Diagnostics of Power Transformer Faults Using Prediction Fusion of Neural Network (MLP) with Naïve Bayes Algorithm Based on DGA Data

Abdelmoumene **HECHIFA** [1], Abdelaziz **LAKEHAL** [2], Ridha **KELAIAIA** [3]

[1] *Department of Mechanical engineering August 20, 1955,*
Skikda University, LGMM Laboratory, Algeria,

[2] *Laboratoire de Recherche en Électromécanique et Sûreté de Fonctionnement,*
LRESF laboratory, Mohamed Chérif Messaadia University, Algeria

[3] *LGMM Laboratory, Faculty of Technology, August 20, 1955, Skikda University, Algeria*

Email of 1st a.hechifa@univ-skikda.dz, Email of 2nd a.lakehal@univ-soukahras.dz, Email of 3rd r.kelaiaia@univ-skikda.dz,

ABSTRACT : *Dissolved gas analysis (DGA) is still one of the most popular techniques for diagnosing transformer faults by adopting it as a primary technique for early detection of faults and increasing the reliability of the electrical system. This paper proposes an efficient fusion method based on DGA data using the two best algorithms accuracy, the neural network (MLP), the naïve bayes (NB) through ppm input vectors and percentages. The fusion method has predictively combined the two classifiers and obtained a higher accuracy than both of them, which reached 95.83%. This indicates that the proposed method is effective and promising for diagnosing faults in power transformers.*

KEYWORDS*: transformer faults; fusion method; MLP; NB; DGA;*

1. INTRODUCTION

Power transformers are among the most important units that are continuously monitored to ensure the quality of their work in the power system, and their failure means the failure of the system and leads to a catastrophe in transmission across networks [1]. Transformer failures result in thermal and electrical stresses, and this causes the oil to decompose, which leads to oxidation from the insulating oil, the most important of which are Hydrogen (H2), Methane (CH4), Ethane (C2H6), Ethylene (C2H4), Acetylene (C2H2), carbon monoxide (CO), carbon dioxide (CO2) [2].There are many ways to analyze these gases and the common approach is Dissolved Gas Analysis (DGA) [3]

The dissolved gases analysis is considered a diagnostic tool with an acceptable efficiency and is recognized by the international committees for diagnosing thermal and electrical faults resulting from oil or paper [4] .Based on the concentration of dissolved gases, rules and methods have been relied upon to simplify the concept of the relationship of gas to the expected fault [5].
For DGA analysis, methods for diagnosing faults in power transformers are necessary. At present, there are traditional methods represented in Dornenburg method [6], Rogers four ratios, IEC 60599 code methods [7], Duval triangle [8],pentagon [9] [10],The traditional methods are not accurate enough to know the type of fault, so they are weak and suffer from decision-making, but by turning to artificial intelligence, diagnosing faults for transformers has become available and with high

reliability. The methods represented are: neural networks [11] , fuzzy logic [12], support vector machines [13] , k-nearest neighbor [14], and Bayesian networks [15].
In this paper, both the neural network algorithm and the naive theory were presented using different input vectors and the method of Fusion them to diagnose the six power transformer faults using the KNIME analytics platform.
Where this paper presented as follows, starting with the methodology, by describing data, faults of transformers and various input vectors. The results of the proposed model were discussed and the input vector was chosen with the highest accuracy in order to combine the two algorithms and achieve better accuracy of the model. The results were presented. Finally, the conclusion is a brief explanation and interpretation of the results and a look at the promising fusion method in diagnosing faults in power transformers.

2.METHODOLOGY

2.1. Data Description:

Data collection is the first and most important step for each study, as it is an integral part of the work, thanks to which we can increase the life of all machines by diagnosing and predicting their faults. This data was collected from this study [16], which consists of (240 samples), including (129) electrical fault samples and (111) thermal fault samples, which in turn are divided into six faults as follows: partial discharges (PD=27 cases), low energy discharges (D1=42 cases), high energy discharges (D2=55 cases), thermal faults < 300 ◦C (T1=70 cases), thermal faults of 300 ◦ C to 700 ◦C (T2=18 cases), and thermal faults > 700 ◦C (T3=28 cases), where 70% was used for the training process (168 samples) and 30% for the testing process (72 samples).

