International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2021 | Issue 49 | Volume 13 | Number 4 | Pages 225-230

FAULT DIAGNOSIS OF POWER TRANSFORMERS WITH MACHINE LEARNING METHODS USING TRADITIONAL METHODS DATA

M. Demirci H. Gozde M.C. Taplamacioglu

Electrical and Electronics Engineering Department, Gazi University, Ankara, Turkey
mervedemirci@gazi.edu.tr, halukgozde@gmail.com, taplam@gazi.edu.tr

Abstract- Since power transformers are one of the most important parts in the electrical power system, fault diagnosis in transformers is one of the most crucial issues. The Dissolved Gas Analysis method is often used to obtain gas concentrations to be used in fault diagnosis in transformers. Traditional methods have been used for many years to diagnose supply using these gas data. However, since these traditional methods cannot diagnose faults in some cases, smart methods such as machine learning methods have started to be used. This study presents a 523 set of data taken from a real power grid and troubleshooting studies done using traditional and intelligent methods. In the present research, the gas ratio and gas percentages used in traditional methods are used with machine learning methods. The obtained results are compared to each other. As a result of the study, it is observed that the accuracy of diagnosis increased by applying these new methods. It has been observed with the results that the content of the input data used for the classification algorithms affects the diagnostic performance.

Keywords: Transformer, Diagnosis, Fault Analysis, Machine Learning, Traditional Methods.

1. INTRODUCTION

Power transformers are one of the most vital equipment located at every stage of power systems. The trouble-free operation of the power system depends on the health of these equipment, the early diagnosis of any malfunction, and most importantly, the elimination of this malfunction before it escalates. Since power transformers are high-cost equipment, any malfunctions that occur can cause costly consequences. Faults that are not diagnosed early can cause major power outages [1, 2, 3]. For this reason, transformer faults have become one of the important research areas, and interest in diagnosing these faults has increased.

There is a Dissolved Gas Analysis (DGA) method used in transformer fault diagnosis, which makes measurements based on the changes occurring in the insulation oil in case of any malfunction. In the DGA method, gas concentrations in the insulating oil are measured and fault diagnosis is made using different interpretation methods [4]. The methods included in the standards and used for many years, where the rules used in fault diagnosis are created according to past fault data and expert experience, are called traditional methods. These methods diagnose malfunctions using different gas rates and gas percentages, rules and graphical representations [4]. In fault diagnosis made in this way, in some cases a wrong diagnosis is made, and in some cases no diagnosis can be made. To overcome these situations, intelligent classification methods have begun to be used. These methods are computer-aided methods and based on artificial intelligence.

In this study, fault diagnosis has been made using gas concentrations obtained from the DGA method for power systems fault diagnosis. It is aimed to increase the fault diagnosis performance by using a combination of traditional methods and intelligent methods, which are DGA interpretation methods. For this purpose, gas ratios and gas percentages used in traditional methods have been used as input data for classification algorithms. Gas ratios and gas percentages used in Rogers Ratio Method, Doernenburg Ratio Method, Duval Triangle Method and Duval Pentagon Method constituted the input data of separate classifier algorithms. Machine learning methods Support Vector Machine (SVM), K Nearest Neighbors (KNN), Naive Bayes (NB) and Decision Tree (DT) have been used as classification algorithms [5, 6, 7, 8]. Firstly, the fault diagnosis performances using gas data and the input data created by traditional methods were obtained and the results have been compared.

This study is organized as follows. In the second part, information about the DGA method and interpretation method is given. In the third section, information about smart methods and application steps is given, and the intelligent methods used in this study are briefly explained. In the fourth section, the data set content and fault diagnosis results are included. The fifth chapter includes the findings obtained in this study.

2. DISSOLVED GAS ANALYSIS (DGA) and INTERPRETATION METHODS

There are insulating oils used in transformers for insulation and cooling purposes. In case of any malfunction in the transformers, differences begin to occur in the structure of this oil. What differentiates them is the gas concentrations that occur in the oil used and increase depending on the type of fault. The temperature and energy at which the gases formed in the insulating oil are different. These differences are used to determine the transformer fault type [4, 9].

The method used to determine the concentrations of gases formed in the insulating oil is Dissolved Gas Analysis (DGA). With this method, the gases that increase in the insulating oil before or at the time of failure are measured by the chromatography method used to separate chemical components [4, 10, 11].

