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 circular 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 fraction 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.

Keywords: Power Transformer, Intelligent Methods, Traditional Methods, Fault Diagnosis.

1. INTRODUCTION

Due to the increasing population of the world especially in recent years, the need for electricity increases rapidly. In parallel, the capacities of electrical power systems have to be largely increased. Because of security and reliability are great importance of electric power systems, any failure that may occur into the system and to fix this fault in a short time also gains great importance. Power transformers have an important place in every stage of power systems. Therefore, fault diagnosis in a power transformer is important to ensure continuity, which is one of the most important issues of electrical power systems [1, 2, 3].

Although there are many techniques to be used in fault diagnosis in electrical power transformers, the most widely used method today is the Dissolved Gas Analysis (DGA) method. By using this method, the gas concentrations in the transformer insulating liquid are measured and evaluated them these data uses for fault diagnosis [4].

On the other hand, to diagnose faults by using DGA data, traditional methods based on personal experiences have been also using for many years. In these methods, fault diagnosis is made by using the ratios and different percentages of gas concentrations obtained from the DGA method [4]. However, since the rule bases used for diagnosis in these methods are generalized data created according to past data and experiences, many deficiencies are encountered. For this reason, recently the computer-based methods called as intelligent methods have started to be used.

The aim of this study is to achieve highest diagnostic accuracy by using the combination of traditional and intelligent methods for fault diagnosis of power transformers using the same data. For this purpose, first of all, the fault diagnosis is made using K- Nearest Neighbors (k-NN), Decision Tree and Support Vector Machine (SVM) methods based on machine learning, by using the data obtained from the DGA method [5, 6, 7]. Following, the gas ratios and gas percentages used in traditional methods are created in a way to form the input data of the machine learning methods, and successful results of the fault diagnosis are obtained and comparisons are made.

Brief information about the DGA method, with the traditional and intelligent methods is summarized in Part 2. Dataset information is given in Part 3. The simulation results are presented and compared in Part 4. Finally, in Part 5, the conclusions and interpretations obtained from the simulations are given in detail.

2. DISSOLVED GAS ANALYSIS (DGA)

In power transformers, different gas molecules start to form when or before any malfunction occurs during operation [4, 8]. The formation temperatures and energies of these gas molecules are different and these are given in Table 1. These gas concentrations can be used in fault diagnosis as different amounts and different types of gases are produced in different faults.

Table 1. The formation temperature and energies of the gases [8]

	Formation Temperature (°C)	Formation Energy (kj/mol)
H_2	<500	>338
CH_4	<500	>607
C_2H_6	<500	>607
C_2H_4	>500	>720
C_2H_2	>800-1200	>960
CO	>105-300	-
CO_2	>105-300	-

The DGA method allows measuring these gas concentrations. The gases measured are nitrogen, ethylene, carbon- monoxide, carbon dioxide, methane, ethane, hydrogen, acetylene and oxygen. The Gas Chromatography method is used to determine the gas type and concentration in the transformer insulating liquid [4]. This method provides a simple analysis to distinguish chemical components in a sample [9, 10]. Figure 1 shows the basic gas chromatograph structure used for Dissolved gas analysis.

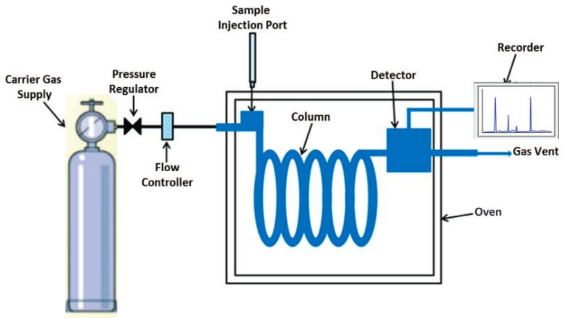

Figure 1. Basic gas chromatograph [10]

The main purpose of DGA method is to detect a potential or current failure and to increase the system security while reducing the cost. It increases the system security by detecting the fault in a transformer in advance and monitoring a transformer suspicious for faults. In addition, it reduces the maintenance costs by preventing transformers from being damaged by failure. The other duties of DGA method are basic risk management, detection of existing anomalies in the system, monitoring of these anomalies, quality assurance measurement and fault type identification [4].

The gas concentrations attained from DGA analysis can be interpreted using various methods, and then the present transformer condition can be commented on. The DGA interpretation methods are divided in two: the traditional methods and the intelligent methods. The faults are commonly classified depending on gas concentrations as partial discharge, electrical faults, and thermal faults. In general, the increase in some gas concentrations indicates a certain malfunction. For example; although the key gas method, which is one of the traditional methods, is not precise, they make predictions for some failures by investigating at gas concentrations [4]. Diagnostic rules of the key gas method are given in Table 2.