2.1.1. Data Preparation:

To ensure the quality of any data mining process, the data must be processed by changing the input vectors and data format to facilitate dealing with them. This process is the most important step to increase the accuracy of the model. In this study, original data (ppm) input vectors and data input vectors were used in the form of percentages, as shown in corresponding Table1

Data Format	
input vectors (ppm)	$H2$ $CH4$ $C2H6$ $C2H4$ $C2H2$
input vectors (Percentage)	[$H2$/ (Total Gases)]× 100% [$CH4$/ (Total Gases)] × 100% [$C2H6$/ (Total Gases)]× 100% [$C2H4$/ (Total Gases)] × 100% [$C2H2$/ (Total Gases)] × 100%

(Total Gases) = $H2 + CH4 + C2H6 + C2H4 + C2H2$

3.Classification Algorithms:

3.1. Neural Network (MLP)

Artificial Neural Networks Neural networks are a symmetric processor that has great interconnectedness because it is an advanced and efficient algorithm for solving problems of data shortage and interference. Neurons are the basic base of neural networks, as they are inspired by the human brain, which consists of complex neurons and a lot of neutrons. [17].

The Multilayer Perceptron (MLP) are described as the most powerful and effective neural networks in dealing with data. They are among the category of supervised neural networks, depending on the expected outputs of machine learning [18]. On the other hand, the neural network (MLP) algorithm consists of three main layers, The input layer is the connection of the data input vector to the network, The hidden level varies from one algorithm to another according to the sensory cells, The output layer is the processing of data for the input vector through the sensory cells and directing it through its activation function, which is represented by [19] .

$$y_i = f\left(\sum_{j=1}^{n} \omega_{ij} x_j + \theta_i\right)$$

x_j: is the jth input of the ith neuron.
ω_{ij}: is the weight from the jth input to the ith neuron.
θ_i: is called the bias of the ith neuron.
y_i: is the output of the ith neuron.
The corresponding Figure 1 represents the classified neural network (MLP) architecture, which consists of three layers, the input layer, the hidden layer, and the output layer.

