

Research Article

Transformer Fault Diagnosis Based on BP-Adaboost and PNN Series Connection

Chun Yan ¹, Meixuan Li,¹ and Wei Liu ²

¹College of Mathematics and System Science Shandong University of Science and Technology, Qingdao 266590, China

²College of Computer Science and Engineering Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Wei Liu; liuwei.doctor@yeah.net

Received 8 May 2019; Accepted 18 June 2019; Published 3 July 2019

Academic Editor: Gaetano Giunta

Copyright © 2019 Chun Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dissolved gas-in-oil analysis (DGA) is a powerful method to diagnose and detect transformer faults. It is of profound significance for the accurate and rapid determination of the fault of the transformer and the stability of the power. In different transformer faults, the concentration of dissolved gases in oil is also inconsistent. Commonly used gases include hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethane (C_2H_6), and ethylene (C_2H_4). This paper first combines BP neural network with improved Adaboost algorithm, then combines PNN neural network to form a series diagnosis model for transformer fault, and finally combines dissolved gas-in-oil analysis to diagnose transformer fault. The experimental results show that the accuracy of the series diagnosis model proposed in this paper is greatly improved compared with BP neural network, GA-BP neural network, PNN neural network, and BP-Adaboost.

1. Introduction

In recent years, with the rapid development of China's economy, power system is developing towards the direction of ultrahigh voltage, large power grid, large capacity, and automation. Domestic demand for electricity has increased dramatically, and the national power industry is experiencing a rapid development stage. At present, the number of 110KV (66KV) and above voltage transformers transported by the State Grid Corporation has reached more than 30,000, with a total capacity of 3.4 TVA. Because the power transformer is in the central position of the power grid, the operation environment is complex, and, under the impact of various bad operating conditions, it is easy that fault occurs. Transformer faults has caused large area of breakage, resulting in a large number of economic losses. Therefore, the effective diagnosis of transformer faults is of great significance.

At present, the main testing and monitoring methods of transformer operation state are DC resistance measurement [1], dissolved gas-in-oil analysis (DGA) [2], oil temperature monitoring (OTM) [3], insulation experiment (IE), acoustic partial discharge measurement (APDM) [4], detection of characteristic curve by repeated pulse method [5], Winding Deformation Testing and Low Voltage Short Circuit

Impedance Testing [6], etc. DGA analysis is a relatively ideal monitoring and analysis method, because it can monitor and analyze oil chromatography online and sample analysis at any time. At present, it is recognized that dissolved gas analysis technology in oil is a powerful measure to find latent faults in transformers at home and abroad [2].

Due to the normal operation of transformer, transformer oil and solid insulation materials will gradually age and decompose into a small amount of gas. However, when the power equipment is faulted, especially in the case of overheating, discharge, or humidity, the amount of these gases will increase rapidly. It has been proved by long-term practice that the content of gas in oil is directly related to the faulty degree of transformer [7]. Over the years, both at home and abroad have devoted themselves to the development of online monitoring devices and systems characterized by the content of dissolved gas in oil. The three-ratio method is the most basic method for fault diagnosis of oil-filled power equipment. However, the three-ratio method exposes that the coding cannot include all DGA analysis results and the coding is too absolute [8]. With the development of research, more and more artificial intelligence methods are introduced into transformer fault diagnosis, for example, artificial neural network [9], BP neural network [10], fuzzy logic reasoning [11],

rough set theory [12], Extreme Learning Machine [13], support vector machine [14], and Bayesian network [15]. However, all kinds of artificial intelligence methods have some limitations, such as the network structure and weight of artificial neural network which are difficult to determine and easy to fall into local minimum and overfitting. The Bayesian network requires a large number of sample data [16]. The inference rules of fuzzy logic and the function of fuzzy membership degree depend to a great extent on experience [6].

BP neural network is a kind of multilayer feedforward neural network, because of its simple structure, many adjustable parameters, many training algorithms, and good maneuverability. BP neural network has been widely used. According to statistics, 80%~90% of the neural network models are based on the BP network or its deformation [17]. The traditional BP neural network has the disadvantage of random initial weights, which lead to low learning efficiency, slow convergence speed, and easy to fall into a local minimum. So many scholars use intelligent algorithm to optimize the weight of BP neural network. Liang uses the genetic algorithm to optimize the weight of BP neural network to realize the inversion method of soil moisture [18]. Salman Predicts Palm Oil Price Using BP Neural Network Based on Particle Swarm Optimization [19]. Kuang uses ant colony algorithm to optimize BP neural network for macroeconomic prediction [20]. But using intelligent algorithms to optimize the weight of BP neural network greatly enlarges the operation time and makes the model diagnosis inefficient [21].

In this paper, a diagnostic model which combines BP-Adaboost algorithm and PNN in series is proposed. Adaboost algorithm is a simple and easy algorithm, which can combine several weak classifiers to form a strong classifier. At the same time, the upper limit of the classification error rate will not increase with the overfitting of training. In addition, the Adaboost algorithm has the advantages of no need to adjust parameters and low generalization error rate. Because the Adaboost algorithm can construct multiple weak predictors with lower accuracy into a strong learner with higher accuracy, therefore, this paper combines BP neural network as a weak classifier with Adaboost algorithm. Considering that Adaboost algorithm is usually used to deal with binary classification problems, and transformer faults are often divided into many types of faults, this paper changes the multiclassification problem into multiple Adaboost binary classification problems to be solved. In the Adaboost algorithm, only the error rate of the weak classifier is slightly less than 1/2, but in the actual training process, there will still be special cases, which will affect the operation of the algorithm. In order to solve this problem, this paper revalues individual variables under special circumstances. Then the transformer fault is diagnosed by the improved BP-Adaboost algorithm. Samples that have not been successfully classified in the diagnosis results (A sample is divided into two or more different faults or not into any faults.) are reclassified as prediction samples, and the original training samples are reclassified as training samples and put into PNN neural network for diagnosis. The advantages of BP-Adaboost and PNN are fully combined by this series model. With samples

which have not been successfully diagnosed by BP-Adaboost algorithm, the second diagnosis can be carried out by PNN, which effectively improves the accuracy of the model.