Table 2. Diagnostic rules of the Key Gas method [4]

Key Gas	Transformer Fault Type
Ethylene	Thermal Mineral Oil
Carbon-Monoxide	Thermal Mineral Oil and Cellulose
Hydrogen	Electrical Low Energy Partial Discharge (PD)
Hydrogen and Acetylene	Electrical High Energy (arching)

2.1. Traditional Methods

Traditional methods classify the transformer state according to some rules using the gas concentrations obtained from DGA method. The rules used consist of graphs and tables created according to the past failure data and personal experiences. These methods are Duval Triangle Method, Rogers Ratio Method and Doernenburg Ratio Method. [4].

2.1.1. Rogers Ratio Method

In this method, three different gas ratios are used and the transformer situation is classified in 6 different ways. The gas ratios used are, CH_4/H_2, C_2H_4/C_2H_6, C_2H_2/C_2H_4.

2.1.2. Doernenburg Ratio Method

In this method, four different gas ratios are used and the transformer situation is classified in 3 different ways. The gas ratios used are CH_4/H_2, C_2H_2/CH_4, C_2H_6/C_2H_2, C_2H_2/C_2H_4.

2.1.3. Duval Triangle Method

In this method, a triangular rule shape is created with three different gas percentages to describe different transformer states. Transformer state is classified in 7 different states. Gas percentages used are $\%CH_4$, $\%C_2H_4$, $\%C_2H_2$.

In traditional methods, while fault classification is made according to the rules created with the experience, some situations do not comply with any rules. This reveals the deficiencies of the established rule bases. For this reason, recently intelligent classification methods are frequently used.

2.2. Intelligent Methods

Intelligent classification methods consist of computer-assisted artificial intelligence-based algorithms. It is widely used in the literature due to the shortcomings of the traditional method. In addition, in order to increase the diagnostic accuracy of the classification methods used, error classification can be made by using gas ratios and percentages used in traditional methods instead of raw DGA data.

Some of the intelligent methods can be summarized as Expert Systems [11], Artificial Neural Networks (ANN) [12], Fuzzy Logic [13] and Machine Learning methods [14].

In this study, machine learning methods are used for fault classification. The processing steps applied for classification in machine learning algorithms are given in Figure 2.

First, the data to be used in the classification algorithm are defined. Second, the data is prepared for the classification algorithm. In the third step, the machine learning algorithm to be used in the classification process is selected. Fourth, the algorithm is trained with the data and the trained model is created. Fifth, the model is evaluated. Finally, the trained model is put into practice. The model is expected to predict using the test data, and the process is completed by checking the estimation results obtained [15].

The intelligent methods based on machine learning, used in this study are K-NN, Decision Tree and SVM algorithms.

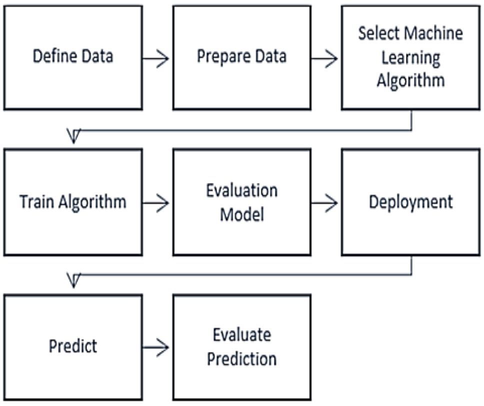

Figure 2. The process steps for machine learning [13]

2.2.1. Support Vector Machine Method (SVM)

Support vector machine method are a machine learning method based on statistical learning introduced by Cortes and Vapnik in 1995 [16].

SVM first maximizes the margin between the data in the dataset. It aims to find the optimal hyperplane with good generalization ability [17]. It is the most used machine learning method.

2.2.2. K Nearest Neighbors Method (k-NN)

The K nearest neighbors algorithm is a method that classifies the data based on the closest data available in the data set. It is preferred since it is easy to understand and apply [18].

In this method, first of all a distance measurement method is determined. With this method, the distance between the unlabeled data and the labeled data is calculated, then the number of k neighbors is selected and the nearest neighbors are determined accordingly. The majority label is chosen by looking at the label of the determined neighbors. Finally, the majority label is assigned as the label of the unlabeled data [5, 19].