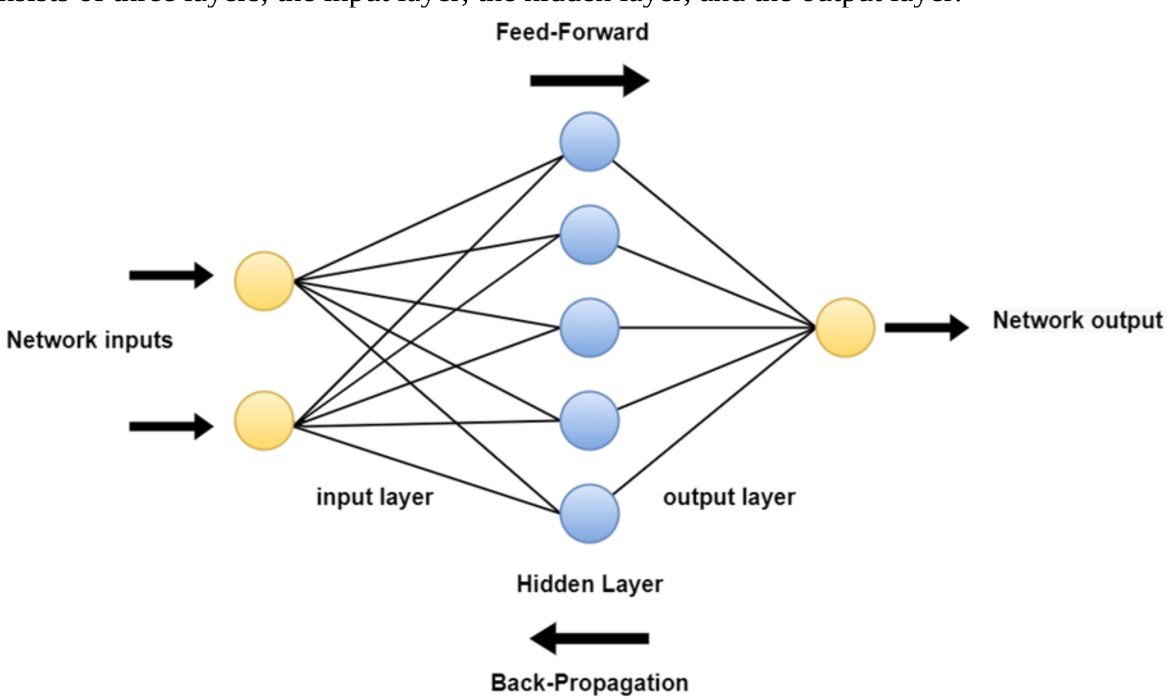


Figure 1 The architecture of the MLP-based classifier

3.2. Naïve Bayes (NB)

The naïve bayes is a classification algorithm inspired by the Bayes theorem. The principle of its work is based on the probability and statistical methods presented by the British scientist Thomas Bayes, which establish independent values based on what is before it [20] . Since this algorithm assumes all variables are independent of class values, this tends to be absent in real life, so it has been called naive and most importantly, it learns quickly [21]. The Bayes Theorem algorithm has the advantage of being a classifier for different objects, and the posterior probability equation can be described as follows [22] :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$P(A|B)$ is the posterior probability of *A* under condition *B*.
$P(A)$ is the prior probability of *A*.
$P(B|A)$ is the posterior probability under condition *A*.
$P(B)$ is the prior probability of *B*.
The posterior probability can also be represented as:

$$Posterior = \frac{likelihood \times prior}{evidence}$$

The principle of the naïve bayes algorithm can be illustrated through a simplified diagram and how to classify things, Figure 2 illustrates the basic structure of a naïve bayes.

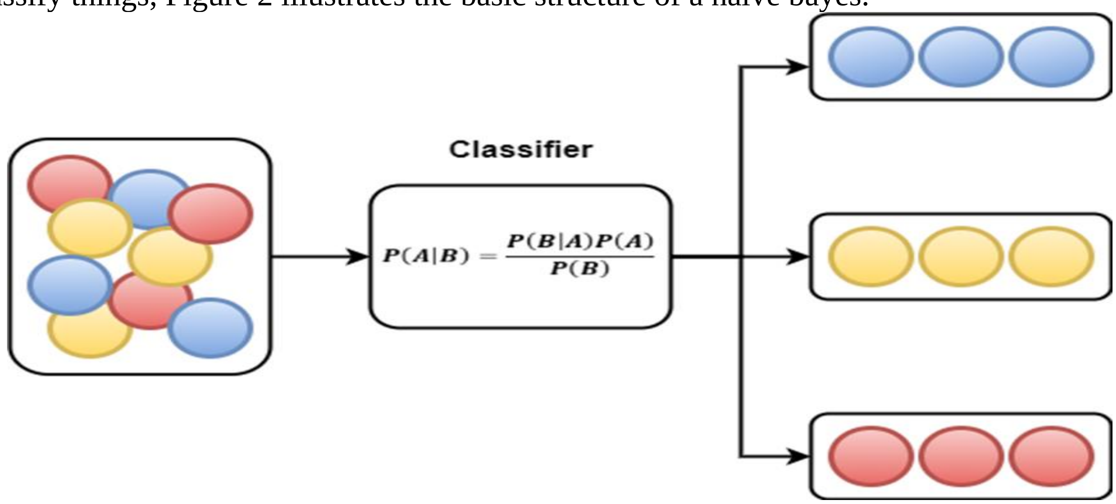


Figure 2 simplified diagram of the naïve bayes algorithm

4. MODEL EVALUATION

In the science of data mining, there is a reliable way to measure the accuracy of data so that the data enables us in documentation support systems, which is a confusion matrix consisting of several terms to evaluate the performance results represented in True Negative (TN), False Positive (FP), True Positive (TP), and False Negative (FN) [23] .
The accuracy is a percentage for evaluating the model, by which the efficiency of the algorithm is known, and it is as follows:

$$\text{Accuracy} = \frac{\text{TN+TN}}{\text{TP+TN+FP+FN}} \times 100\%$$

5.RESULTS AND DISCUSSIONS

To evaluate the effectiveness of the proposed model for both ppm and percentage input vectors, a data set consisting of 240 samples was used, which are as follows: 160 samples for training and 70 samples for testing. Where this paper represents one of the most powerful and accurate algorithms, namely, the naive rule and neural networks in diagnosing transformer faults, which in turn, their predictions were combined through ensemble learning to increase the accuracy of the model, using the KNIME analytics platform. Figure 3 represents the proposed model structure.

