significant		<0.3	<0.1	>0.4	>0.2
Arcing	>0.1–<1	>0.01–<0.1	>0.75	>1	>0.3	>0.1	<0.4	<0.2

Table 3. Limit (L1) for the Doernenburg-c ratio method.

Key Gas	L1 Concentrations (ppm)
Hydrogen (H_2)	100
Methane (CH_4)	120
Acetylene (C_2H_2)	35
Ethylene (C_2H_4)	50
Ethane (C_2H_6)	65

Rogers ratio method: This method employs three-key gas ratios as shown in Table 4 and was developed based on Doernenburg's method [11]. Similarly to the Doernenburg ratio method, the Rogers ratio method may result in out-of-code values for some DGA cases when the gases concentration used in the ratios are not large enough.

Table 4. Faults classification with Rogers' ratio method.

C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6	Range of Gas Ratio
0	1	0	<0.1
1	0	0	0.1–1
1	2	1	1–3
2	2	2	>3
Characteristic Fault			
0	0	0	Normal ageing
2	1	0	Partial discharge of low energy density
1	1	0	Partial discharge of high energy density
1–2	0	1–2	Continuous sparking
1	0	2	Discharge of high energy
0	0	1	Thermal fault of low temp <150 °C
0	2	0	Thermal fault of low temp between 150–300 °C
0	2	1	Thermal fault of medium temp between 300–700 °C
0	2	2	Thermal fault of high temp >700 °C

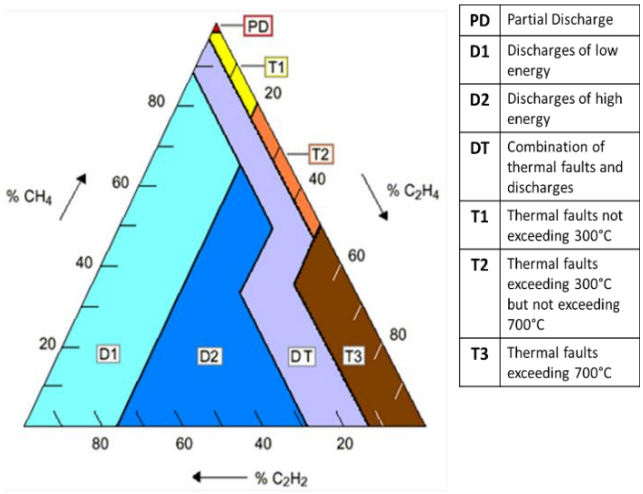
IEC ratio method: This method uses the same ratios as Rogers', but with ratio ranges and analysis as shown in Table 5 [12]. While this method showed some improvement, it still suffers from the common drawback of all ratio methods: a significant amount of gases used in the ratios must exist to result in a valid code; otherwise, interpretation of DGA results cannot be conducted using this method.

Table 5. IEC ratio method.

Case	Characteristic Fault	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
PD	Partial discharges	NS	<0.1	<0.2
D1	Discharges of low energy	>1	0.1–0.5	>1
D2	High energy discharges	0.6–2.5	0.1–1	>2
T1	Thermal faults <300 °C	NS	>1 but NS	<1
T2	Thermal faults >300 °C and <700 °C	<0.1	>1	1–4
T3	Thermal faults >700 °C	<0.2	>1	>4

NS = Not significant whatever the value.

Duval Triangle Method: This is a graphical method developed by Duval to analyze DGA data using three gases, CH_4, C_2H_2, and C_2H_4, which are plotted along three sides of a triangle as shown in Figure 1 [13]. The triangle is divided into seven zones, indicating various transformer faults including partial discharge, thermal faults at various temperatures, and electric arcing. While the Duval triangle provides more accurate diagnoses than the above ratio methods, it does not encompass a fault-free zone; hence, this method cannot be used to detect incipient faults.



