ISSN 2286-4822
www.euacademic.org

EUROPEAN ACADEMIC RESEARCH
Vol. VI, Issue 1/ April 2018

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

Power transformer fault diagnosis based on noise detection using clustering and classifier ensemble

MEHDI SHOKRI ASRAMI
Department of Information Technology
Faculty of Computer Science and Multimedia
Lincoln University College, Malaysia, Branch Iran
EBRAHIM AKBARI
Department of Computer Engineering, Sari Branch
Islamic Azad University, Sari, Iran

Abstract:

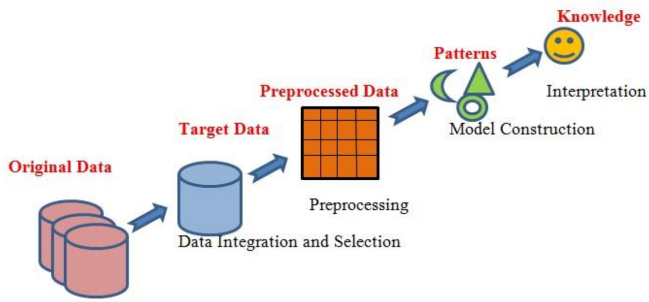
Today, due to existing complexity and changing organizations using accurate and modern tools is compulsory needed on performing the maintenance strategies. Human's improvements during last decades about gathering and storage the results and data cause the organizations have a huge dimension of data related to maintenance. Given the above description and rapid progress in the field of data mining and logging industry equipment repairs and maintenance on our plants that in this research we are going to power transformer fault diagnosis based on noise detection using clustering and classifier ensemble by oil transformers dissolved gas analysis (DGA) results and reduce costs and increase system reliability in the power generation industry through timely diagnosis and prognosis Defects in transformers which are considered one of the most critical power plant equipment in the power industry take an important and effective step. Dissolved gas analysis (DGA) of transformer oil is an important way for early fault detection. If the data the correct results of the gas chromatographic tests of the transformers are extracted and using the correct CE methods, we can detect the noise in the Data sets and remove them and perform the new Classification Data sets. We will get the best and highest accuracy in the output and our purposed method is called ensemble classification method with majority voting have more reliability than basic classification methods.

Key words: Power transformer, noise detection, techniques

INTRUDOCTION

Data mining and knowledge discovery in databases including subjects that coincided with the creation and use of databases were instigated in the early 1980s in the search for knowledge in the data. Lowell (1983) can be considered as the first person to provide a report on data mining report entitled "simulating the data mining activity". At the same time, researchers and specialists in the field of computer science, statistics, artificial intelligence, and machine learning have conducted researches in this field and related fields. Considering the importance of maintenance costs and control in an organization, it can be realized that having the data knowledge is extremely important to ensure the application of new data mining methods. Currently, although many organizations are aware of their information needs and use a large amount of budgets to use information systems, applying such new, accurate and analytical location-based data mining seems to be a crucial step and serves as the main objective of this research.

The research on the subject of data mining began seriously in early 1990s, and since then, a lot of research and studies have been done in this area (Daniel, 2004). On the other hand, the intensity of competition in the scientific, social, economic, political and even military field has intensified the importance of speed and access to the information and knowledge. On the one hand, the need to design systems capable of fast exploration of users' information with minimal emphasis on humans' intervention and the need for appropriate statistical methods for assessing high volume of data, on the other hand, seem extremely important. For achieving these goals, various methods such as statistical analysis and database search tools were proposed, but each of them has got some limitations which make them inaccessible and infeasible in all circumstance, for example, the statistical analysis generally examines the description of the statistical test and the data are referred to as the entire population data.



Therefore, a new method is needed to put aside the deficiencies of other methods, but uses the other methods and take advantage of their benefits. Clustering is considered as one of the main rules with which the data are analyzed and categorized based on certain criteria in which the same samples in one category will find maximum differences with the samples in other categories. Clustering has been employed in different sciences including statistics, pattern recognition, machine learning, data mining and bio-informatics (Shi and Malik, 2000; Arbelaez et al., 2011; Eisen et al., 1998; Sˆırbu et al., 2012).