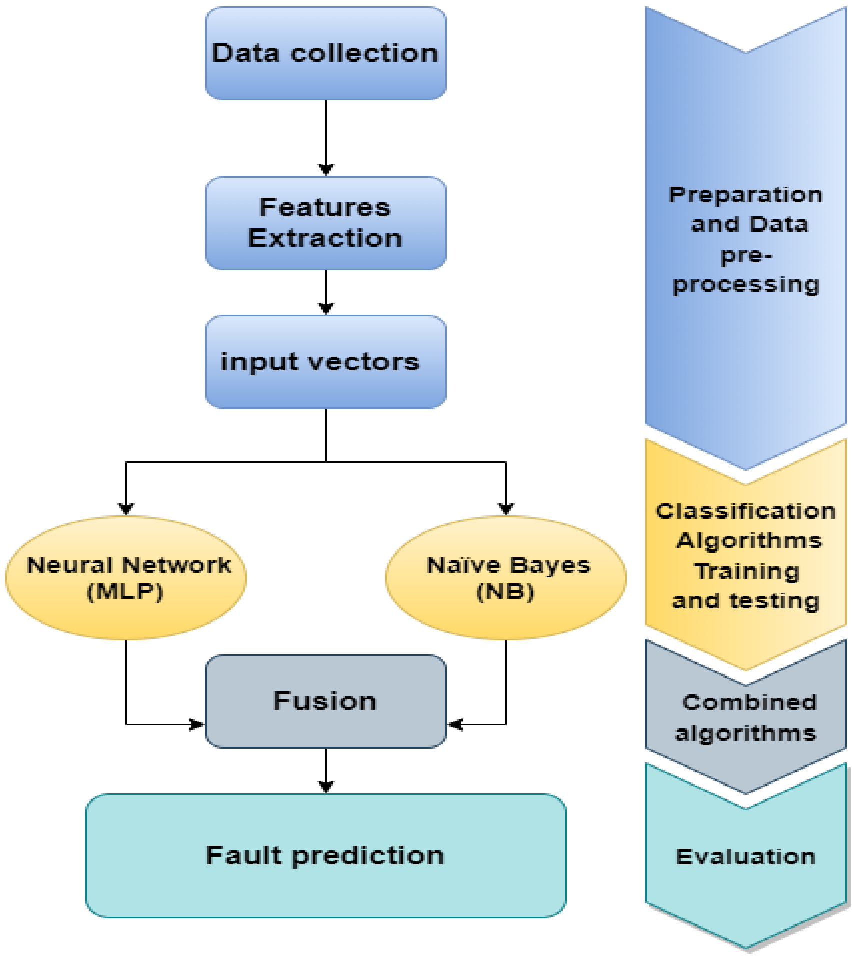
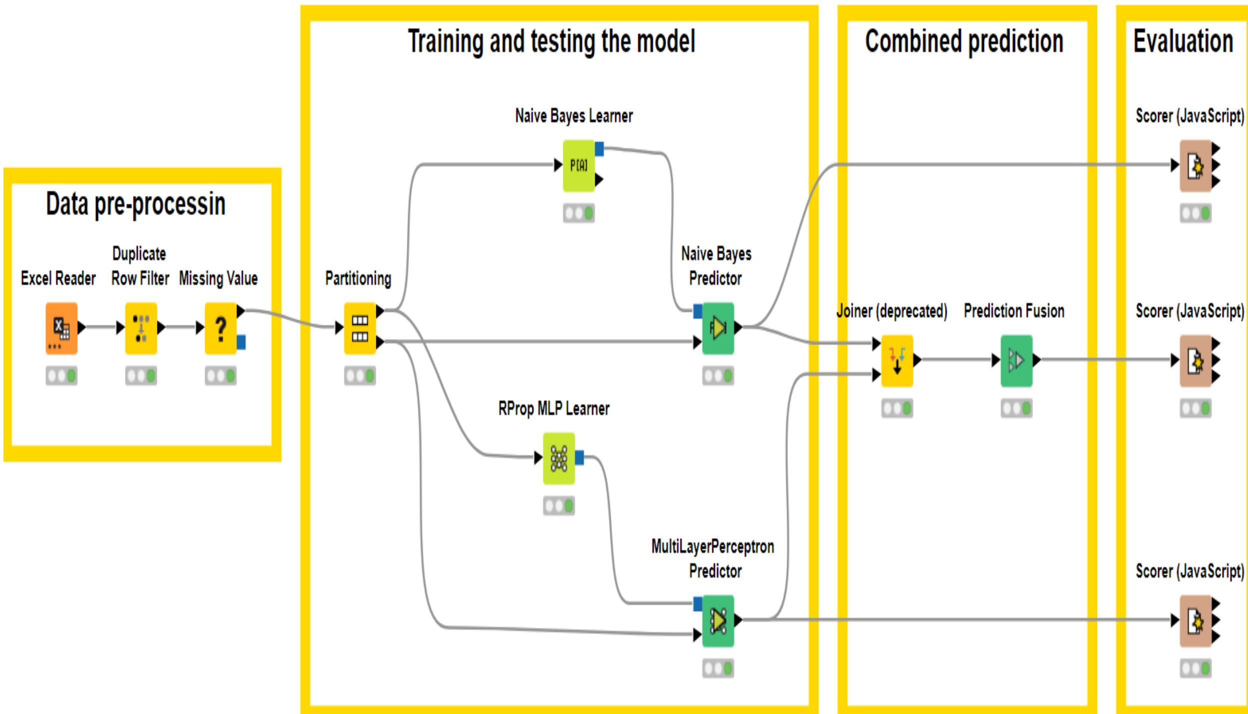