The aim of the DGA method is to diagnose the current fault and prevent potentially costly problems. For this reason, gas concentrations formed in the oil are used. The most frequently used gases formed in transformer oil and measured in the DGA method are: CH_4- Methane, H_2-Hydrogen, C_2H_4- Ethylene and C_2H_6-Ethane. Each of these gas concentrations is associated with a malfunction, and when that gas is in excess, there can be a prediction about the type of malfunction. For example, an increase in hydrogen and acetylene may be related to arc failures, while an increase in ethylene may be related to thermal failure. Only the increase in hydrogen may be related to partial discharge. These examples are given in the Key Gas Method, which is also included in the standards created through experience [4].

These measured gas concentrations are used with different interpretation methods in the standards and the transformer status is determined. These methods are: Rogers and Doernenburg Ratio Method, Duval Triangle and Pentagon Method. Below, the gas ratios and gas percentages used in these methods from gas concentrations are given below [4, 12].

- Rogers Ratio Method:
C_2H_2/C_2H_4, CH_4/H_2, C_2H_4/C_2H_6
- Doernenburg Method:
CH_4/H_2, C_2H_2/C_2H_4, C_2H_2/CH_4, C_2H_6/C_2H_2
- Duval Triangle Method:
$\%CH_4$, $\%C_2H_4$, $\%C_2H_2$
- Duval Pentagon Method:
$\%H_2$, $\%CH_4$, $\%C_2H_2$, $\%C_2H_4$, $\%C_2H_6$

3. INTELLIGENT CLASSIFICATION METHODS

Intelligent classification algorithms consist of computer and artificial intelligence-based methods. The use of these methods in transformer fault diagnosis has become widespread because the desired efficiency cannot be achieved with traditional methods. Intelligent classification algorithms: Machine learning methods, fuzzy logic methods, expert systems and artificial neural network methods [13, 14, 15].

There are process steps followed when classifying with machine learning methods [16].

- First, the data set to be used in the classification process is defined.
- Next, the data set is prepared for the algorithm. If necessary, it is passed through a pre-processing step.
- The classification algorithm to be used is selected.
- The selected algorithm is trained with the training data set.
- The trained data set is finally tested with the test data set and the results are evaluated.

The classification algorithms used in this study are machine learning algorithms SVM, NB, DT and KNN.

3.1. Support Vector Machine (SVM)

It is one of the machine learning classification algorithms that was introduced in 1995 as a statistical learning theory. In the plane where the data in the used data set is placed, a margin is determined between the data and this margin is aimed to be maximum. The fact that the algorithm is easy to optimize and responds quickly makes it the most preferred machine learning method [7, 18, 19].

3.2. Decision Tree (DT)

It is one of the preferred machine learning algorithms because the rule bases created when classifying are easy to understand and offer the user fast and high accuracy classification. Structurally, it is similar to the real-life tree structure. It consists of its components: roots, branches and leaves. It presents the rules used in classification to the user through these components [20, 21, 22].

3.3. K-Nearest Neighbors (KNN)

It is a machine learning algorithm based on the principle of determining the class of unlabeled data, depending on the distance to the labeled data in the data set. It is one of the preferred machine learning algorithms because its working principle is easy to understand and easy to apply [23, 24].

The important components of this algorithm are the number of neighborhoods and the distance measurement method. By changing these components, the most accurate result can be obtained from the classification algorithm [23, 24].

3.4. Naive Bayes (NB)

Naive Bayes algorithm is one of the statistical learning algorithms. It operates according to Bayes' theorem. It is used due to its performance and simplicity [8].

4. SIMULATION RESULTS

4.1. Dataset Content

The data set used in this study includes 523 DGA data. There are 5 gas concentrations obtained from the DGA method in the data set. In the data set, transformers are classified in 6 different states, faulty and normal, depending on these gas concentrations. Faulty data includes 5 fault types. These are: Low energy Discharge, High Energy Discharge, Thermal Fault (300°C<T<700°C-T_2), Thermal Fault (T>700°C-T_3), Thermal Fault (T_2/T_3).

4.2. Classification Results

In this study, a study has been conducted on the combination of machine learning classification algorithms and traditional methods when diagnosing faults in the transformer. The sight ratios and gas percentages used by traditional methods when diagnosing faults constitute the input data of the classification algorithms.

First, 5 gas concentrations have been used as input data of the classifiers, then 3 gas ratios used by the Rogers Ratio method and 4 gas ratios used by the Doernenburg Ratio method have been used, respectively. Afterwards, the classification has made with the percentages of the 3 gases used by the Duval Triangle method in fault diagnosis, and finally, the classification has made with the percentages of the 5 gases used in the Duval Pentagon method.