PD	Partial Discharge
D1	Discharges of low energy
D2	Discharges of high energy
DT	Combination of thermal faults and discharges
T1	Thermal faults not exceeding 300°C
T2	Thermal faults exceeding 300°C but not exceeding 700°C
T3	Thermal faults exceeding 700°C

Figure 1. Duval triangle and associated fault zones.

As can be seen from the above discussion, all existing DGA interpretation techniques comprise shortcomings that make them unreliable to some extent. As such, DGA must be conducted and analyzed by expert personnel. This makes the DGA interpretation process inconsistent and different conclusions may be reported for the same oil sample if analyzed by different personnel. This inspired researchers to develop artificial intelligence-based techniques for DGA interpretation, including fuzzy logic [14], which is also employed to identify transformer criticality and remnant life based on DGA data [15,16], neural network [17], gene expression programming [18,19], support vector machine [20], and particle swarm optimization [21]. While AI-based DGA interpretation techniques published so far in the literature provide more reliable diagnoses than the conventional interpretation methods, they fail to diagnose some faults in oil and cellulose, and engineering judgment is still required [22].

2. Proposed DGA Machine Learning Technique

Logistic regression is widely employed in many applications to improve machine learning approaches, as it is suitable for systems of discrete and historical data such as DGA [23]. As the logistic regression is a method for classifying data into discrete outcomes, it is an ideal method for DGA applications. In supervised learning algorithms, overfitting is a likely problem with many input features. In un-regularized models such as fuzzy logic models, tuning is mainly conducted through training-error minimization. On the other hand, regularized logistic regression methods are widely used to solve problems with numerous features [24]. In these methods, the outcomes are classified using a cost function that is solved by logistic regression [25]. The regularization helps to avoid over-fitting due to either small number of training samples or large number of features, and it is often used for proper feature selection by filtering out irrelevant features [26].

Some algorithms employ conditional maximum entropy models such as the generalized iterative scaling (GIS) [27]. However, this algorithm is considered as one-sided Laplacian prior which can be extended towards the regularized logistic regression. Another AI-based technique is called grafting, which consists in steadily constructing a subclass of the parameters [28]. Moreover, the generalized LASSO is an algorithm developed based on the regularized least squares problem [29].

This paper is taking a forward step into the development of reliable and automated DGA interpretation techniques that can be continuously enhanced through self-learning processes based on historical and future DGA data. One of the cutting-edge techniques of artificial intelligence is deployed to understand the DAG data through machines without the need of human intervention. The regularized logistic regression is implemented based on the iteratively reweighted least squares (IRLS). The technique utilizes a one-vs-all

method in order to analyze and then classify DGA data into one fault out of a set of possible transformer faults.

Regularized logistic regression is a distinguished tool utilized in machine learning. The proposed algorithm is designed to classify the condition of DGA oil samples automatically into designated faults. Classified faults include partial discharge (PD), low and high energy discharge (D1 and D2, respectively), and low, medium, and high thermal faults (T1, T2, and T3, respectively). DGA measurements are defined based on five features that are used as inputs to the proposed machine learning algorithm. DGA results, including H_2, CH_4, C_2H_6, C_2H_4, and C_2H_2 are fed into the algorithm as a percentage of the total concentration of these gases (gas-ratio). The proposed model is developed using 446 DGA samples collected from the literature and Egypt electric utilities, and are divided into two sets. The first set (335 samples) is used for training processes including model validation (67 samples), while the second set (111 samples) is used for testing the developed model. Table 6 shows the actual fault classifications of the collected 446 DGA samples.

Table 6. Dataset samples per each fault type for training and testing sets.

	PD	D1	D2	T1	T2	T3	Total
Training set	27	50	79	69	40	70	335
Test set	8	17	26	24	13	23	111

The proposed model is based on a regularized supervised learning approach in which all data are labelled. The number of labelled samples used in this model is given by $m = 446$. Each collected sample is represented by x and y coordinates $\left[\left(x^{(i)}, \ y^{(i)} \right); \ i = 1, \cdots, m \right]$ that represent the input and output, respectively. Each input is represented as an N-dimensional vector $\left[x^{(i)} \in \mathbb{R}^N \right]$, $N = 5$, that represents the input feature vector of the DGA five-gas ratios. Moreover, each output is a class labeled $\left[y^{(i)} \in \{1, 2, \ldots, 6\} \right]$ that represents one of the possible six fault types mentioned above.