STATEMENT OF THE PROBLEM

Nowadays, due to existing complexity and changing organizations, using accurate and modern tools is compulsory for performing the maintenance strategies. Human's improvements during last decades about gathering and storage of the results and data causes the organizations to have a huge dimension of data related to maintenance. The main point in such database is the information and knowledge that is discovered and the use of intelligent and structural methods is needed. Data mining is a new scientific field of recycling information from databases. The designing of description, explanatory, prediction and control models by using data mining techniques would be easier. These techniques would make analysis, programming, controlling and inspecting accessible by recognizing effective factors on events (Pariazar, Zaeri and Shahrabi,2007). Today the electric networks become larger and more complex with big data received from a lot of events in different sections, among which power transformer is one of the most important sections in power systems. Any fault in the transformer can cause a severe outage, which, therefore, necessitates continuous monitoring and diagnosis of its operation.

In this sense, any faults caused in power transformers will produce a lot of alarms, some of which are uncertain, incomplete and misinformed, thus, it is necessary to develop a good method to help dispatchers evaluate where the faults are and which transformers fail. However, transformer fault diagnosis decision-making based on dissolved and free gas analysis (DGA) diagnostic methods may give conflicting analysis results and complicate the final decision making by operators (Bacha, K., Souahlia, S. and Gossa, M, 2012). In fact, intelligent fault diagnosis systems are necessary to deal with changes in the typology of power network to fast diagnose the fault state and location of power transformers faults(Wang, T., Zhang, G.X., Zhao, J., He, Z., Wang, J. and Pérez-Jiménez, M.J, 2014).

According to the purpose of this study, the following objectives have been identified:

(i) To obtain the noise sampling using clustering ensemble.
(ii) To improve the classifier algorithm for fault diagnosis through noise detection.
(iii) To compare the results based on accuracy.

Based on the main description of the research, the main questions addressed in this study is as follows: How can the analysis of dissolved gases (DGA) in the power transformer oil fault be diagnosed? In order to be able to answer this question, a set of research questions is presented as follows:

1. How can noise sampling be obtained using clustering ensemble?
2. How can the classifier algorithm for fault diagnosis be improved by noise detection?
3. How can the results based on accuracy be compared?

METHODOLOGY

The methodologies used to meet the research objectives were explained comprehensively. The three different phases and subsequently the research strategy with its own steps were presented which starts by problem identification and finishes by research completion. Research problems and research objectives which are obtaining the noise sampling using clustering ensemble and improving the classification algorithm for fault diagnosis through noise detection and comparing the results based on accuracy, were stated in this chapter and the methods that showed how to solve the problems and meet the objectives with implementation. Two publicly available datasets (real dataset and artificial dataset) will be used in this research and result evaluation and validation are discussed. The evaluation of noise sampling, improving the classification algorithm and comparing the results methods are SVM, Naive Bayes and Decision Tree, KNN and RF .

Based on the given description and the rapid advance of data mining and its entry into the power industry equipment repairs and maintenance, in this study, we sought to apply data mining to reduce costs and increase power system reliability through timely diagnosis and prediction of power plant transformers faults which are one of the most sensitive equipment in power industry; and to do so, we need to take some samples from a number of transformers and conduct dissolved gases in oil analyses. After obtaining the data, they should be clustering and classified; and subsequently by available clustering groups, they are analyzed in order to understand and predict we transformer faults.