In the classification process, cross validation (cv) has been applied to show that the classification results are independent of the training and test data set used in the training and testing of the algorithm. The results have been classified as cv=10 and compared.

The fault diagnosis has examined in 5 different cases according to the data set used and the results have been obtained. The results for different situations are listed below.

4.2.1. Case-1

In this case, the 5 gas concentrations obtained from the DGA method have been used as input data for the classification algorithms. The results of the classification algorithms are obtained with cv = 10 and the results are given in Table 1.

Table 1. Fault Diagnosis Accuracies for Case 1

Classification Algorithm	Fault Diagnosis Accuracy (%)
SVM	73.4%
KNN	66.1%
NB	58%
DT	79.3%

When the results obtained have been compared, the highest fault diagnosis accuracy of 79.3% has been obtained in the DT algorithm. The lowest diagnostic accuracy belongs to the NB algorithm. Confusion Matrix of the DT algorithm is given in Figure 1. Confusion matrix is used to show how many transformer fault labels in the data set are classified correctly and how many are classified incorrectly.

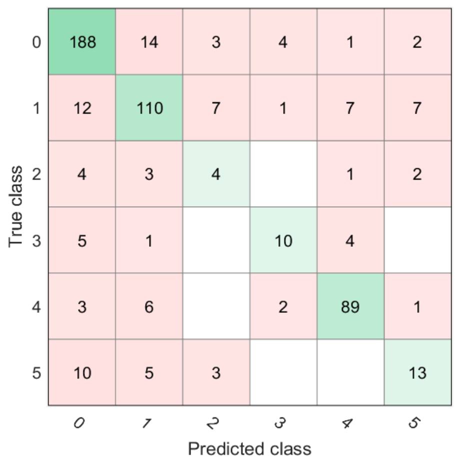


Figure 1. Confusion matrix of DT algorithm for Case 1

4.2.2. Case-2

In this case, 3 gas ratios (C_2H_2/C_2H_4, CH_4/H_2, C_2H_4/C_2H_6) used in fault diagnosis in the Rogers Ratio Method have been used as input data for the classification algorithms. The effect of these gas ratios on the diagnostic performance of the algorithms has been examined. The classification results obtained are given in Table 2.

Table 2. Fault Diagnosis Accuracies for Case 2

Classification Algorithm	Fault Diagnosis Accuracy (%)
SVM	57.5%
KNN	47%
NB	38%
DT	78.4%

In the case where Rogers ratios have been used, the highest diagnostic accuracy has been achieved with the DT algorithm and is 78.4%. The confusion matrix of the DT algorithm is shown in Figure 2. When the performances of other classification algorithms are examined, it is seen that the gas ratios used do not increase the performance of the classifiers, but rather reduce them.

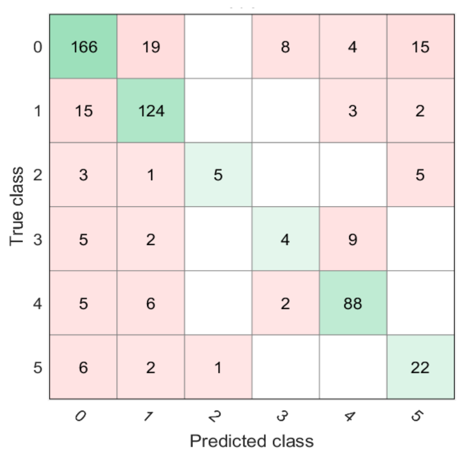


Figure 2. Confusion matrix of DT algorithm for Case 2

4.2.3. Case-3

Fault diagnosis has been made using the 4 gas ratios (CH_4/H_2, C_2H_2/C_2H_4, C_2H_2/CH_4, C_2H_6/C_2H_2) used in the Doernenburg Ratio Method as the data set. The diagnostic performances of machine learning classification algorithms are given in Table 3.

Table 3. Fault Diagnosis Accuracies for Case 3

Classification Algorithm	Fault Diagnosis Accuracy (%)
SVM	58.6%
KNN	49.8%
NB	35%
DT	84.7%

The highest diagnostic accuracy of 84.7% in fault diagnosis using Doernenburg ratios after Rogers Ratio ratios has been obtained with the DT algorithm and the confusion matrix is given in Figure 3.