As shown in Figure 2, the data preparation stage includes two steps: data normalization and samples shuffling. In the first step, the collected samples are normalized to ensure the input features (DGA measurements) have fair influence on the output. In the second step, the collected samples are shuffled to remove any undesired pre-order effects.

After the preparation stage, the collected samples are divided into three sets as follows:

(1) The training set, consisting of 268 samples or 60% of the entire data set, is utilized to form the initial model;
(2) The validation set, consisting of 67 samples or 15% of the entire data set, is utilized to optimize the regularization parameter (λ) to form the final model;
(3) The testing set, consisting of 111 samples or 25% of the entire data set, is utilized for testing the final model.

It is worth mentioning that while data splitting is performed randomly, all data classes, i.e., all fault types, are included in each data set. The percentages of data sets were chosen based on the usual practices used in the literature [30].

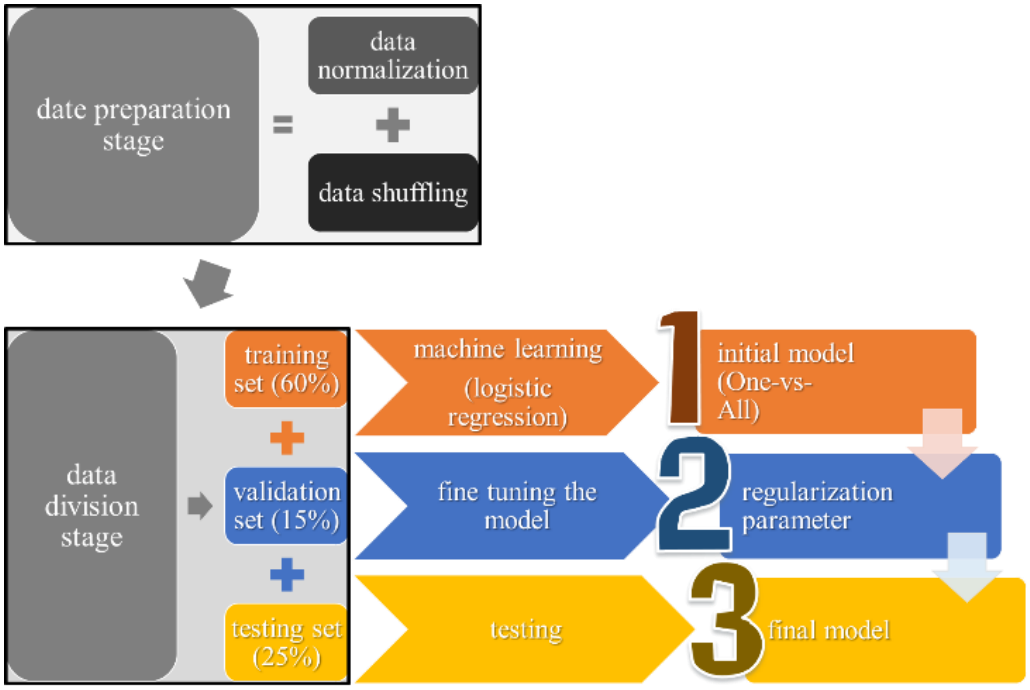


Figure 2. Flowchart of the regularized logistic regression.