2.2.3. Decision Tree Method

Decision tree algorithm is one of the preferred classification algorithms due to its simple structure in terms of computational complexity and high classification accuracy [20, 21].

The rule base for the classification process is similar to a tree structure. It consists of branches and nodes. The classification rule is obtained by following the structure from the root node to the leaf node [22]. In the Decision tree algorithm, the user can easily obtain the decision mechanism and classification rules from the tree structure created. The decision-making mechanism is very useful for the users. When the decision tree algorithm is used for transformer fault diagnosis, maintenance personnel can obtain the rule base to predict and interpret the condition of the transformer.

3. OBTAINING THE DATASET

Five of the gas concentrations acquired from DGA analysis are used. These are Methane, Ethylene, Ethane, Acetylene and Ethylene.

In the present work, the data consists of 523 sets of data taken from the real electric power system. The dataset classifies the transformer in 6 different states. These data are divided into training dataset and test dataset.

In the present study, first of all, the fault classification is made with machine learning algorithms. In this case, the raw DGA data is used as input data. Following, the gas ratios and gas percentages used in fault classification in traditional methods are used as input data of the machine learning algorithms and the classification is made again. The classification result is evaluated by using two measurements: Training Success Rate and Test Success Rate.

- The training success rate is the measure of how much this dataset is learned as a result of training the classification algorithm with the training dataset.
- The test success rate is a measure of how well the trained classification algorithm classifies the test data.

4. SIMULATION RESULTS

The simulations are performed in four cases:

- *Case-1:* In this case, the input data of the classification algorithm consists of 5 gas concentrations obtained from the DGA method. First, the training of the classifiers is completed with the training dataset and the training success rates are obtained. Then, the test data set is tested in the trained model and the test success rate is obtained. Success rates are given in the Table 3.

Table 3. Success Rates for Case 1

	Training Success Rate (%)	Test Success Rate (%)
Decision Tree	88.12	74.42
k-NN	99.62	70.61
SVM	75.47	70.40

When fault diagnosis is made using raw 5 DGA gases, the highest diagnostic accuracy is obtained in the Decision Tree Algorithm with 74.42%. The rule base created by the Decision Tree Algorithm for fault diagnosis is given in Figure 3 as a tree structure. Confusion matrix of the decision tree is given in Figure 4. With the confusion matrix structure, it can be observed how many of the fault types the classification algorithm diagnoses correct and how many of them are incorrect.

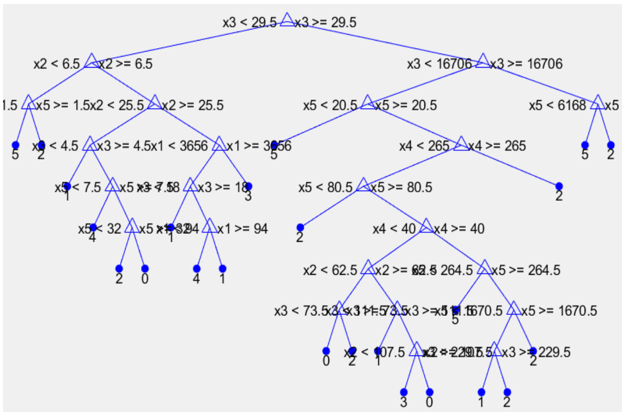

Figure 3. Tree Structure of Case 1

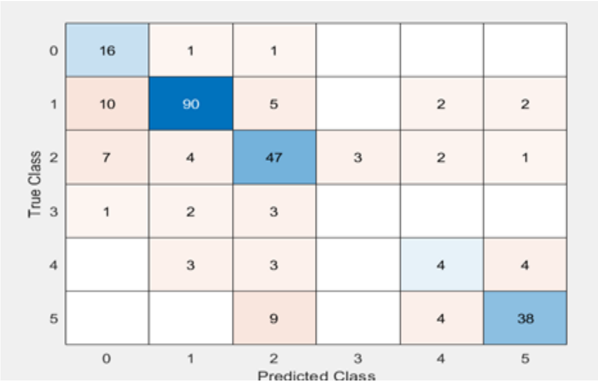

Figure 4. Decision Tree's Confusion matrix in Case 1

- *Case-2:* In this case, 3 gas ratios used in Roger's gas ratios method are used as inputs for machine learning based classification algorithms and their effect on the fault diagnosis result is observed. Obtained classification achievements are given in Table 4.