As can be seen, one of the steps in this process is data mining. It is clear that the success in the data mining affected by three previous steps so that if any such steps are not performed correctly, the results of data mining may not be useful but rather they can be misleading. The data mining techniques are modern scientific techniques for explaining, describing, predicting and controlling the phenomenon (Hair، Joseph F, 2005). These techniques are used to measure, explain and predict the dependence degree between variables. Data mining methods not only influence the analytical aspects of the studies, but also impact on the design and data gathering tools for making decisions and solving problems. Data mining process can be shown in the following steps as in Figure 1 (Pang-Ning Tan, Steinbach, 2005):

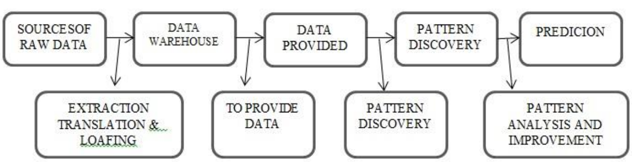


Figure 2: Data mining process

DATA COLLECTION

In this study, used the test results of oil Gas Chromatography (DGA) of 15 transformers in Neka power plant which have been continuously carried out by the approved laboratory of the Energy ministry for the past 8 years at certain times. Totally, 151 samples were able to record dissolved oil and gases and 8 parameters according to the IEEE standard were listed in the required tables for each transformer oil sampling of and the corresponding numbers were also registered based on which and in accordance with each gas dissolved in transformer oil and according to IEEE and IEC60599 standards, transformers defects can be predicted. Among these gases there are 5 important gases based on which each transformer fault can be predicted and diagnosed according to the proportion of each of these gases. In the next step, after the formation of 4 table called D, E, D1 and E1which will be explained accordingly, by using methods such as Noise Detection Using Clustering Ensemble and removing them, new data will be obtained (D2, E2…………D32, E32); and consistent with Classifier Ensemble, more accuracy and ultimately comparing them, the best method for early detection of transformers defects can be determined and introduced. What follows are the tables and explanations of each of the data.

FINDING OF STUDY

According to the description given in this Section, we run all data sets with different clustering ensemble algorithms including cspa, hgpa and mcla in MATLAB software, and record their accuracy results according to the Table 1 .

Table 1: % Accuracy of four data set in CE (D, E, D1, E1)

	D	**E**	**D1**	**E1**
CSPA	61.58	64.23	56.39	84.4
HGPA	56.95	64.29	56.39	84.26
MCLA	59.6	42.38	43.6	31.46

Now, according to the above table, we see the curve for the Accuracy of four data set in CE (D, E, D1, E1) in Fig. 4.2. According to the description given in Section 4.2, we run all data sets with different clustering ensemble algorithms including cspa, hgpa and mcla in MATLAB software, and record their Noise Number results according to the Table 2 .

Table 2 : The Noise Number of four data set in CE (D,E,D1,E1)

	D	E	D1	E1
CSPA	58	54	58	15
HGPA	65	53	58	14
MCLA	61	87	75	61

Now, according to the above table, we see the curve for the Noise Number of four data set in CE (D, E, D1, E1) in Table 2.

Conclusion of Noise Detection Using CE:

Given the above diagrams and the results of the accuracy obtained in the four datasets, it can be seen that among the algorithms of the clusterin ensemble, the cspa algorithm has the highest accuracy in detecting the noise, and also the E1 Data set, among the total data, has the highest Accuration And the minimum number of noise.

Classifier ensemble before Noise Detection

In this section first, all data sets D2, E2, D3, and E3, in accordance with the explanations given in Chapter 3, by the Classifier ensemble algorithms (SVM, RF, NAIVE BAYES, DECISION TREE, KNN) in MATLAB software before Noise Detection is performed and the accuracy value of each category is checked after 10 times according to the Accuracy that the software gives us. It is worth noting that each of the classification algorithms (SVM, RF, NAIES BAYES, DECISION TREE, KNN) and the advantages and disadvantages of each of them are explained.

Data-sets Test for Classifier ensemble before Noise Detection

According to the description given in Section 4.3, we run all data sets with different Classifier ensemble algorithms including SVM, RF, NAIES BAYES, DECISION TREE and KNN in MATLAB software, and record their Accuracy average results after 10 times according to the Table 3.