For each output class $y^{(i)}$, the one-vs-all logistic regression method is used to train a hypothesis classifier $h_\theta^{(i)}(x)$. In this method, the hypothesis calculates the probability of the output corresponding to one of the possible faults, e.g., $y = i$. This process is repeated for all samples to calculate the probability of each fault corresponding to the used DGA sample and the output with the highest probability considered as the classified fault. The cost function for such logistic regression J as a function of model parameters θ is given by:

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} Cost\left(h_\theta\left(x^{(i)}\right).y^{(i)}\right) + \lambda\sum_{j=1}^{n}\theta_j^2 \tag{1}$$

$$Cost(h_\theta(x).y) = \begin{cases} -\log(h_\theta(x)) & y = i \\ -\log(1 - h_\theta(x)) & y \neq i \end{cases} \tag{2}$$

The hypothesis classifier defines the probability distribution for each class label $y = i$ given a feature vector x as follows:

$$h_\theta(x) = p(y = i|x;\theta) = g\left(\theta^T x\right) = \frac{1}{1 + e^{-\theta^T x}} \tag{3}$$

The classifier is trained for each output class i by utilizing the training sample set. Therefore, for a given input feature x, the algorithm optimizes the classifier to predict a class i that maximizes the hypothesis $\max_\theta h_\theta^{(i)}$. A developed MATLAB code is utilized to train the classifier by minimizing the cost function $J(\theta)$.

The first term is always positive since the hypothesis classifier goes only from 0 to 1. The second term in Equation (1) represents the regularization term that is used to avoid model overfitting. This term is optimized using the validation sample set. In Equation (3), all model parameters are penalized with a ratio that minimizes the cost function with the λ parameter. After spanning through a range of values for λ from 0.001 to 10, the obtained optimum value is $\lambda = 1$ as shown in Table 7 and Figure 3. The training error is still acceptable at this value while a significant reduction is achieved in the validation error.

Table 7. Selection of the regularization parameter (λ) using cross validation samples.

λ	Training % Error	Validation % Error
0.000	3.358	13.433
0.001	3.358	14.925
0.003	3.731	13.433
0.010	4.104	13.433
0.030	3.731	11.940
0.100	4.478	11.940
0.300	5.224	11.940
1.000	6.343	8.955
3.000	11.194	11.940
10.000	13.806	16.418

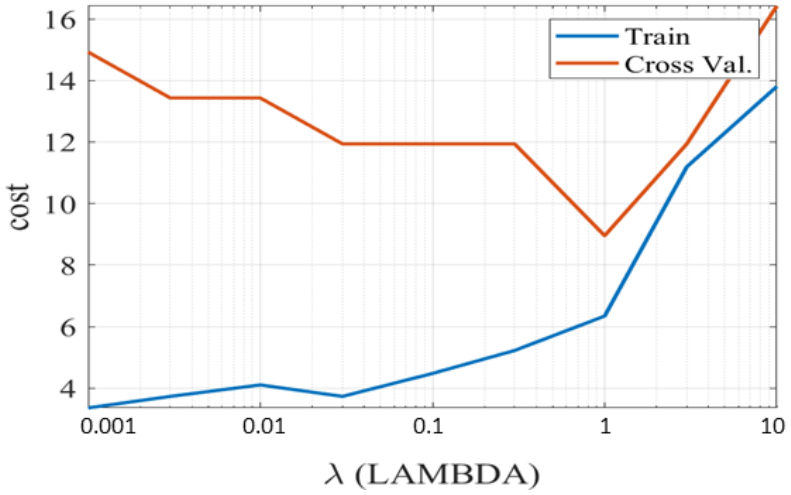


Figure 3. Selection of the regularization parameter using cross validation samples.

The algorithm process goes through three main steps. Firstly, the training sample set is used to train the model and find the model parameter θ_j that correlates the input and output factors. Secondly, the validation sample set is utilized to find the proper regularization parameter (λ). Thirdly, the test sample set is used to test the model which was kept apart from the system modeling to ensure proper independent validation.