The main contributions of this paper are as follows.

(1) Firstly, the Adaboost algorithm is improved. It solves the defect that the diagnostic error of each weak classifier in the traditional Adaboost algorithm can only be within (0, 1/2). The two-class Adaboost algorithm is improved to multiclass algorithm.

(2) Then, BP-Adaboost multiclassification diagnosis algorithm is formed by combining the BP neural network as a weak classifier with multiclassification Adaboost algorithm. For samples diagnosed wrong in BP-Adaboost diagnosis model, they are put into PNN neural network for diagnosis again.

(3) Finally, the sample set is selected. Inspired by IEC three-ratio method, this paper not only takes five commonly used gas as characteristic parameters, but also takes C_2H_2/C_2H_4 , CH_4/H_2 , C_2H_4/C_2H_6 as characteristic parameters of transformer fault diagnosis.

Section 1 introduces the background significance of transformer fault diagnosis and the methods commonly used to diagnose transformer fault in recent years. Section 2 first introduces the model used in this paper and finally introduces the series multiclassification algorithm proposed in this paper. Section 3 introduces the selection of the sample set and sample characteristic parameters. Section 4 compares the diagnostic results of the proposed model with those of the other four models.

2. Materials and Methods

2.1. Improved BP-Adaboost Diagnostic Model. BP neural network is an error back propagation algorithm, which uses the steepest descent method to continuously adjust the weights and biases of the network are continuously adjusted to minimize the sum of square errors of the network. BP neural network consists of one input layer, one output layer, and several hidden layers. The training process is as follows.

(1) Establish BP neural network and initialize the weight and biases of BP neural network.

(2) Preprocess the sample data and set the number of neurons in each layer. Suppose $x = (x_{ij})$ ($i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, m$) is a sample input matrix. The output of the hidden layer is b_j . The biases of neurons in the hidden layer and the output layer are θ_j and θ_k , respectively. The output b_j of the j -th neuron in the hidden layer is

$$b_j = f_1 \left(\sum_{i=1}^n w_{ij} x_i - \theta_j \right) \quad (1)$$

The output layer output y_k is

$$y_k = f_2 \left(\sum_{j=1}^m w_{jk} b_j - \theta_k \right) \quad (2)$$

In formulas (1) and (2), f_1 and f_2 are S-type tangent function and S-type logarithmic function, respectively.

(3) Error e of actual output y_k and expected output t_k in BP neural network are

$$e = \sum_{k=1} (t_k - y_k)^2 \quad (3)$$

(4) If the errors produced do not meet the requirements, the steepest descent method is used to backpropagate the errors and adjust the weights and biases. Iterative cycle until the error meets the requirement.

Adaptive boosting (Adaboost) is a strong efficient algorithm that combines weak classifiers into strong classifiers. It was proposed by Yoav Freund and Robert Schapire in 1995. The main idea is as follows. Firstly, each training sample is given the same weight. Then the weak classifier is used to run iteratively T times; after each operation, the weight of training data is updated according to the classification results of training samples, and the wrong samples are usually given larger weight. For multiple weak classifiers, after running T times, a sequence of classification results of training samples is obtained. Each classification function is given a weight. The better the classification result, the greater the corresponding weight. The steps of the Adaboost algorithm are as follows.

Step 1 (randomly select m samples from the samples as training data). Initialization data distribution weights $D_t(i) = 1/m$. The structure of the neural network is determined according to the input and output dimensions of the samples, and the weights and biases of the neural network are initialized.

Step 2 (calculate the prediction error sum of weak classifier). When the t -th weak classifier is trained, the prediction error of prediction sequence $g(t)$ and e_t are obtained. In formula (4), $g_i(t)$ is the predicted result of the i -th sample and y_i is the expected classification result of the i -th sample. In BP neural network, the default output of more than 0 belongs to "1" category in classification, and the output of less than 0 belongs to "-1" category in classification.

$$e_t = \sum_i D_i(i) \quad i = 1, 2, \dots, m \quad (g_i(t) \neq y_i) \quad (4)$$

Step 3. Calculation of sequence weights a_t based on prediction error is

$$a_t = \left(\frac{1}{2}\right) \ln \left(\frac{(1 - e_t)}{e_t}\right) \quad (5)$$

Formula (5) shows that when e_t is less than 1/2, $(1 - e_t)/e_t$ increases with the decrease of e_t . $\ln(1 - e_t/e_t)$ is an increasing function, so the weight a_t of weak classifier increases with the decrease of error sum. However, as the number of iterations increases, the classification error decreases gradually. According to formula (5), when the error is reduced to a certain extent, it is easy to be unable to calculate the weight of weak classifiers, thus affecting the classification results [22]. Therefore, this paper reassigns the sum of errors in this case. Because there are 80 training samples in this paper, formula

(5) cannot calculate the weight; only the error is zero. So when the error of weak classifier predicting training data is zero, this paper makes error e_t be 0.0125, that is, only one sample predicting error. The calculation formula of sequence weight is established in the case of error and less than 1/2, but in the actual case, it is not ruled out that the prediction accuracy of weak classifier to the sample is less than 1/2. Therefore, this paper improves this situation. For the case where e_t is more than 1/2, it shows that the accuracy of weak classifier is less than 1/2 in the binary classification (-1 and 1), so that the output of weak classifier can be taken as the opposite number $e_t = 1 - e_t$, which can also be calculated by formula (5).