Figure 3: Flowchart of the proposed model.

The KNIME analytics platform environment allows engineers to develop and implement algorithms in a short time through a group of interconnected nodes, each of which performs a specific function and enables the expert to enter, output and modify data [24]. Figure 4 represents the proposed model and prediction Fusion with a simplified explanation of the steps.

- ✓ **Data pre-processing:** Data is entered by the reader node, which scans the input file to determine the number and types of columns, Then the duplicate rows node removes all the repeating Rows From the input table and then processes the missing values.

- ✓ **Training and testing the model:** first, the data is divided into training and testing by the Partitioning node, the model is trained and tested by the node (Learner and Predictor)
- ✓ **Combined prediction:** The predictions are collected first by the joiner node, which keeps the class and prediction columns, and the Prediction Fusion node is collected using the mean predictions of both naive bayes (NB) and neural network (MLP) algorithms.



- ✓ **Evaluation:** The proposed model is evaluated by Confusion matrix and classifier accuracy.

Figure 4: The proposed model using Prediction Fusion to combine predictions of different classifiers.

It was noted from the table1 that the highest accuracy is always when using a percentage input vector for both (MLP) and (NB) algorithms, respectively 93.06%, 91.67%, so the predictions for both will be compiled using percentage input vectors.

Table 1 shows accuracy results for both classifiers using the proposed input vectors.

	input vectors (ppm)	input vectors (Percentage)
MLP	80.56%	93.06%
NB	62.50%	91.67%

Confusion matrix is one of the most important tools for evaluating model accuracy between fault actual and prediction results (PD=1, D1=2, D2=3, T1=4, T2=5, T3=6) for the input vector as percentages.

Figure 5 The horizontal cells in blue slashes indicate the number of correctly categorized data, and the rest of the cells indicate the incorrectly categorized data by the proposed model. Figure 5.A shows that the faults were correctly rated for PD, D1, T1 and T2, while the D2 fault was classified three times as D1 and once as T3 and the fault T3 was classified once as D1. Figure 5.B shows that the faults were correctly rated for PD, D1, T2, and T3, but the D2 fault was incorrectly rated twice T1 and once T3 and the T1 fault was incorrectly rated once PD, once D2 and once T2.