3. Results and Nonlinear Approximation

The prediction accuracy (η) can be estimated as follows:

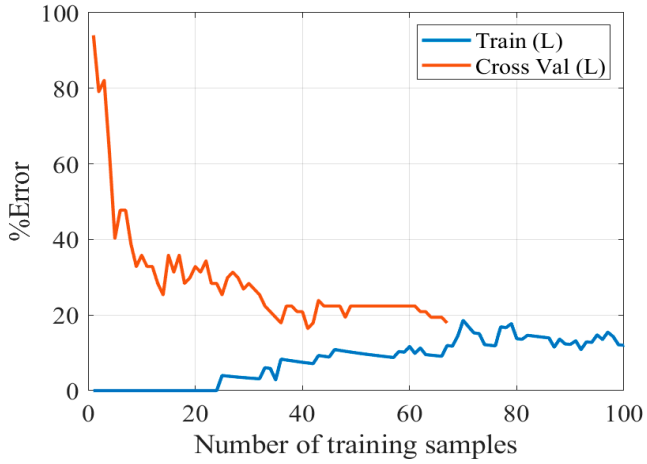
$$\%\eta = \frac{Total\ number\ of\ correct\ predicting\ samples}{Total\ number\ of\ samples\ used} \times 100 \quad (4)$$

Using the basic input five features, the algorithm can be used to predict the output fault with an accuracy of 82.9% ($\lambda = 0$). After initial tuning of the regularization parameter, the system accuracy is slightly increased to 83.8% ($\lambda = 0.01$). When the inputs are used as percentages instead of absolute values, the accuracy (for test samples) is improved as can be seen in Table 8. The polynomial regression shows an increase in the accuracy up to 86%.

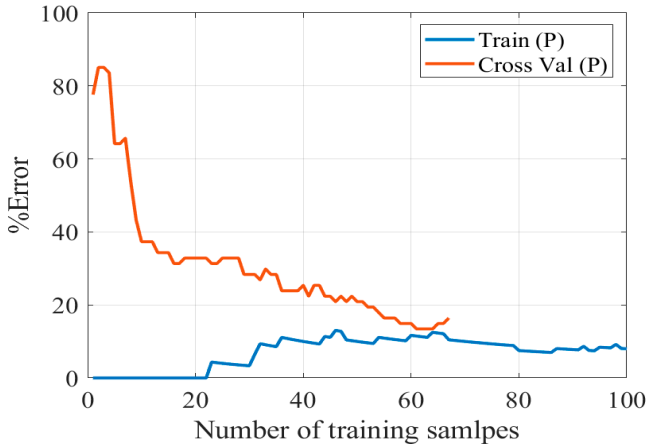
Table 8. Error corresponding to different polynomial orders.

p	Number of Features	Percentage Error
4	80	17.9%
6	180	16.4%
8	320	9.0%
10	500	21.0%
12	720	16.4%

The learning curves with the prediction error ($Error = 1 - \eta$) for the linear approximation of the proposed model is shown in Figure 4a. At the beginning, the training set comprises low error since the system can easily approximate the function over very few samples, but as the number of samples increases, the training error also increases but settles at a level of less than 20% when the number of samples increases to 100. On the other hand, the error of the validation set is high when a few samples are used, but it drops as the number of samples increases.



(**a**)



(**b**)

Figure 4. Learning curve for logistic regression (L: linear, P: polynomial) (**a**) Linear combination of features; (**b**) Polynomial combination of features (p=8).

Figure 4a,b show the linear and polynomial learning profiles for the logistic regression; respectively. The error results shown in this figure reveal the nonlinear characteristic

of the investigated problem. Hence, it is not accurate to express the system using a linear combination of the features. One way to approximate the nonlinear feature of the investigated problem is through using polynomials. As noticed, by increasing the number of samples, the cross-validation error decreases. Moreover, the training error curve shows that the model is biased and, therefore, additional features are required where the nonlinear combination of features is considered. To select an optimum polynomial order, p is changed in the range 4–12 and the number of features along with the corresponding error are recorded in Table 8. It can be seen that for p = 8, the lowest percentage error is obtained. For values of p above 8, the error increases, again showing an overfitting problem at p = 10.