Table 4. Success Rates for Case 2

	Training Success Rate (%)	Test Success Rate (%)
Decision Tree	86.60	75.57
k-NN	87.36	64.89
SVM	49.80	43.51

In this case, the classification results are in the Decision Tree algorithm with the highest diagnostic accuracy. While the gas ratios used as the data set increased both the test success rate in the Decision tree algorithm, the success rates for other classification algorithms decreased. In this case, the classification results are in the Decision Tree algorithm has the highest diagnostic accuracy. While the gas ratios used as the data set increased the test success rate in the Decision tree algorithm, the success rates for other classification algorithms are decreased. The Decision Tree's Confusion matrix is given in Figure 5.

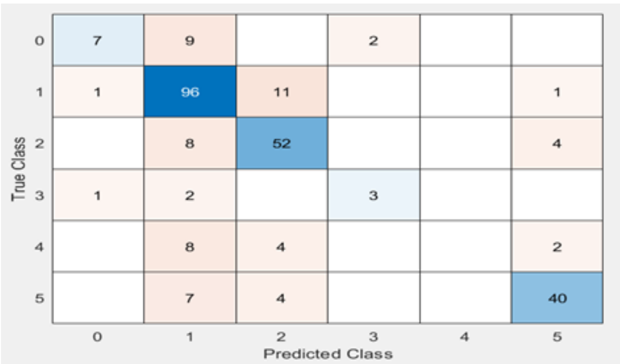

Figure 5. Decision Tree's Confusion matrix in Case 2

- *Case-3:* In this case, the 4 gas ratios used in the fault classification in the Doernenburg gas ratio method constitute the input data for the machine learning based classification algorithms. The success rates obtained according to these data are given in Table 5.

Table 5. Success Rates for Case 3

	Training Success Rate (%)	Test Success Rate (%)
Decision Tree	88.89	82.06
k-NN	93.11	77.09
SVM	60.15	58.78

In this study, the results obtained in all classifiers increased. While the highest classification accuracy is obtained with Decision Tree, the lowest classification success is obtained with SVM. Decision Tree's Confusion matrix is given in Figure 6.

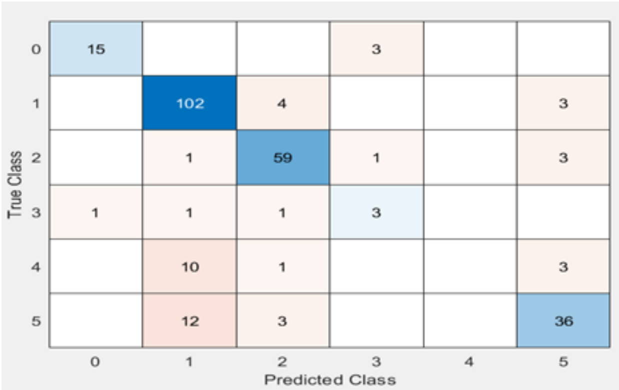

Figure 6. Decision Tree's Confusion matrix in Case 3

- *Case-4:* In this case, the gas percentages used in the Duval Triangle Method are used as the input data of the machine learning based classification algorithms, unlike the gas ratios used in other traditional methods. Obtained training success rates and test success rates are given in the Table 6.

Table 6. Success Rates for Case 4

	Training Success Rate (%)	Test Success Rate (%)
Decision Tree	95.40	93.13
k-NN	99.62	89.31
SVM	89.65	88.55

Looking at the results obtained, the gas percentages used in the Duval Triangle Method significantly increased the diagnostic accuracy. While the highest accuracy is obtained in the Decision tree algorithm, the lowest accuracy is obtained in the SVM method. The rule base of the Decision Tree algorithm is given in Figure 7 in the form of a tree structure. Confusion Matrix of Decision Tree is given in Figure 8.

When fault diagnosis is made using gas ratios and percentages used in traditional methods, the highest accuracy is obtained in the Decision Tree Algorithm in all cases. When the results of the Decision tree algorithm are given together, the highest diagnostic accuracy is obtained when the gas percentages used in the Duval Triangle Method is taken as input. In Table 7, these results are given comparatively.