Table 3 : The Accuracy of four data set in CL before Noise Detection (D2, E2, D3, E3)

	D2	E2	D3	E3
SVM	0.9566	0.8333	0.8461	0.8647
RF	0.9583	0.9579	0.8066	0.8163
NB	1	0.9756	0.9612	0.9482
DT	0.9866	0.9566	0.9192	0.8235
KNN	0.9733	0.8733	0.8269	0.8882

Now, according to the above table, we see the curve for the Accuracy of four data set in CL before Noise Detection (D2, E2, D3, and E3) in Table 3.

Conclusion of Classifier ensemble before Noise Detection

Based on the above diagrams and the results obtained from the classifier ensemble before noise detection with different algorithms including SVM, RF, NAIVE BAYES, DECISION TREE, KNN, two results are obtained in all Data sets. First, the NB (naïve bayes) algorithm has highest Accuracy in all of the Data sets and second, the NB algorithm the highest Accuracy value in the Data set D2.

Classifier ensemble After Noise Detection by Relabel Noises

In this section, initially, all D2, E2, D3, and E3 data sets are identified by the CSPA algorithm described in Chapter 3, NOISE, and after moving the initial results of those noises, in the last column instead Enter the number one, zero, and enter the number 1 instead of zero, and finally, by entering this information in the new Data sets D21, E21, D31, E31 with the unit sorting algorithms (SVM, RF, NAIVE BAYES, DECISION TREE, KNN) in MATLAB software again re-classifies each of these data and examines the accuracy of each of these categories after 10 times according to the Accuracy that the software gives us.

Data-sets Test for Classifier ensemble After Noise Detection By Relabel Noises

According to the description given in Section 4.4, we run all data sets with different Classifier ensemble algorithms including SVM, RF, NAIES BAYES, DECISION TREE and KNN in MATLAB software, and record their accuracy average results after 10 times according to the Table 4.

Table 4: the Accuracy of four data set in CL After Noise Detection By Relabel Noises (D21,E21,D31,E31)

	D21	E21	D31	E31
SVM	0.96	0.9	0.9576	0.9823
RF	0.9678	0.9322	0.9285	0.9941
NB	0.9966	0.9487	0.9664	0.947
DT	0.9933	0.9266	1	0.9823
KNN	0.9833	0.9366	1	0.947

Now, according to the above table, we see the curve for the Accuracy of four data set in CL After Noise Detection By Relabel Noises (D21,E21,D31,E31) in Table 4.

Conclusion of Classifier ensemble After Noise Detection by Relabel Noises

According to the above diagrams and the ACCURACY RESULTS obtained from all DATA SETs after classifier ensemble after Noise Detection by Relabel Noises and comparing them with the results obtained in this study, there is an increase in ACCURACY value in all tables of data. So that in Data set D31, the Accuracy DT & KNN value is 100% **.**

Classifier ensemble After Noise Detection by Remove Noises

In this section, first all D2, E2, D3, E3 data sets are identified by the CSPA algorithm described in Chapter 3, NOISE, and then all of the noise data is deleted and eventually recorded by Information in the new DATA SETs D22, E22, D32, E32 with the unit categorization algorithms (SVM, RF, NAIVE BAYES, DECISION TREE, KNN) in the MATLAB software re-classifies each of these data and the accuracy of each category, We will examine them after 10 times according to the ACCURACY that the software gives us.

Data-sets Test for Classifier ensemble After Noise Detection by Remove Noises

According to the description given in Section 4.5, we run all data sets with different Classifier ensemble algorithms including SVM, RF, NAIES BAYES, DECISION TREE and KNN in MATLAB software, and record their accuracy average results after 10 times according to the Table 4.5 .

Table 5: The Accuracy of four data set in CL After Noise Detection By Remove Noises (D22,E22,D32,E32)

	D22	E22	D32	E32
SVM	0.9722	0.9634	1	0.9928
RF	0.9694	0.9894	1	1
NB	1	0.9947	0.9933	0.9714
DT	1	0.9894	1	1
KNN	1	0.9684	1	0.9714

Now, according to the above table, we see the curve for the Accuracy of four data set in CL after Noise Detection by Remove Noises (D21, E21, D31, and E31) in Table. 5.