The regularization term is more prominent when the number of features increases as in the polynomial features case. The total number of features for 8th order polynomial (n_{poly}) with five independent variables ($n = 5$) including homogeneous ($n_0 = n \times p = 5 \times 8 = 40$) and nonhomogeneous n_x terms is calculated as follows:

$$n_{poly} = n_0 + n_x \tag{5}$$

$$n_x = \sum_{j=1}^{p-1} j \times \sum_{i=1}^{n-1} i \tag{6}$$

The nonhomogeneous term is given by $n_x = 28 \times 10 = 280$ and hence the total n_{poly} reaches 320 features.

Results of the algorithm using the polynomial features can achieve an accuracy of 86%, as shown in the training curve in Figure 4b. The last column of Table 9 demonstrates the predicted faults using polynomial features for logistic regression with testing dataset samples.

Table 9. Error of the linear and polynomial regression models.

Sample Set	Training Error	Validation Error	Testing Error
Linear regression	15.2%	13.4%	17.1%
Polynomial regression	6.3%	9.0%	14.4%

Table 10 presents a summary of the obtained results for each fault using the testing dataset (111 samples). In order to highlight the system capability to detect various faults using DGA results, the predicted fault types using polynomial regression are compared with the actual faults. Results attest the capability of the proposed model in detecting different fault types, especially low and high thermal faults (T1 and T3) with a high degree of accuracy. As can be seen in Table 10, the overall prediction accuracy of the testing samples is 85.6%.

Table 10. Results of the proposed model using 111 testing samples.

	Actual	PD	D1	D2	T1	T2	T3	%η
PD	8	7	0	0	0	1	0	87.5
D1	17	1	12	3	1	0	0	70.6
D2	26	0	4	21	1	0	0	80.5
T1	24	0	0	0	23	0	1	95.8
T2	13	0	0	0	2	11	0	84.6
T3	23	0	0	0	0	2	21	91.3
All	111	8	16	24	28	13	22	85.6

4. Model Validation

The proposed model is validated by comparing its output (for the 111 testing samples) with conventional and AI-based techniques recently published in the literature. Conventional methods include the Duval triangle, IEC, and Rogers 4-ratio methods. AI methods

include the clustering method [30], conditional probability [31], and modified IEC and Rogers 4-ratio methods [32]. Table 11 indicates the number of samples successfully predicted by each method and the actual number of samples for each fault. Overall, the proposed method in this paper achieves the highest prediction accuracy (85.6%), when compared to all other methods investigated in the table.

Table 11. Comparison between the proposed model and other methods.

	PD	D1	D2	T1	T2	T3	%η
Actual	**8**	**17**	**26**	**24**	**13**	**23**	
Duval	6	16	23	13	7	22	78.4
IEC	4	10	16	16	12	18	68.5
Rog	4	0	19	21	7	13	57.7
Mod-Rog	7	11	22	21	12	21	84.7
Mod-IEC	7	9	22	21	12	23	84.7
Prob.	8	7	23	20	13	21	82.9
Cluster	7	10	20	23	4	21	76.6
Proposed	7	12	21	23	11	21	85.6

5. Conclusions

A logistic regression algorithm is implemented in this paper to improve the reliability and accuracy of the DGA interpretation process. The comparison between linear and polynomial logistic regression algorithms for predicting transformer fault types based on DGA results shows the better performance of the polynomial algorithm. The logistic regression parameter is estimated to obtain high prediction accuracy during training and testing stages. The proposed model is designed based on 335 training and 111 testing DGA samples of pre-known fault types. Two validation techniques are used for the proposed model. These are the cross-validation technique during the training stage and a comparison with other methods in the literatures. The results reveal the high prediction accuracy of the proposed logistic regression technique (85.6%) over conventional and recently published techniques in the literature.