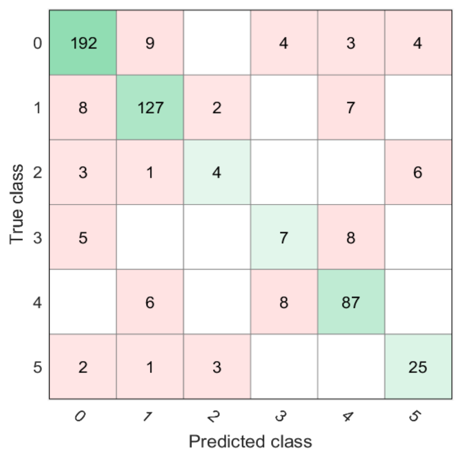


Figure 3. Confusion matrix of DT algorithm for Case 3

4.2.4. Case-4

In this case, 3 gas percentages (%CH_4, %C_2H_4, %C_2H_2), which are placed on the three sides of the triangle in the Duval Triangle Method and used to diagnose the fault, have used as input data for the algorithms. The obtained diagnostic results are shown in Table 4.

Table 4. Fault Diagnosis Accuracies for Case 4

Classification Algorithm	Fault Diagnosis Accuracy (%)
SVM	89.3%
KNN	90.6%
NB	80.8%
DT	90.2%

In fault diagnosis using three gas percentages, the highest diagnostic accuracy of 90.6% has been obtained with the KNN algorithm. The confusion matrix of KNN is given in Figure 4. When the performances of other algorithms have been examined, it has been seen that the use of Duval gas percentages increased the performance of the classifiers.

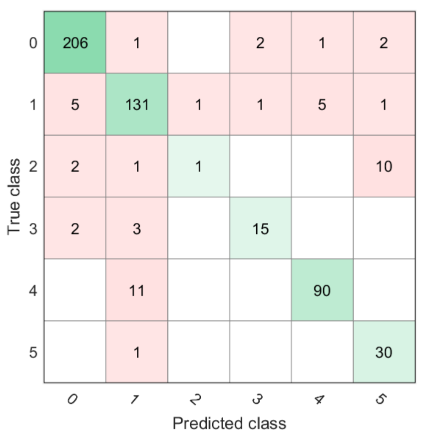


Figure 4. Confusion matrix of KNN algorithm for Case 4

4.2.5. Case-5

Fault diagnosis has been made using 5 gas percentages (%H_2, %CH_4, %C_2H_2, %C_2H_4, %C_2H_6), which are used in the Duval Pentagon method and placed on the edges of the pentagon. The diagnostic performances obtained by cross validation using this data set are given comparatively in Table 5.

Table 5. Fault Diagnosis Accuracies for Case 5

Classification Algorithm	Fault Diagnosis Accuracy (%)
SVM	87.4%
KNN	84.9%
NB	71.8%
DT	83.9%

The highest diagnostic accuracy obtained with the 5 gas percentage used in the Duval Pentagon method has been obtained in the SVM algorithm with 87.4%. The confusion matrix of SVM is given in Figure 5.

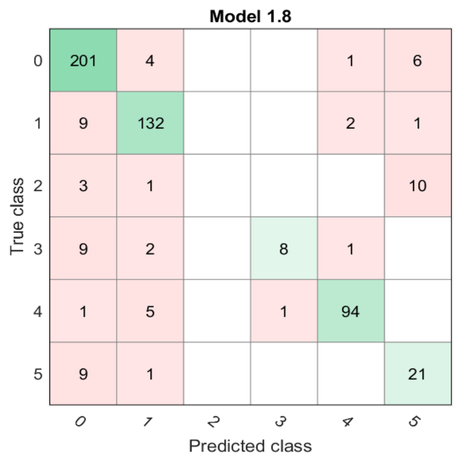


Figure 5. Confusion matrix of SVM algorithm for Case 5

In this study, where gas ratios and gas percentages of traditional methods have been compared as input data for classification algorithms, the results show that the highest diagnostic accuracy has been obtained with 3 gas percentages used in the Duval Triangle method. The highest diagnostic accuracies of the input data are given in Table 6.

Table 6. Highest Diagnosis Accuracies for Different Cases

Input Data	Used Algorithm	Fault Diagnosis Accuracy
Case-1: Raw 5 gas concenrations	DT	79.3%
Case-2: Rogers ratios	DT	78.4%
Case-3: Doernenburg ratios	DT	84.7%
Case-4: Duval Triangle percentages	KNN	90.6%
Case-5: Duval Pentagon percentages	SVM	87.4%

The graph showing the classification accuracies obtained according to the input data for all classifiers is given in Figure 6. As can be seen from this graph, the highest diagnostic accuracy is obtained with the gas percentages used in the Duval Triangle method in the KNN algorithm.