Step 4 (adjust the data weight). Adjusting the weight of the next training sample according to the sequence weight,

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} \times \exp[-a_t y_i g_t(x_i)] \quad i = 1, 2, \dots, m \quad (6)$$

In formula (6), B_t is a normalization factor. The purpose is to make the sum of distribution weights equal to 1 under the condition that the proportion of weights is unchanged. When the predicted results are not the same as the actual results, $y_i g_t(x_i)$ in formula (6) is less than 0, and the greater the absolute value of the predicted results, the greater the value of $\exp[-a_t y_i g_t(x_i)]$, thus satisfying the condition that the samples with wrong classification are usually given a larger weight.

Step 5 (construct a strong classifier). After several weak classifiers are trained by T -rounds, the T -group weak classifier functions $f(g_t, a_t)$ are obtained, and then the T -group weak classifier functions are combined to form a strong classifier function $h(x)$.

$$h(x) = \text{sign} \left[\sum_{t=1}^T a_t \cdot f(g_t, a_t) \right] \quad (7)$$

The algorithm flow based on BP-Adaboost model is shown in Figure 1.

2.2. PNN Neural Network. Probabilistic neural network (PNN) is a parallel algorithm based on Bayes classification rules and Parzen window for probability density function estimation. PNN is a kind of artificial neural network with a simple structure, simple training, and wide application. Its structure consists of an input layer, mode layer, summation layer, and output layer. Its basic structure is shown in Figure 2.

The values of training samples are first received through the input layer, and the number of neurons in the input layer is equal to the dimension of the input sample vector. Then the data information is transmitted to the pattern layer of the second layer through the input layer.

The pattern layer is used to calculate the matching relationship between input samples and each pattern in the training set. The number of neurons in the pattern layer is equal to the total number of training samples. Assuming that the vector of the input layer is $X = (x_1, x_2, \dots, x_n)^T$, the data

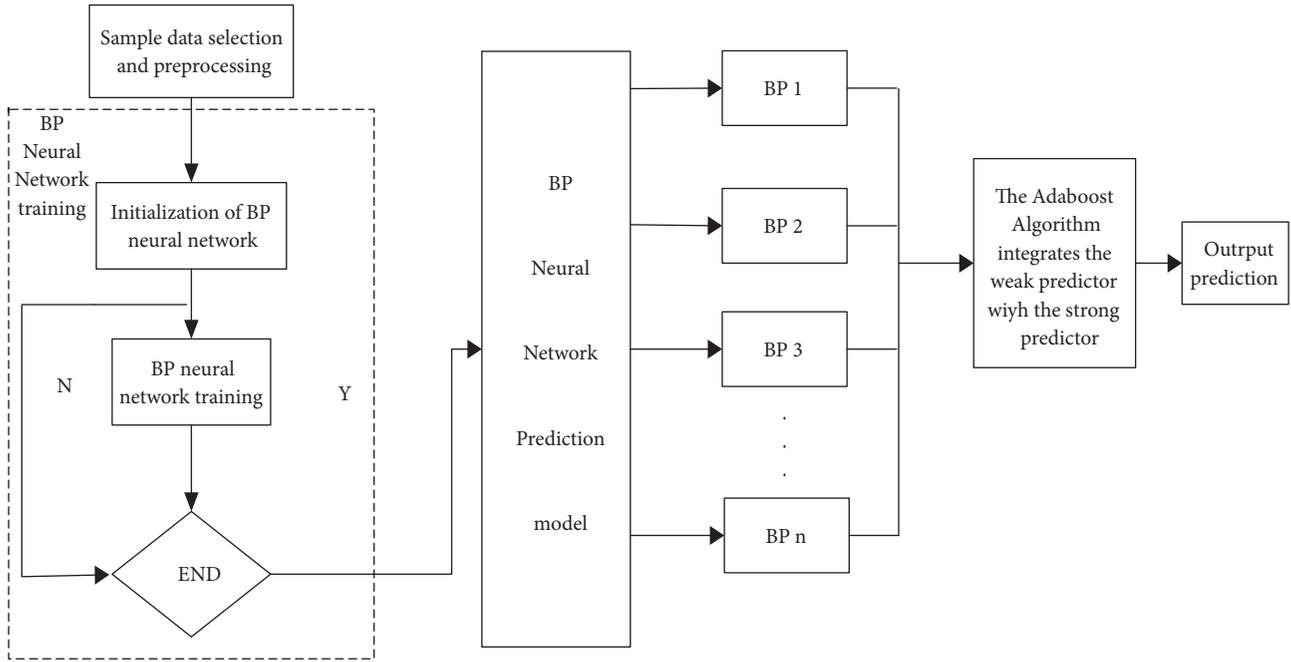


FIGURE 1: BP-Adaboost algorithm flow.