Author Contributions: Y.D.A.: Software, Validation, Conceptualization, Methodology, and visualization. I.B.M.T.: Conceptualization, Software, Validation, Formal analysis, Writing original draft. M.I.M.: Methodology, Software, Validation, Investigation, Resources, Data curation, Writing-review & editing. L.N.: Investigation, editing. A.A.-S.: Methodology, Software, Validation, Investigation, Resources, Data curation, Writing-review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to acknowledge the financial support received from Taif University Researchers Supporting Project Number (TURSP-2020/61), Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Norazhar, A.B.; Abu-Siada, A.; Islam, S. A Review of Dissolved Gas Analysis Measurement and Interpretation Techniques. *IEEE Electr. Insul. Mag.* **2014**, *30*, 39–49.
2. STM. *Standard Test Method for Analysis of Gases Dissolved in Electrical Insulating Oil by Gas Chromatography*; ASTM D3612-02; American Society for Testing and Materials: West Conshohocken, PA, USA, 2009.
3. Abu-Siada, A.; Islam, S. A new approach to identify power transformer criticality and asset management decision based on dissolved gas-in-oil analysis. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 1007–1012. [CrossRef]
4. Abu-Siada, A. Correlation of furan concentration and spectral response of transformer oil-using expert systems. *IET Sci. Meas. Technol.* **2011**, *5*, 183–188. [CrossRef]
5. Norazhar, A.B.; Abu-Siada, A.; Islam, S. A Review on Chemical Diagnosis Techniques for Transformer Paper Insulation Degradation. In Proceedings of the Australasian Universities Power Engineering Conference, AUPEC'13, Hobart, Australia, 29 September–3 October 2013.