Conclusion of Classifier ensemble After Noise Detection by Remove Noises

According to the above diagrams and ACCURACY RESULTS, obtained from all DATA SETs after Classifier ensemble After Noise Detection by Remove Noises and comparing them with the results obtained in study, ACCURACY Classifier ensemble for DATA SETs only occurs if First, Noise Detection and further Remove Noises will increase the ACCURACY value significantly so that the number of data with an ACCURACY maximum of 100% is higher, indicating that this method can be The best way to detect and troubleshoot the equipment in the shortest time possible is to minimize the error. In the Classifier ensemble After Noise Detection, Remove Noises will have more ACCURACY than Relabel Noises according to table 6.

Table 6: Compare the increase rate of Accuracy all of the data set in single Classifier ensemble algorithms (S.CL)

	Data Set	Accuracy	Avr Accuracy
CL Without Noise Detection	**D2**	0.9749	0.9085
	E2	0.9193	
	D3	0.872	
	E3	0.8681	
CL With Noise Detection By Relabel Noises	**D21**	0.9802	0.9625
	E21	0.9288	
	D31	0.9705	
	E31	0.9705	
CL With Noise Detection By Remove Noises	**D22**	0.9883	0.9887
	E22	0.9810	
	D32	0.9986	
	E32	0.9871	

Compare the increase rate of Accuracy all of the data set in single Classifier ensemble algorithms(S.CL) .Now, according to the above table, we see the curve for the increase rate of Accuracy all of the data set in single Classifier ensemble algorithms(S.CL) .

Combined Classifier ensemble

To this end, we get the accuracy of categorizing each of the base-case algorithms on each of the data sets. In the next step, we want to combine the basic classification algorithms and, using the combination of the results obtained from each of the algorithms with the majority vote approach, we create a new result, regain its accuracy, and with the classification algorithms compare the base.

As explained above, we first give the algorithms, giving test samples and instructional samples to each algorithm, and our test samples are labeled. Now we put together the answers of each of the five algorithms, and we look at each instance of what each algorithm has given a tag to a sample. If among the five algorithms more than half of them, for example, a positive label to the test sample we have given, the final tag given to the sample is positive, and if more than half of the algorithms have tested the negative label to the sample, the final tag for that instance will be negative. After reviewing and final labels given to each test case, by comparing the final result and the labels assigned to each of the samples in the original dataset, the classification accuracy for the combination method is obtained by the majority vote approach comes. First, a variety of combinational categorization methods are described below:

Types of Combined Classifier ensemble Methods

Below, various methods in the Combined Classifier ensemble Methods and their application are described briefly.

Majority voting scheme

In the voting design method, the majority vote for each algorithm generates responses, and if more than half the number of algorithms are individually given the same value of the tag as the output, the combined category response is the same as the output of the number The majority of algorithms.

Bagging training algorithm

This concept is used to combine predictable classifications from several models. Specify that you intend to create a model for prediction classification, and that the data set is small. You can select examples (by substitution) from the data set. And for samples derived from the classification tree (for example, C & RT and CHAID). In general, you will get different trees for different specimens. Then, a simple vote is taken to predict with the help of different trees obtained from the samples. The final classification will be the classification that different trees have foreseen.

Training Boosting Algorithm

Boosting is a hybrid algorithm in the field of machine learning that is used to reduce imbalances as well as variances. This method is used in supervised learning and is a family of machine learning algorithms. This technique is a method of transforming weak learning systems into strong based on the combination of different class results. The initial idea of this methodology is based on the question posed by Cairns and Braid in 1988 and 1989, which can create a strong learning system by combining a set of poor learning systems. A poor learning system is a learner who acts just as slightly randomly as a classifier (guesses sample labels better than random). In contrast to the strong learner, the classifier can alone predict sample labels well.