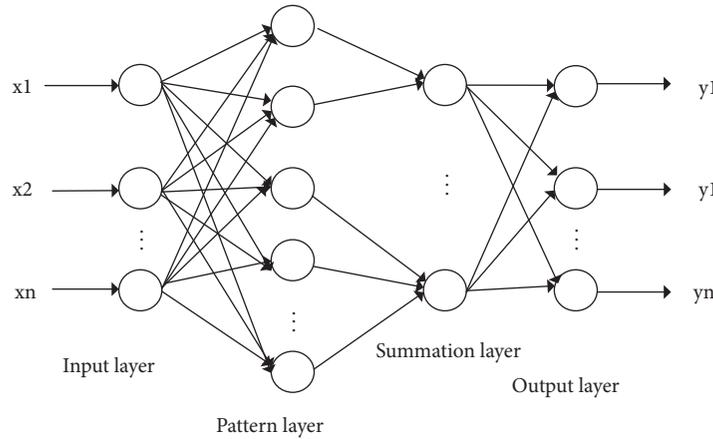


FIGURE 2: Basic structure of probabilistic neural network.

is mapped from the input layer to the pattern layer through the mapping mechanism, then the input of the j -th neuron of the pattern layer is X , and the output of the j -th neuron of the pattern layer is

$$f_{ij}(X) = \left(\frac{1}{2\pi^{p/2} \delta^p N_{R_1}} \right) \exp \left(-\frac{(x - w_{ij})^T (x - w_{ij})}{2\delta^2} \right) \quad (8)$$

In formula (8), p is the total number of categories, m is the number of neurons in the mode layer, w_{ij} is the connection weight from the output layer to the mode layer, and δ is the smoothing factor.

The summation layer is the accumulation of probabilities belonging to the same class, and its conditional probability density is

$$P(i | X) = \sum_{j=1}^m w_{ij} f_{ij}(X) \quad i = (1, 2, \dots, p) \quad (9)$$

of which

$$\sum_{j=1}^m w_{ij} = 1, \quad i = 1, 2, \dots, p. \quad w_{ij} \in [0, 1] \quad (10)$$

The output layer receives the probability density function of each class of summation layer output and then finds the maximum probability $P(X) = \max(P(i | X), (i = 1, 2, \dots, P)$ of one of them.

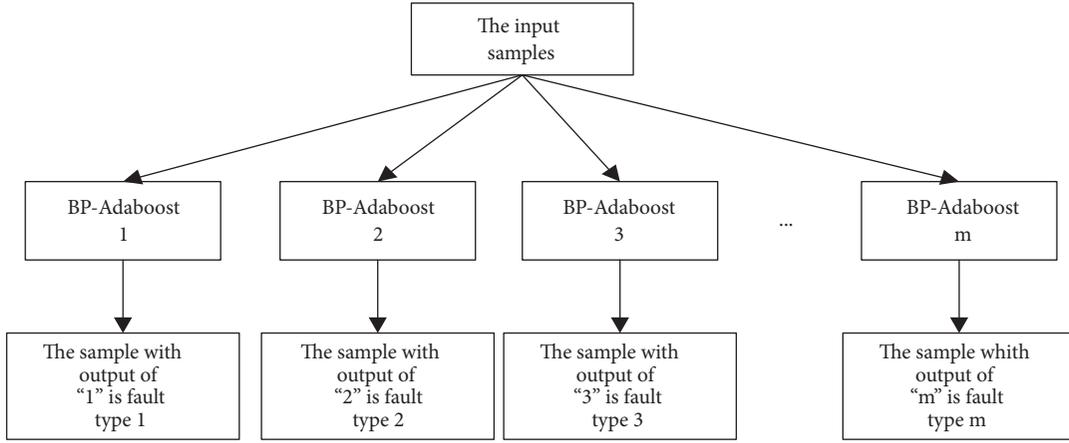


FIGURE 3: BP-Adaboost multiclassification model.

TABLE 1: Multiple BP-Adaboost classification result.

Sample	a	b	c	d	e
BP-Adaboost classification result 1	-1	1	-1	1	-1
BP-Adaboost classification result 2	-1	-1	1	1	-1
BP-Adaboost classification result 3	-1	-1	-1	-1	1

TABLE 2: BP-Adaboost multiclassification results.

Sample	a	b	c	d	e
BP-Adaboost classification result 1	-1	1	-1	1	-1
BP-Adaboost classification result 2	-1	-1	1	1	-1
BP-Adaboost classification result 3	-1	-1	-1	-1	1
Classification result	0	1	2	0	3

2.3. *A Multiclassification Series Model for BP-Adaboost and PNN Neural Networks.* Generally, Adaboost combines with weak classifier to form a strong classifier to solve the problem of binary classification. But in transformer fault classification, there are not only two types of transformer faults. Therefore, BP-Adaboost needs to be transformed into multiclassification problems. In this paper, according to the total fault types to be classified, several BP-Adaboost two-classification models are established to classify each fault in turn. The specific classification operation is shown in Figure 3.

In Figure 3, we make the output of each BP-Adaboost model “- 1” or “1”, so that the output of “A” fault in the training sample is “1”, and the output of the other fault types is “- 1”. In this way, the test sample whose result is “1” can be classified by BP-Adaboost, and its prediction result can be regarded as “A” fault. Each fault type can be diagnosed by binary classification in this way. Although Adaboost improves the prediction results of the weak classifier by its powerful result correction ability, such a multiclassification model still has some defects. For example, the same test sample is divided into different fault models (e.g., Sample “d” in Table 1), or a test sample is not divided into any type (e.g., Sample “a” in Table 1). The occurrence of these two cases can determine that the multiclassification model can not accurately classify these samples.