6. IEEE. *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*; IEEE Std C57.104-1991; IEEE: Piscataway, NJ, USA, 1992.
7. IEEE. *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*; IEEE Std C57.104-2008 (Revision of IEEE Std C57.104-1991); IEEE: Piscataway, NJ, USA, 2019; p. C1-27.
8. IEEE. *IEEE Guide for the Detection and Determination of Generated Gases in Oil-Immersed Transformers and Their Relation to the Serviceability of the Equipment*; ANSI/IEEE Std C57.104-1978; IEEE: Piscataway, NJ, USA, 1978.
9. Bakar, N.A.; Abu-Siada, A. A New Method to Detect Dissolved Gases in Transformer Oil using NIR-IR Spectroscopy. *IEEE Trans. Dielectr. Electr. Insul.* **2017**, *24*, 409–419. [CrossRef]
10. Muhamad, N.A.; Phung, B.T.; Blackburn, T.R.; Lai, K.X. Comparative Study and Analysis of DGA Methods for Transformer Mineral Oil. In Proceedings of the 2007 IEEE in Power Tech, Lausanne, Switzerland, 1–5 July 2007; pp. 45–50.
11. Rogers, R.R. IEEE and IEC Codes to Interpret Incipient Faults in Transformers, Using Gas in Oil Analysis. *IEEE Trans. Electr. Insul.* **1978**, *EI-13*, 349–354. [CrossRef]
12. De Pablo, A.; Ferguson, W.; Mudryk, A.; Golovan, D. On-line condition monitoring of power transformers: A case history. In Proceedings of the Electrical Insulation Conference (EIC), Annapolis, MD, USA, 5–8 June 2011; pp. 285–288.
13. Duval, M. A review of faults detectable by gas-in-oil analysis in transformers. *IEEE Electr. Insul. Mag.* **2002**, *18*, 8–17. [CrossRef]
14. Abu-Siada, A.; Hmood, S.; Islam, S. A New Fuzzy Logic Approach for Consistent Interpretation of Dissolved Gas-in-Oil Analysis. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 2343–2349. [CrossRef]
15. Abu-Siada, A.; Hmood, S.; Islam, S. Fuzzy Logic Approach to Identify Transformer Criticality using Dissolved Gas Analysis. In Proceeding of the 2010 IEEE PES General Meeting, Minneapolis, MN, USA, 25–29 July 2010.
16. Bakar, N.A.; Abu-Siada, A. Fuzzy Logic Approach for Transformer Remnant Life Prediction and Asset Management Decision. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 3199–3208. [CrossRef]
17. Miranda, V.; Castro, A. Improving the IEC Table for Transformer Failure Diagnosis With Knowledge Extraction From Neural Networks. *IEEE Trans. Power Deliv.* **2005**, *20*, 2509–2516. [CrossRef]
18. Abu-Siada, A. Improved Consistent Interpretation Approach of Fault Type within Power Transformers Using Dissolved Gas Analysis and Gene Expression Programming. *Energies* **2019**, *12*, 730. [CrossRef]
19. Abu-Siada, A.; Lai, S.P.; Islam, S. Remnant Life Estimation of Power Transformer using Oil UV-Vis Spectral Response. Presented at the 2009 IEEE PES Power Systems Conference & Exhibition (PSCE), Seattle, WA, USA, 15–18 March 2009.
20. Bacha, K.; Souahlia, S.; Gossa, M. Power transformer fault diagnosis based on dissolved gas analysis by support vector machine. *Electr. Power Syst. Res.* **2012**, *83*, 73–79. [CrossRef]
21. Taha, I.B.M.; Hoballah, A.; Ghoneim, S.S.M. Optimal ratio limits of rogers' four-ratios and IEC 60599 code methods using particle swarm optimization fuzzy-logic approach. *IEEE Trans. Dielectr. Electr. Insul.* **2020**, *27*, 222–230. [CrossRef]
22. Abu-Siada, A.; Hmood, S. Fuzzy Logic Approach for Power Transformer Asset Management Based on Dissolved Gas-in-Oil Analysis. In Proceedings of the Prognostics and System Health Management Conference, Gaithersburg, ML, USA, 8–11 September 2013. Available online: https://www.aidic.it/cet/13/33/167.pdf (accessed on 1 May 2021).
23. Lee, S.I.; Lee, H.; Abbeel, P.; Ng, A.Y. Efficient L_1 regularized logistic regression. *Am. Assoc. Artif. Intell.* **2006**, *6*, 401–408.
24. Ng, A.Y. Feature selection, L1 vs. L2 regularization, and rotational invariance. In Proceedings of the twenty-first international conference on Machine learning, Banff, AB, Canada, 4–8 July 2004.
25. Goodman, J. Exponential priors for maximum entropy models. In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: HLT-NAACL 2004, Boston, MA, USA, 2–7 May 2004; pp. 305–312.
26. Perkins, S.; Theiler, J. Online feature selection using grafting. In Proceedings of the 20th International Conference on Machine Learning (ICML-03), Washington, DC, USA, 21–24 August 2003; pp. 592–599.
27. Roth, V. The Generalized LASSO. *IEEE Trans. Neural Netw.* **2004**, *15*, 16–28. [CrossRef] [PubMed]
28. Shadabi, F.; Sharma, D. Comparison of Artificial Neural Networks with Logistic Regression in Prediction of Kidney Transplant Outcomes. In Proceedings of the 2009 International Conference on Future Computer and Communication, Kuala Lumpar, Malaysia, 3–5 April 2009; pp. 543–547.
29. Plan, Y.; Vershynin, R. The Generalized Lasso with Non-Linear Observations. *IEEE Trans. Inf. Theory* **2016**, *62*, 1528–1537. [CrossRef]
30. Ghoneim, S.S.; Taha, I.B. A new approach of DGA interpretation technique for transformer fault diagnosis. *Int. J. Electr. Power Energy Syst.* **2016**, *81*, 265–274. [CrossRef]
31. Taha, I.B.; Mansour DE, A.; Ghoneim, S.S.; Elkalashy, N. Conditional probability-based interpretation of dissolved gas analysis for transformer incipient faults. *IET Gener. Transm. Distrib.* **2017**, *11*, 943–951. [CrossRef]
32. Hoballah, A.; Mansour, D.-E.A.; Taha, I.B.M. Hybrid Grey Wolf Optimizer for Transformer Fault Diagnosis Using Dissolved Gases Considering Uncertainty in Measurements. *IEEE Access* **2020**, *8*, 139176–139187. [CrossRef]