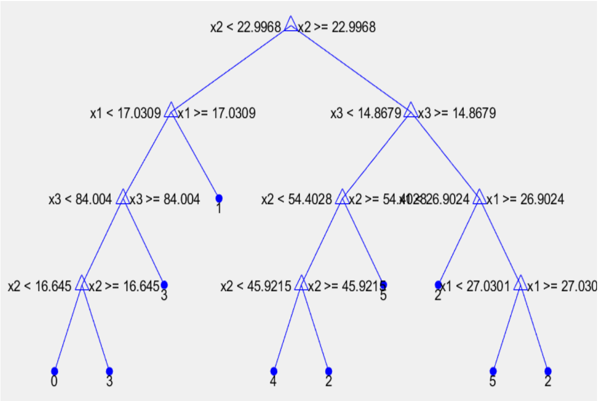

Figure 7. Tree Structure of Case 4

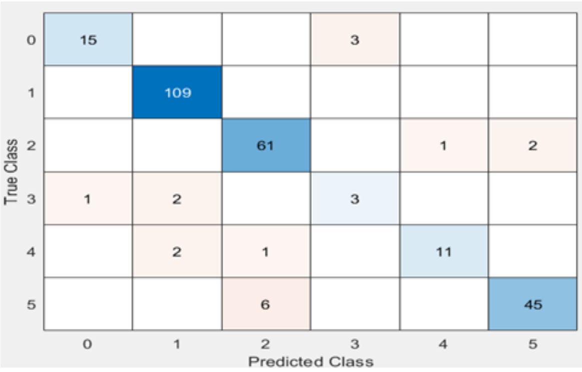

Figure 8. Decision Tree's Confusion matrix in Case 4

Table 7. Diagnosis Accuracy of Decision tree

	Diagnosis Accuracy
5 gas Concentrations	74.42
Roger's Ratios	75.52
Doernenburg's Ratios	82.06
Duval's Percentages	93.13

The graph showing the classification accuracies obtained according to the input data for all classifiers is given in Figure 9. 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 Decision Tree algorithm.

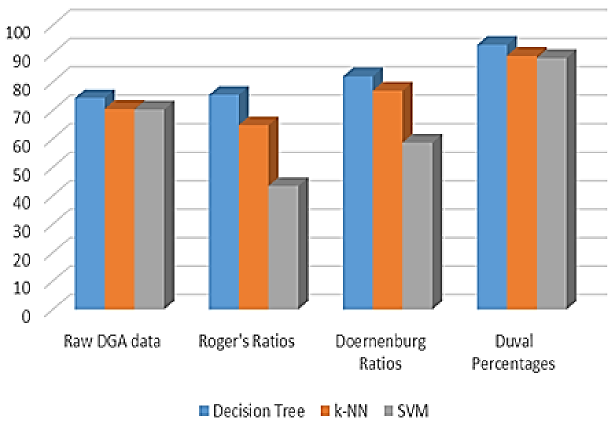

Figure 9. Diagnostic Accuracy Based on Input Data of Classification Algorithm

5. CONCLUSIONS

In this study, a number of analyses have been carried out for fault diagnosis in power transformers. It is aimed to increase the diagnostic accuracy by using machine learning methods as well as the traditional methods used in fault diagnosis for many years. Gas ratios and percentages of traditional methods are used as input data in machine learning methods. Diagnostic results for different input data are compared.

The results show that the Decision Tree algorithm has the highest accuracy diagnostic capability in all cases. When the diagnostic accuracy of the Decision Tree algorithm is examined, it is seen that the highest diagnostic ability is obtained with the gas percentages in the Duval triangle method. As a result of this study, it has been observed that the classification algorithm and the input data greatly affect the diagnostic accuracy.

REFERENCES

[1] L. Luan, W. Mo, L. Kong, K. Zhou, "A refined classification method for transformer fault", Journal of Physics: Conference Series, Vol. 1302, 2019.

[2] F. Yang, J. Xi, L. Zhida, "A neural network approach to power transformer fault diagnosis", 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), Vol. 11, No. 1, pp. 42-49, 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, 2019.

[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", 2019 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] M. Sezer, "Transformer Maintenance and Insulation Oil", Cagdas Publishing House, Balikesir, Turkey, 2014.

[9] 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.

[10] 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.

[11] 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, Jan. 1993.
[12] 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.
[13] 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.
[14] H.I. Bulbul, O. Unsal, "Comparison of Classification Techniques used in Machine Learning as Applied on Vocational Guidance Data", 10th International Conference on Machine Learning and Applications and Workshops, pp. 298-301, Honolulu, HI, USA, 18-21 Dec. 2011.
[15] 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.
[16] V.N. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, USA, 1995.
[17] 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 & Distribution, Vol. 12, pp. 5672-5682, 2018.
[18] Y. Benmahamed, M. Teguar, A. Boubakeur, "Diagnosis of Power Transformer Oil Using PSO-SVM and KNN Classifiers," 2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), pp. 1-4, Algiers, Algeria, 28-31 Oct. 2018.
[19] 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.
[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 & Distribution, Vol. 12, pp. 3673-3679, 2020.

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.