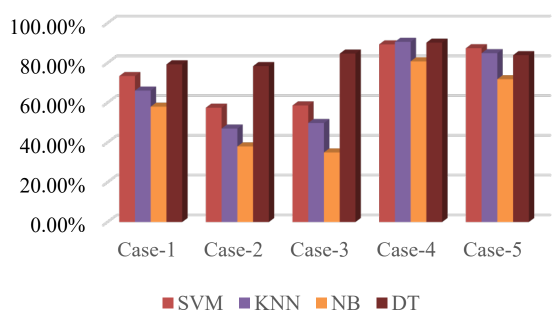


Figure 6. Diagnostic Accuracy Based on Input Data of Classification Algorithm

5. CONCLUSIONS

In this study, machine learning methods, which are intelligent methods for fault diagnosis in power transformers, and traditional methods included in the standards have been used together. The 5 gas concentrations obtained from the DGA method and the gas ratios and gas percentages used in traditional methods have been used as input data for the classification algorithms. The classification algorithms used are SVM, KNN, DT and NB. When the input data of these algorithms are different, the diagnostic performances have been obtained and the results have been compared. The aim is to increase the diagnostic performance of the algorithms by using different input data.

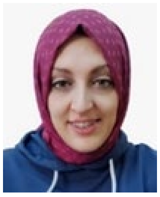
When the results obtained have been compared, it has observed that the highest diagnostic accuracies have been obtained with the 3 gas ratios used in the Duval Triangle Method. The use of different input data did not always affect the classification performance positively, and it has been observed that the classifier performances decreased in Case-2, where Rogers ratios have been used. This shows that the data set used does not carry the necessary information for the classification algorithms to make an accurate diagnosis. The results show that the content of the dataset used in training and testing classification algorithms has a great impact on performance.

REFERENCES

[1] L. Luan, W. Mo, L. Kong, K. Zhou, "A refined Classification Method for Transformer Fault", Journal of Physics, Conference Series, 1302, 2019.
[2] F. Yang, J. Xi, L. Zhida, "A Neural Network Approach to Power Transformer Fault Diagnosis", The Sixth International Conference on Electrical Machines and Systems, Vol. 1, pp. 351-354, Beijing, China, 2003.
[3] M. Zile, "Design of Power Transformers Using Heuristic Algorithms", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 38, Vol. 11, No. 1, pp. 42-47, March 2019.
[4] "IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers", IEEE Std C57.104TM- 2019. IEEE Power and Energy Society.
[5] H. Xie, J. Chen, P. Zhang, F. Zhang, X. Meng, J. Cheng, "Knowledge Acquisition for Transformer Condition Assessment Using Synthetic Minority Over-Sampling Technique and Decision Tree Algorithm", IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), pp 4186-4191, Chengdu, China, 2019.
[6] X. Wu, V. Kumar, J.R. Qinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. Mclachlan, A. Ng, B. Liu, P.S. Yu, Z.H. Zhou, M. Steinbach, D.J. Hand, D. Steinberg, "Top 10 Algorithms in Data Mining", Knowledge and Information Systems, Vol. 14, pp. 1-37, 2008.
[7] K. Bacha, S. Souahlia, M. Gossa, "Power Transformer Fault Diagnosis Based on Dissolved Gas Analysis by Support Vector Machine", Electric Power System Research, Vol. 83, pp. 73-79, 2012.
[8] G. Odongo, R. Musabe, D. Hanyurwimfura, "A Multinomial DGA Classifier for Incipient Fault in Oil-Impregnated Power Transformers", Algorithms, Vol. 14, No. 4, 2021.
[9] M. Sezer, "Transformer Maintenance and Insulation Oil", Cagdas Publishing House, Balikesir, Turkey, 2014.
[10] N. Lelekakis, D. Martin, W. Guo, J. Wijaya, "Comparison of Dissolved Gas-in-Oil Analysis Methods Using a Dissolved Gas-in-Oil Standard", IEEE Electrical Insulation Magazine, Vol. 27, No. 5, pp. 29-35, September-October 2011.
[11] N.A. Bakar, A. Abu-Siada, S. Islam, "A Review of Dissolved Gas Analysis Measurement and Interpretation Techniques", IEEE Electrical Insulation Magazine, Vol. 30, No. 3, pp. 39-49, May-June 2014.
[12] 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.
[13] C.E. Lin, J. Ling, C. Huang, "An Expert System for Transformer Fault Diagnosis Using Dissolved Gas Analysis", IEEE Transactions on Power Delivery, Vol. 8, No. 1, pp. 231-238, January 1993.
[14] F. Zakaria, D. Johari, I. Musirin, "Artificial Neural Network (ANN) Application in Dissolved Gas Analysis (DGA) Methods for the Detection of Incipient Faults in Oil-Filled Power Transformer", IEEE International Conference on Control System, Computing and Engineering, pp. 328-332, Penang, Malaysia, 23-25 November 2012.
[15] K. Tomsovic, M. Tapper, T. Ingvarsson, "A Fuzzy Information Approach to Integrating Different Transformer Diagnostic Methods", IEEE Transactions on Power Delivery, Vol. 8, No. 3, pp. 1638-1646, July 1993.
[16] H.I. Bulbul, O. Unsal, "Comparison of Classification Techniques used in Machine Learning as Applied on Vocational Guidance Data", The 10th International Conference on Machine Learning and Applications and Workshops, pp. 298-301, Honolulu, HI, USA, 18-21 Dec. 2011.
[17] A. Basuki, Suwarno, "Online Dissolved Gas Analysis of Power Transformers Based on Decision Tree Model", Conference on Power Engineering and Renewable Energy (ICPERE), pp. 1-6, Solo, Indonesia, 29-31 October 2018.
[18] V.N. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, USA, 1995.
[19] T. Kari, W. Gao, D. Zhao, K. Abiderexiti, W. Mo, Y. Wang, L. Luan, "Hybrid Feature Selection Approach for Power Transformer Fault Diagnosis Based on Support Vector Machine and Genetic Algorithm", IET Generation