	1 (Predicted)	2 (Predicted)	3 (Predicted)	4 (Predicted)	5 (Predicted)	6 (Predicted)	
1 (Actual)	9	0	0	0	0	0	100.00%
2 (Actual)	0	12	0	0	0	0	100.00%
3 (Actual)	0	3	16	0	0	1	80.00%
4 (Actual)	0	0	0	17	0	0	100.00%
5 (Actual)	0	0	0	0	5	0	100.00%
6 (Actual)	0	1	0	0	0	8	88.89%
	100.00%	75.00%	100.00%	100.00%	100.00%	88.89%	

(A) Confusion matrix (MLP) Input vector percentages

	1 (Predicted)	2 (Predicted)	3 (Predicted)	4 (Predicted)	5 (Predicted)	6 (Predicted)	
1 (Actual)	9	0	0	0	0	0	100.00%
2 (Actual)	0	12	0	0	0	0	100.00%
3 (Actual)	0	0	17	2	0	1	85.00%
4 (Actual)	1	0	1	14	1	0	82.35%
5 (Actual)	0	0	0	0	5	0	100.00%
6 (Actual)	0	0	0	0	0	9	100.00%
	90.00%	100.00%	94.44%	87.50%	83.33%	90.00%	

(B) Confusion matrix (NB) Input vector percentages

Figure 5: Confusion matrix Model

Through the confusion matrix, it is clear that there is a difference in the predictions for each of the two classifiers, meaning that each algorithm is effective in predicting a specific fault. Hence the idea of compiling the predictions based on the strengths of both classifiers. Where the results of Prediction Fusion showed that when compiling the predictions, they depend mainly on the mean of the predictions. For example, when Predict (1) for the neural network (MLP) algorithm, the value was 0.985 and for the naïve bayes (NB) algorithm, the value was 0.984, i.e. by summing the mean of the two values, we get 0.99 and thus the result is included as Class (PD=1) and so on for all classes. The corresponding table 2 shows how to integrate the mean predictions of the proposed model.

Table 2 Fusion prediction using mean in KNIME analytics platform.

		classes					
		PD=1	D1=2	D2=3	T1=4	T2=5	T3=6
weight	1	Predict(1) MLP	Predict(2) MLP	Predict(3) MLP	Predict(4) MLP	Predict(5) MLP	Predict(6) MLP
	1	Predict(1) NB	Predict(2) NB	Predict(3) NB	Predict(4) NB	Predict(5) NB	Predict(6) NB

After performing the fusion method, the accuracy of the proposed model was improved, as shown in the corresponding table 3.

Table 3 Clarification to improve the accuracy of the proposed model.

	MLP	NB	MLP with NB
input vectors (Percentage)	93.06%	91.67%	95.83%

Figure 6 shows the confusion matrix for fusion classifiers and improving predictions for the six different faults (PD=1, D1=2, D2=3, T1=4, T2=5, T3=6). PD, D1, T1, T2, and T3 are all correctly classified except for D2, which is incorrectly classified twice as D1 and once T3.

	1 (Predicted)	2 (Predicted)	3 (Predicted)	4 (Predicted)	5 (Predicted)	6 (Predicted)	
1 (Actual)	9	0	0	0	0	0	100.00%
2 (Actual)	0	12	0	0	0	0	100.00%
3 (Actual)	0	2	17	0	0	1	85.00%
4 (Actual)	0	0	0	17	0	0	100.00%
5 (Actual)	0	0	0	0	5	0	100.00%
6 (Actual)	0	0	0	0	0	9	100.00%
	100.00%	85.71%	100.00%	100.00%	100.00%	90.00%	

Figure 6: Confusion matrix (MLP) with (NB) Input vector percentages

Figure 7 shows bar graphs of the accuracy ratios, respectively, from the accuracy of the classifiers to the compilation of predictions and the improvement of the proposed model.