Based on the possible error classification, the classification results of BP-Adaboost multiclassification model are set as follows. Firstly, the diagnosis type is coded and then the fault type is diagnosed according to the coding order. Then, each BP-Adaboost binary classification result is constructed into a $T_{m \times n}$ matrix, where m is the total fault type and n is the number of test samples. Finally, according to the position

of “1” in column i , $i = 1, 2, \dots, n$ of matrix T , the class of the i -th sample is determined. That is, the diagnostic result of sample i -th in multiclassification results is the location of “1” in i column. When there are more than one “1” in column j or no “1” in column j , this indicates that the BP-Adaboost multiclassification diagnostic model is wrong in classifying the j -th sample, so the classification result of the j -th individual in the BP-Adaboost multiclassification result is “0”. The specific operation results are shown in Table 2.

Usually we do not know the accuracy of the diagnostic results of the diagnostic model before comparing them with the real results. But for BP-Adaboost multiclassification diagnostic results in this paper, when “0” appears in the diagnostic results, it shows that the diagnosis must be wrong. In order to improve the accuracy of one algorithm, many scholars usually combine one algorithm with another to make the final experimental results better than any of them [23, 24]. In order to improve the diagnostic accuracy, a PNN neural network is connected in series after the diagnosis result of BP-Adaboost. By utilizing the recognition ability of PNN neural network, the samples of BP-Adaboost diagnosis error are rediagnosed. The algorithm diagram is shown in Figure 4. This fully combines the excellent results of BP-Adaboost multiclassification diagnosis model with those of PNN diagnosis model, so as to improve the accuracy of sample prediction.

3. Results and Discussion

3.1. *Selection of Sample Sets.* In transformer fault diagnosis, selecting representative data samples is more conducive to

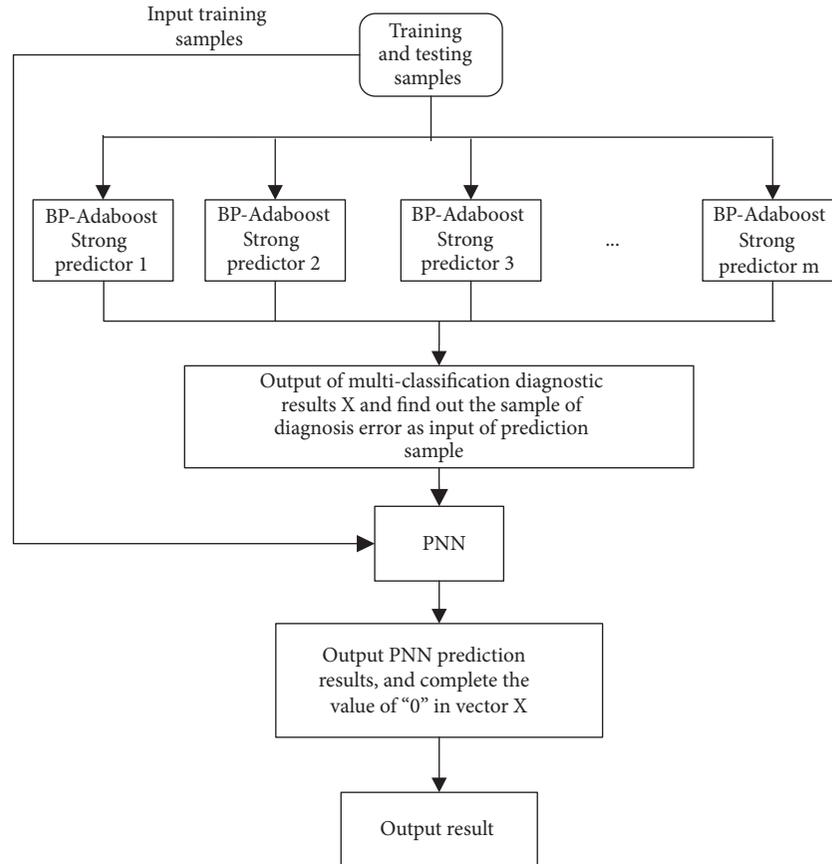


FIGURE 4: Diagnostic model of BP-Adaboost in tandem with PNN.

TABLE 3: Sample data description.

The fault types	code	Total number of samples	Number of training samples	Number of testing samples
medium and low temperature overheating	1	29	23	6
arc discharge	2	12	10	2
discharge and overheating	3	11	9	2
low energy discharge fault	4	18	14	4
high temperature overheating	5	30	24	6

the establishment of a simulation model. Therefore, the basic principles of sample selection in this paper are as follows. (1) The fault samples selected are representative. (2) The selected samples should involve more complete fault type. (3) The samples should be compact. Therefore, this paper selected 100 representative samples from the historical faults of several oil-immersed power transformers in a 220V substation for empirical analysis. Types of transformer faults include medium and low temperature overheating, arc discharge, discharge and overheating fault, low energy discharge fault, and high temperature overheating fault.

For selection of training and test sets, in this paper, 20 samples were randomly selected as test data, and the remaining 80 samples were used as training data. In order to diagnose transformer faults accurately, the test samples are randomly selected according to the proportion of different

types of samples. Specific fault codes and sample numbers are shown in Table 3.