Although the boosting is not algorithmic, most of the bosting-based algorithms train poor learners as repetitive ones and add to the previous set to eventually reach a strong classifier. Poor learners are weighted when added to the set, which is usually based on the degree of accuracy in the sample classification. After adding each category, the existing samples (data) are weighed (their weight is corrected). The weighing of the samples is such that, at each stage, the weight of the samples that are categorized correctly is reduced, and the weight of the samples not properly classified is increased up to the next (by the learners new) are more attentive and more accurately classified. Therefore, the focus of the new weak learners will be on the data that the system could not correctly classify in the previous stages. So far, there have been many algorithms for boosting, but the original version of these algorithms has been provided by Robert Schepper and Uvo Ferond, which is not compatible and does not make full use of the benefits of poor learners. Later, these two-odd Adabost algorithms, which are compatible with the boost algorithm, have been presented with a specimen and Gladl's credible prize.

Consensus voting scheme

In the Consensus voting scheme, all classification of algorithms should all be reached in the same answer. If all algorithms recognize that the data is healthy, then the combination of the classification category classifies the data into a healthy category.

Data-sets Test for Combined Classifier ensemble by a Majority voting scheme(MVS)

Here, the accuracy of each of the above datasets (D2, E2, D3, E3, D21, E21, D31, E31, D22, E22, D32, E32) with the Combined Classifier ensemble algorithm by a majority voting scheme after ten times Calculate RUN in MATLAB and compare with the results obtained in the previous steps in table 8 .Now, according to the above table, we see the curve for the Accuracy all of the data set with the Combined Classifier ensemble by a majority voting scheme (C.CL in MVS).

According to the chart above, it can be seen that despite the relative increase of the Accuracy value of all Data sets at this stage, the test with the combination classifier ensemble algorithm with the majority voting scheme compared to the results of sections 4.3.2, 4.4.2, 4.5.2 and 4.6.2 The Accuracy data of NOISE DETECTION BY REMOVE NOISES (D22, E22, D32, E32) is much higher than the other data, and Combined classification method uses a majority vote design to have a higher reliability and accuracy than the single classification methods, this is what we expect from this research according to the table 8.

Table 8 : Compare the Accuracy all of the data

ACCURACY	Average of Accuracy S.CL[2]	Average of Accuracy C.CL[1] MVS
D2	0.9749	**1**
E2	0.9193	**0.9798**
D3	0.872	**0.9363**
E3	0.8681	**0.9514**
D21	0.9802	**0.9809**
E21	0.9288	**0.9931**

[1] C.CL : Combined Classifier ensemble
[2] S.CL : Single Classifier ensemble

D31	0.9705	**0.9923**
E31	0.9705	**1**
D22	0.9883	**1**
E22	0.9810	**0.9947**
D32	0.9986	**1**
E32	0.9871	**1**

Table 8 shows Compare the Accuracy all of the data set with the Combined Classifier ensemble by a majority voting scheme and the single classifier ensemble algorithms. Now, according to the above table, we see the curve for the increase rate of the Accuracy all of the data set with the Combined Classifier ensemble by a majority voting scheme and the single classifier ensemble algorithms and compere them in Table. 8.

In this section, we want to summarize the average growth rate of ACCURACY of all DATA SETS in the three research phases of this project between the single classifier ensemble and the combined classifier ensemble by the majority voting scheme algorithms in Table 9.

Table 9 : Compare the increase rate of Accuracy all of the data

	Data Set		**Accuracy**		**Avr Accuracy**	
	S.CL	C.CL MVS	S.CL	C.CL MVS	S.CL	C.CL MVS
Without Noise Detection	D2	D2	0.9749	1	0.9085	0.9668
	E2	E2	0.9193	0.9798		
	D3	D3	0.872	0.9363		
	E3	E3	0.8681	0.9514		
With Noise Detection By Relabel Noises	D21	D21	0.9802	0.9809	0.9625	0.9915
	E21	E21	0.9288	0.9931		
	D31	D31	0.9705	0.9923		
	E31	E31	0.9705	1		
With Noise Detection By Remove Noises	D22	D22	0.9883	1	0.9887	0.9986
	E22	E22	0.9810	0.9947		
	D32	D32	0.9986	1		
	E32	E32	0.9871	1		

Table 9 shows compare the increase rate of Accuracy all of the data set in three phase between single Classifier ensemble and combined Classifier ensemble by the majority voting scheme algorithms(S.CL & C.CL MVS). Now, according to the above table, we see the curve for the increase rate of Accuracy all of the data set in three phase between single Classifier ensemble and combined Classifier ensemble by the majority voting scheme algorithms(S.CL & C.CL MVS) .