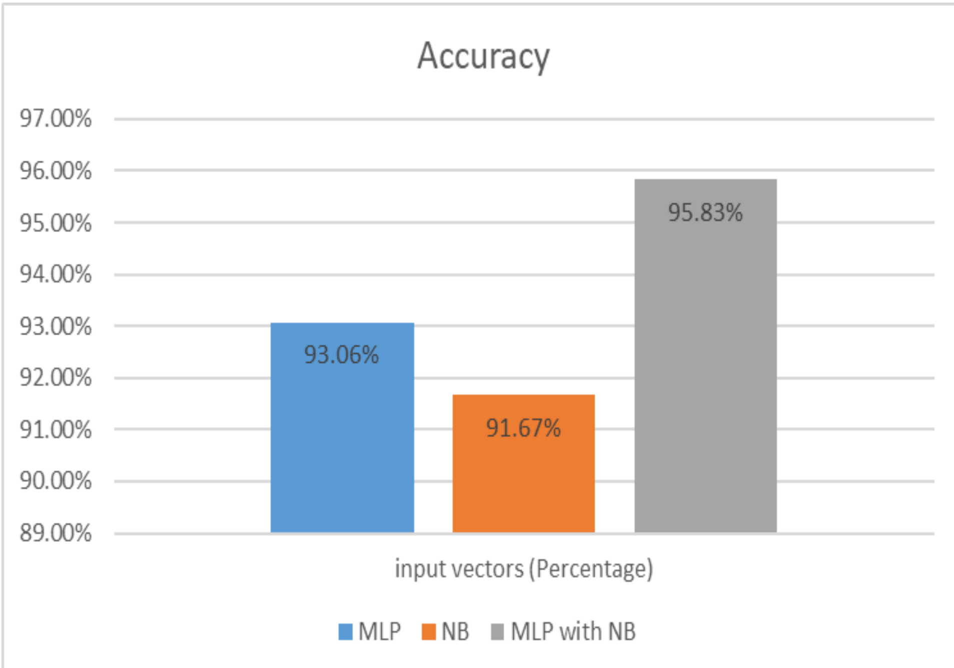


Figure 7: The different accuracy values obtained.

6. CONCLUSION

This paper proposes a promising technique for DGA data diagnostics using Fusion predictions for classification algorithms of both neural network (MLP) and naïve bayes (NB). The workflow for this paper was as follows First, by comparing the proposed input vectors, the preference for both classifiers was for the input vector in percentages, where in neural network (MLP) the diagnostic accuracy was 93.06%, and in the naïve bayes (NB) it was 91.67%, and when using the fusion method, the accuracy became 95.83% and a better result was obtained. This technique relied on the majority of votes by combining the mean predictions of the two algorithms and improving the efficiency of the model in diagnosing power transformer faults.

REFERENCES

[1] H. Nasirul, J. Aadil, C. Kingshuk and C. Soumya, "Accurate Sensing of Power Transformer Faults From Dissolved Gas Data Using Random Forest Classifier Aided by Data Clustering Method," *IEEE Sensors Journal*, vol. 22, no. 6, p. 5902 – 5910, 2022.

[2] I. B. Taha and D. A. Mansour, "Novel power transformer fault diagnosis using optimized machine learning methods," *Intelligent Automation & Soft Computing*, vol. 28, no. 3, pp. 739-752, 2021.

[3] C. Passara and V. Chanvichet, "Feature engineering and feature selection for fault type classification from dissolved gas values in transformer oil," *International Computer Science and Engineering Conference (ICSEC),* pp. 75-80, 2021.

[4] U. M. Rao, I. Fofana, K. N. V. P. S. Rajesh and P. Picher, "Identification and Application of Machine Learning Algorithms for Transformer Dissolved Gas Analysis," *IEEE Trans Dielectr Electr Insul.,* vol. 28, no. 5, p. 1828–1835, 2021.

[5] I. B. M. Taha, S. Ibrahim and D.-E. A. Mansour, "Power Transformer Fault Diagnosis Based on DGA Using a Convolutional Neural Network with Noise in Measurements," *IEEE Access,* vol. 9, p. 111162 –111170, 2021.

[6] E. Wannapring, C. Suwanasri and T. Suwanasri, "Dissolved Gas Analysis methods for distribution transformers," *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI CON),* pp. 1-6, 2016.

[7] I. B. M. Taha, S. S. M. Ghoneim and A. S. A. Duaywah, "Refining DGA Methods of IEC Code and Rogers Four Ratios for Transformer Fault Diagnosis," *IEEE Power and Energy Society General Meeting (PESGM),* p. 1–5, 2016.

[8] M. S. A. Khiar, S. A. Ghani, I. S. Chairul, Y. H. M. Thayoob and Y. Z. Y. Ghazali, "On-site OLTC monitoring using Duval triangle and DWRM," *In Proceedings of the 2nd International Conference on Technology, Informatics, Management, Engineering and Environment, TIME E 2014, Bandung, Indonesia,* vol. 1, p. 216–221, 2014.