3.2. Feature Fault Selection. Because of the difference of transformer internal faults, the gas produced by each fault is not completely the same. Principally, the fault related gases commonly used are hydrogen (H_2), carbon monoxide (CO), carbon dioxide (CO_2), methane (CH_4), acetylene (C_2H_2), ethane (C_2H_6), and ethylene (C_2H_4). Since the components of H_2 , CH_4 , C_2H_6 , C_2H_4 , and C_2H_2 are closely related to the fault types of transformers, all the five gases are taken as the characteristic parameters of transformer fault diagnosis in this paper [25]. Inspired by IEC three-ratio method, this paper also takes C_2H_2/C_2H_4 , CH_4/H_2 , and C_2H_4/C_2H_6 as the characteristic parameters of transformer fault diagnosis.

TABLE 4: Number of samples per diagnostic error.

Number of diagnostic error samples											Average error rate	Average running time
BP-Adaboost	4	4	2	6	4	2	5	2	4	4	18.5%	10.495s
BP	7	7	9	8	9	6	8	7	9	9	39.5%	0.6159s
PNN	5	5	8	5	9	5	7	1	6	5	28%	0.1877s
GA-BP	2	0	3	4	4	1	3	3	3	2	12.5%	723.52s
BP-Adaboost_PNN	1	3	1	4	4	0	2	2	2	2	10.5%	11.253s

3.3. *Parameter Setting and Running Environment.* In the series model of BP-Adaboost and PNN, the number of BP neural networks with weak classifiers is 20. The training target of each BP neural network is 0.00004, the learning rate is 0.1, and the training time is 5. The selection of SPREAD in PNN neural network is 1.5. In the diagnosis of BP neural network, the training target is 0.01, the learning rate is 0.1, and the training time is 1000. In traditional genetic algorithm optimization of BP neural network (GA-BP), the population is 20 and the number of iterations is 50 (Testing environment: Core i5-3230M dual-core processor, running in the 2016a version of MATLAB).

3.4. *Comparative Analysis of Prediction Examples.* Eight variables are selected as input vectors of the model, and different output vectors are set according to different BP-Adaboost binary classification models. According to Table 3, the number of samples of each type is randomly selected as the training data and test data of five models: BP neural network, GA-BP neural network, BP-Adaboost model, PNN neural network, and the diagnostic model proposed in this paper. The test samples are input into the five models after training, and the corresponding prediction results are obtained. Then the results are compared with the real data to evaluate the good degree of each model for transformer fault diagnosis.

Because the test samples produced in each experiment are random, this paper tests the five models 10 times, and the test samples of the five models are the same as the training samples. The error results and the average running time of the four models tested 10 times are shown in Table 4.

From Table 4, we can see that the test results of BP-Adaboost are superior to BP neural network almost every time, so it can be effectively explained that Adaboost can combine weak classifiers into a strong classifier. In addition to the proposed BP-Adaboost and PNN series model in transformer fault diagnosis, each test result is better than the BP-Adaboost and PNN test results, which show that the proposed series method effectively combines the advantages of BP-Adaboost and PNN model recognition. Although the diagnostic accuracy of BP-Adaboost and PNN series model is similar to that of GA-BP model, the time spent by GA-BP greatly exceeds that of BP-Adaboost and PNN series model.

In order to illustrate the validity of the model proposed in this paper, the results of one of the tests are taken as an example to analyze the effectiveness of the proposed model. Figures 5, 6 and 7 are diagrams comparing the diagnostic results of BP neural network, PNN neural network, and GA-BP neural network with the real results. It can be seen from the figure that the diagnostic accuracy of BP is only 65%, and

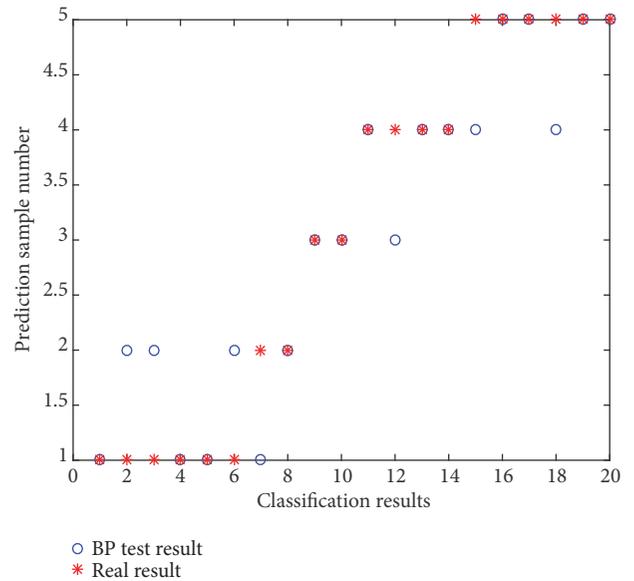


FIGURE 5: Diagnosis result of BP neural network.

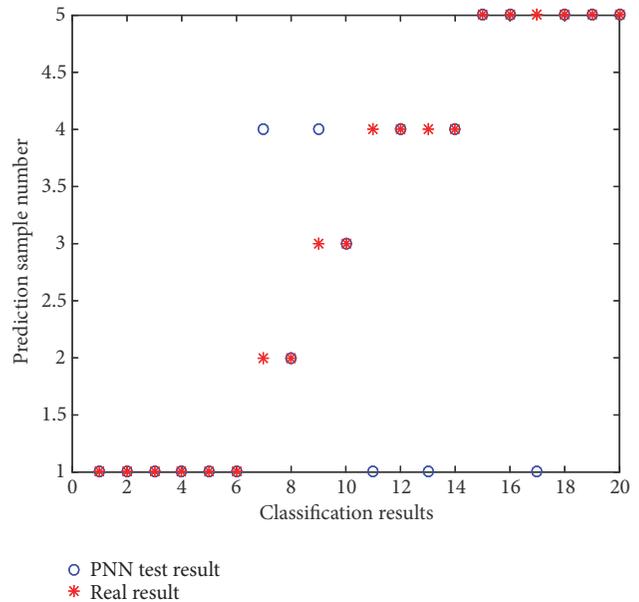


FIGURE 6: Diagnosis result of PNN neural network.

the diagnostic accuracy of PNN is relatively high, but only 75%; the diagnostic accuracy of GA-BP model is 90%.