As predicted from the above curve, among the three research phases in classifier ensemble of this project, our data sets in phase 3, Classifier ensemble with Noise Detection By Remove Noises, have the highest accuracy, and in the case of using the Combined Classifier ensemble algorithm by a Majority voting scheme will greatly increase the accuracy of the single classifier ensemble for all data, which has been the main objective of this study.

CONCLUSIONS

Knowledge discovery techniques and learning efficiency in data mining depend on a variety of criteria, including quality. The quality of ourselves is due to factors such as accuracy, accuracy and correct recognition rate that our focus is on the health factor.

Clustering ensemble is an approach widely adopted in clustering research to improve the quality and robustness of clustering results. In the context of machine learning, classification is supervised learning and clustering is unsupervised learning .Also have a look at Classification and Clustering at Wikipedia. If you have asked this question to any data mining or machine learning persons they will use the term supervised learning and unsupervised learning to explain you the difference between clustering and classification. Combined Classification learning methods use multiple categorization methods to create a set of prediction models whose output is combined into a unit prediction and aims to enhance classification accuracy and classification reliability. The advantage of these algorithms is their unit reliability algorithms, so that a single algorithm may have a high accuracy for a dataset and has lower accuracy for another dataset. Hence, using a combination category algorithm that is roughly the average of the values obtained from unit sorting algorithms, it has higher reliability. In this study can be deduced that the best of systems in the CSPA, HGPA, MCLA , CSPA system, due to the high total accuracy and the lowest number of errors in the total data noise detection, as well as from Data sets, the data for Table E1 has the highest Accuracy and the lowest noise . In Section 4.3, this can be deduced the Accuracy results obtained from D3, E2, D2, and E3 Data sets after Classification without Noise Detection with 5 different SVM, RF, Naive Bayes, Decision Tree, KNN algorithms, KNN, are observed in all Data tables The NB system (Naive Bayes) has the highest Accuracy, and the NB (Naive Bayes) algorithm with the Data sets D2 has the highest Accuracy.

According to the accuracy results obtained from all Data sets after Classification Whit Noise Detection by Relabeled Noises and comparing them with the results obtained in Section 4.3.4, there is an increase in Accuracy value in all tables of data. So that in Data sets D31, the Accuracy DT & KNN value is 100% . Accuracy Results, obtained from all Data sets after Classification whit Noise Detection by Remove Noises and comparing them with the results obtained in Sections 4.3 and 4.4, Accuracy Classification for Data sets only occurs if First, Noise Detection and further Remove Noises will increase the Accuracy value significantly so that the number of data with an Accuracy maximum of 100% is higher, indicating that this method can be The best way to detect and troubleshoot the equipment in the shortest time possible is to minimize the error. In the Classification Whit Noise Detection, Remove Noises will have more Accuracy than Relabeled Noises. In addition, it can be seen that despite the relative increase of the Accuracy value of all Data sets at this stage, the test with the combination classification algorithm with the majority vote plan compared to the results of the study. The Accuracy data of Noise Detection by Remove Noises (D22, E22, D32, E32) is much higher than the other data, and this is what we expect from this research.