[9] 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, p. . 9–12, 2014.

[10] D.-E. A. Mansour, "Development of a new graphical technique for dissolved gas analysis in power transformers based on the five combustible gases," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 22, no. 5, p. 2507–2512, 2015.

[11] A. Li, X. Yang, Z. Xie and C. Yang, "An optimized GRNN-enabled approach for power transformer fault diagnosis," *IEEJ Trans. Electr. Electron. Eng,* vol. 14, no. 8, p. 1181–1188, 2019.

[12] I. B. M. Taha, S. S. M. Ghoneim and H. G. Zaini, "A Fuzzy Diagnostic System for Incipient Transformer Faults Based on DGA of the Insulating Transformer Oils," *International Review of Electrical Engineering (I.R.E.E.),* vol. 11, no. 3, pp. 305-313, 2016.

[13] Y. Zhang, X. Li, H. Zheng, H. Yao, J. Liu, C. Zhang, H. Peng and J. Jiao, "A fault diagnosis model of power transformers based on dissolved gas analysis features selection and improved Krill Herd algorithm optimized support vector machine," *IEEE Access,* vol. 7, p. 102803–102811, 2019.

[14] Y. Benmahamed, Y. Kemari, M. Teguar and A. Boubakeur, "Diagnosis of power transformer oil using KNN and nave bayes classifiers," *IEEE 2nd International Conference on Dielectrics (ICD),* p. 1–4, 2018.

[15] J. I. Aizpurua, V. M. Catterson, B. G. Stewart, S. D. J. McArthur, B. Lambert and B. Ampofo, "Power transformer dissolved gas analysis through Bayesian networks and hypothesis testing," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 25, no. 2, p. 494–506, 2018.

[16] S. I. Ibrahim, S. S. Ghoneim and I. B. Taha, "DGALab: an extensible software implementation for DGA," *IET Gener. Transm. Distrib,* vol. 12, no. 18, pp. 4117-4124, 2018.

[17] K. D. PATEKAR and B. CHAUDHRY, "DGA analysis of transformer using Artificial neutral network to improve reliability in Power Transformers," *IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON),* p. 1–5, 2019.

[18] M. Ramkumar, C. G. Babu, K. V. Kumar, D. Hepsiba4, A. Manjunathan and R. S. Kumar, "ECG Cardiac arrhythmias Classification using DWT, ICA and MLP Neural Networks," *Journal of Physics: Conference Series,* vol. 1831, no. 1, p. 012015, 2021.

[19] W. Xiang, H.-D. Tran and T. T. Johnson, "Output reachable set estimation and verification for multilayer neural networks," *EEE Transactions on Neural Networks and Learning Systems,* vol. 29, no. 11, p. 5777–5783, 2018.

[20] Ketjie, V. C. Mawardi and N. J. Perdana, "Prediction of credit card using the Naïve Bayes method and C4.5 algorithm. IOP Conference Series," *Materials Science and Engineering,* vol. 1007, no. 1, p. 012161, 2020.

[21] Y. Mahamdi, A. Boubakeur, A. Mekhaldi and Y. Benmahamed, "Power Transformer Fault Prediction using Naive Bayes and Decision tree based on Dissolved Gas Analysis," *ENP Engineering Science Journal,* vol. 2, no. 1, pp. 1-5, 2022.

[22] V. Balaji, S. Suganthi, R. Rajadevi, V. K. Kumar, B. S. Balaji and S. Pandiyan, "Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier," *Measurement: Journal of the International Measurement Confederation,* vol. 163, p. 107922, 2020.

[23] F. Rahmad, Y. Suryanto and K. Ramli, "Performance Comparison of Anti-Spam Technology Using Confusion Matrix Classification," *IOP Conference Series: Materials Science and Engineering,* vol. 879, no. 1, p. 012076, 2020.

[24] G. Nicola, M. R. Berthold, M. P. Hedrick and M. K. Gilson, "Connecting proteins with drug-like compounds: Open source drug discovery workflows with BindingDB and KNIME," *Database-Oxford,* p. 1–22, 2015.