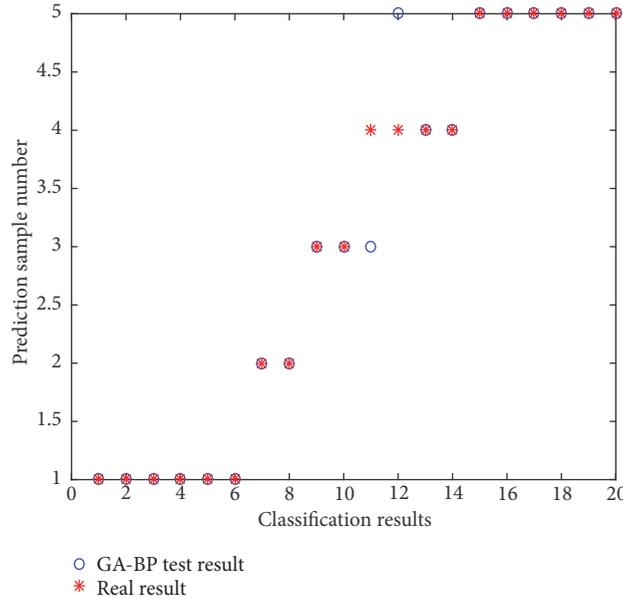


FIGURE 7: Diagnosis result of GA-BP neural network.

BP-Adaboost Output Matrix T

$$\begin{matrix}
 \text{samples} \\
 \text{BP-Adaboost1} \\
 \text{BP-Adaboost2} \\
 \text{BP-Adaboost3} \\
 \text{BP-Adaboost4} \\
 \text{BP-Adaboost5}
 \end{matrix}
 T = \begin{bmatrix}
 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
 -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1
 \end{bmatrix} \quad (11)$$

BP-Adaboost Output Matrix T Converted to Vector

$$X = [1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 2 \ 2 \ 0 \ 3 \ 4 \ 0 \ 4 \ 4 \ 5 \ 5 \ 5 \ 5 \ 5] \quad (12)$$

Equation (12) is the diagnostic result X of BP-Adaboost. X is transformed from the BP-Adaboost output matrix T of (11) according to Table 2. The diagnostic result “0” in X is the undiagnosed sample of BP-Adaboost diagnostic model. From (12), we can see that the output results by 2, 6, 9, and 12 in the test samples are 0. Therefore, it is shown that the four samples in the BP-Adaboost multiclassification result have certain classification errors. These four wrong classification samples are used as new test samples and input into PNN neural network. Through Figure 6, we can see that the prediction accuracy of PNN for 2,6,9,12 samples is 75%. The predicted results of PNN for these four samples are put into the corresponding position of the predicted results in BP-Adaboost, and the final diagnostic results of BP-Adaboost and PNN in series are obtained. The comparison between the diagnostic results and the real results is shown in (12). From Figure 8, we can see that the diagnostic accuracy of BP-Adaboost in series with PNN is as high as 95%, which is

obviously higher than that of BP-Adaboost, GA-BP, and PNN models.

In this paper, BP neural network and improved Adaboost are used to form a strong classifier. At the same time, several BP-Adaboost binary classifiers are established to form a multiclassifier. For the result matrix T formed by multiple binary classifiers, we can directly find some samples of classification errors and then put these samples of diagnosis errors into PNN neural network for rediagnosis. The reason why the proposed method can combine the diagnostic advantages of BP-Adaboost model and PNN model is that in BP-Adaboost multiclassification recognition, the classification accuracy of those samples which are only classified into one category is relatively high. Because it not only requires such samples to be classified into one type, but also ensures that the samples are not classified into other types. Under such stringent requirements, the classification accuracy of samples classified into only one category in BP-Adaboost is relatively high. Finally, the experimental results show that the diagnostic accuracy of the proposed series model in transformer fault diagnosis is significantly higher than that of BP neural network model, BP-Adaboost model, and PNN model. Although the accuracy is slightly better than that of GA-BP model, the diagnostic

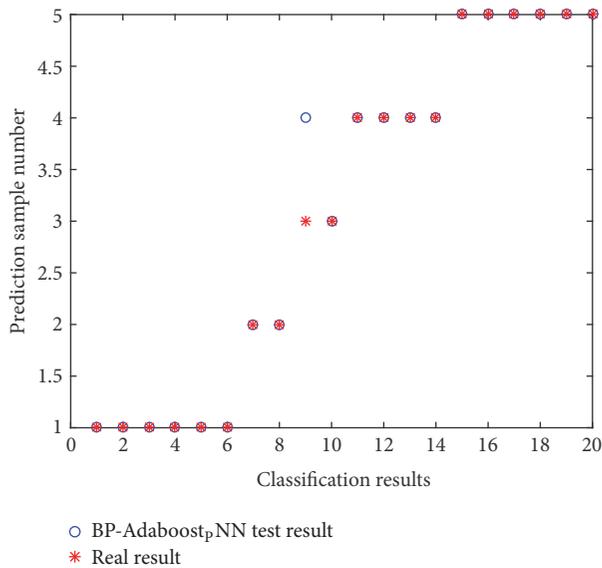


FIGURE 8: Diagnostic results of BP-Adaboost and PNN tandem model.

time of the proposed series model is obviously better than that of GA-BP model.