Transmission and Distribution, Vol. 12, pp. 5672-5682, 2018.
[20] F. Zhao, H. Su, "A Decision Tree Approach for Power Transformer Insulation Fault Diagnosis", The 7th World Congress on Intelligent Control and Automation, pp. 6882-6886, Chongqing, China, 25-27 June 2008.
[21] M.G. Broulday, E.G. Yegin, P. Mahouti, F. Gunes, "Diagnosing liver Diseases with Decision Tree Algorithm", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 33, Vol. 9, No. 4, pp. 67-70, December 2017.
[22] M.E.A. Senoussaoui, M. Brahami, I. Fofana, "Combining and comparing Various Machine Learning Algorithms to Improve Dissolved Gas Analysis Interpretation", IET Generation, Transmission and Distribution, Vol. 12, pp. 3673-3679, 2020.
[23] Y. Benmahamed, M. Teguar, A. Boubakeur, "Diagnosis of Power Transformer Oil Using PSO-SVM and KNN Classifiers", International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), pp. 1-4, Algiers, Algeria, 28-31 October 2018.
[24] M. Demirci, H. Gozde and M. C. Taplamacioglu, "Comparative Dissolved Gas Analysis with Machine Learning and Traditional Methods", The 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-6, Ankara, Turkey, 11-13 June 2021.

BIOGRAPHIES



Merve Demirci graduated from Department of Electrical and Electronics Engineering, Ataturk University, Erzurum, Turkey in 2014. She received M.Sc. degree in 2018 and currently is a Ph.D. student in Electrical and Electronics Engineering Department, Gazi University. She is currently working as a Research Assistant at Gazi University, Ankara, Turkey since 2016. Her main research interests and subjects are power systems analysis, artificial intelligence.

Haluk Gozde received the B.Sc. degree in Electrical and Electronics Engineering from Karadeniz Technical University, Trabzon, Turkey in 1997. He received the M.Sc. and the Ph.D. degrees in Electrical and Electronics Engineering from Gazi University, Ankara, Turkey in 2004 and 2010, respectively. He is working as Assoc. Prof. of the Electrical and Electronics Engineering since 2016. His main research area consists of power system dynamics and control, renewable energy systems, smart grid technologies, artificial intelligence-based control methods, and swarm intelligence-based optimization algorithms.

M. Cengiz Taplamacioglu graduated from Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey. He received the M.Sc. degrees in Industrial Engineering from Gazi University and also in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey. He received his Ph.D. degree in Electrical, Electronics and System Engineering from University of Wales, Cardiff, UK. He is a Professor of the Electrical and Electronics Engineering since 2000. His research interests and subjects are high voltage engineering, corona discharge and modelling, electrical field computation, measurement and modelling techniques, optical HV measurement techniques, power systems control and protection, lighting techniques, renewable energy systems and smart grid applications.