Finally, according to research carried out by researchers in previous years on the diagnosis of transformer errors, according to the analysis of the gases in the transformer oil (DGA) based on the clustering ensemble and classification and the calculation of its accuracy, it can be concluded that if the data The correct results of the gas chromatographic tests of the transformers are extracted and using the correct CE methods, we can detect the noise in the Data sets and remove them and perform the new Classification Data sets. We will get the best and highest accuracy in the output. That this could be a huge help to us in defining the defects Transformers So that it will prevent the high costs of this important and precious equipment in the electric power industry, and in order to better recognize and increase Accuracy, these Data sets are also used by the majority vote plan type algorithm We tested the results we expected to increase Accuracy and then compare all of these results in this study. In sum, combined classification method uses a majority vote design to have a higher reliability than the single classification methods for Power transformer fault diagnosis based on noise detection using clustering and classifier ensemble. In this method, five basic algorithms such as backup vector machine, random forest, decision tree, bias, and nearest neighbor are used Is. We used nine sets of data to prove the functionality of the proposed function and examine the results.

REFERENCES

1. Arbelaez, P., Maire, M., Fowlkes, C. and Malik, J. (2011). Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 33(5), 898–916.

2. Bacha, K., Souahlia, S. and Gossa, M. (2012) Power Transformer Fault Diagnosis Based on Dissolved Gas Analysis by Support Vector Machine. Electric Power Systems Research, 83, 73-79. http://dx.doi.org/10.1016/j.epsr.2011.09.012.
3. Chan, F.T.S., Lau, H.C.W., Ip, R.W.L., Chan, H.K., Kong, S.(2005). Implementation of total productive maintenance: a case study . International Journal of Production Economics 95,71–94.
4. Fei SW, Sun Y.(2008). Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm. Electric Power Systems Research, 78:507–514.
5. Fei, S.W. and Zhang, X.B. (2009). Fault Diagnosis of Power Transformer Based on Support Vector Machine with Genetic Algorithm. Expert Systems with Applications , 36, 11352-11357. http://dx.doi.org/10.1016/j.eswa.2009.03.022
6. Gao. Y., X.S. Wang, Y.H. Cheng, and J. Pan.(2010).Fault diagnosis using a probability least squares support vector classification machine. Mining Science and Technology, Vol.20, pp.917-921.
7. Hao. X., and S. Cai-xin.(2007).Artificial Immune Network Classification Algorithm for Fault Diagnosis of Power Transformer.IEEE TRANS. ON POWER DELIVERY, vol. 22, no. 2, APRIL, pp. 930-935.
8. Hsu. C.W., C.J. Lin.(2002).A Comparison of Methods for Multiclass Support Vector Machines.IEEE Transactions on Neural Networks, vol 13, no 2.
9. Kantardzic, M. (2011). Data mining: concepts, models, methods, and algorithms. John Wiley & Sons.
10. Pang-Ning Tan, Steinbach.(2005). Introduction to Data Mining.

11. Pariazar M, Zaeri MS and Shahrabi J.(2007). Using Data Mining Techniques in Maintenance. IDMC07, Amir Kabir University.
12. Peng, H., Wang, J., Pérez-Jiménez, M.J., Wang, H., Shao, J. and Wang, T.(2013). Fuzzy Reasoning Spiking Neural P System for Fault Diagnosis. Information Sciences, 235, 106-116. http://dx.doi.org/10.1016/j.ins.2012.07.015
13. Rogers R.(1978). IEEE and IEC codes to imterpret incipient faults in transformer,using gas in oil analysis[J]. IEEE Trans on Elect rical Insulation.13 (5):3492354.
14. Sun. Y.J., S. Zhang, C.X. Miao, and J.M. Li.(2007).Improved BP Neural Network for Transformer Fault Diagnosis. Journal of China University of Mining and Technology, Vol.17, pp.138-142.
15. Werner Feilhauer, Edmund Handschin (2006) Interpretation of dissolved gas analysis using dempster-shafer's theory of evidence. KTh, Stockholm, Sweden-june , 11-15.
16. Zhang.Y, X. Ding, and Y. Liu.(1996).IEEE Transactions on Power Delivery 11, 1836-1841.
17. Zhao. X.Y., K.S. Lai, and D.M. Dai.(2007).An Improved BP Algorithm and Its Application in Classification of Surface Defects of Steel Plate. Journal of iron and steel research, international, Vol.14, pp.52-55.