4. Conclusions

Power transformer plays an important role in power transmission and distribution. The performance of power transformer directly affects the operation of the whole power system. Therefore, it is very important to discover the faults of the transformer in advance. Whether early fault of the transformer can be eliminated as soon as possible is the key to ensure the stability of the power supply for users.

This paper presents a new diagnostic model of BP-Adaboost in series with PNN. By transforming BP-Adaboost biclassification model into a multiclassification model, on the basis of advantages of BP-Adaboost biclassification model, it also provides double guarantee for accurate classification samples of BP-Adaboost multiclassification model. For the type to be diagnosed, it is necessary to satisfy the need not to reclassify the classified samples into this type, but also to satisfy that the samples belonging to this type are diagnosed. Obviously, this is not always possible, which has a very high recognition accuracy for the two-class BP-Adaboost. Therefore, this paper transforms the result matrix T into vector X and then finds out the samples which have not diagnosed the fault type by vector X and puts them into PNN for further diagnosis. By connecting BP-Adaboost with PNN in series, we cannot only improve the defect of the BP-Adaboost algorithm that does not diagnose samples, but also improve the defect that the diagnostic accuracy of the PNN model is not very high.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was financially supported by the Project of National Natural Science Foundation of China (nos. 61502280, 61472228).

References

- [1] B. S. Li, Y. Guo, and S. Zhao, "Feasibility analysis of transformer winding DC resistance measuring on-line," in *Proceedings of the International Conference on Systems & Informatics*, pp. 568–571, IEEE, Yantai, China, 2012.
- [2] C. Xiang, Z. Huang, J. Li, Q. Zhou, and W. Yao, "Graphic approaches for faults diagnosis for Camellia insulating liquid filled transformers based on dissolved gas analysis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 5, pp. 1897–1903, 2018.
- [3] D. Susa, J. Palola, M. Lehtonen, and M. Hyvarinen, "Temperature Rises in an OFAF Transformer at OFAN Cooling Mode in Service," *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2517–2525, 2005.
- [4] T. Boczar, A. Cichon, and S. Borucki, "Diagnostic expert system of transformer insulation systems using the acoustic emission method," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 21, no. 2, pp. 854–865, 2014.
- [5] Z. P. Tang, M. F. Pen, G. M. L. X. Wan, and R. S. Liu, "Inter-turn short circuit fault diagnosis of transformer windings based on repetitive pulse method," *Power Automation equipment*, vol. 38, no. 10, pp. 153–158, 2018.
- [6] K. Wang, J. Z. Li, and S. Q. Zhang, "A new characteristic parameter of dissolved gas in oil for transformer fault diagnosis," *Chinese Journal of Electrical Engineering*, vol. 36, pp. 6570–6578, 2016.
- [7] R. L. Tian, J. L. Chai et al., "Power transformer fault detection based on correlated gas and genetic algorithm optimization support vector machine," *Electrical applications*, vol. 19, pp. 40–45, 2018.
- [8] R. R. Zhen, J. Y. Zhao et al., "Power transformer fault diagnosis based on genetic support vector machine and grey artificial immune algorithm," *Chinese Journal of Electrical Engineering*, vol. 21, no. 7, pp. 56–63, 2011.
- [9] S. Seifeddine, B. Khmais, and C. Abdelkader, "Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network," in *Proceedings of the 1st International Conference on Renewable Energies and Vehicular Technology, REVET 2012*, pp. 230–236, IEEE, Hammamet, Tunisia, March 2012.
- [10] G. M. Fa, Y. N. Liu et al., "Research on fault diagnosis of power transformer based on improved bat algorithm optimization BP neural network," *Journal of Shandong University of Science and Technology (Natural Science Edition)*, vol. 36, no. 3, pp. 70–74, 2017.
- [11] M.-C. Nitu, A.-M. Aciu, C.-I. Nicola et al., "Power transformer fault diagnosis using fuzzy logic technique based on dissolved gas analysis and furan analysis," in *Proceedings of the 2017 International Conference on Optimization of Electrical and*

- Electronic Equipment (OPTIM) & 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP)*, pp. 184–189, IEEE, Brasov, Romania, May 2017.
- [12] N. A. Setiawan, Sarjiya, and Z. Adhiarga, “Power transformer incipient faults diagnosis using dissolved gas analysis and rough set,” in *Proceedings of the 2012 IEEE International Conference on Condition Monitoring and Diagnosis, CMD 2012*, pp. 950–953, IEEE, Bali, Indonesia, September 2012.
- [13] W. X. Du, X. P. Zhao et al., “Fault diagnosis of power transformer based on extreme learning machine,” *Journal of Shandong University of Science and Technology (Natural Science Edition)*, vol. 36, no. 5, pp. 29–36, 2017.
- [14] Y. H. Wu and Q. Y. Xu, “Optimization of support vector machine fault diagnosis based on improved swarm algorithm,” *Mechanical strength*, vol. 40, no. 2, pp. 287–292, 2018.
- [15] J. I. Aizpurua, V. M. Catterson, and B. G. Stewart, “Power transformer dissolved gas analysis through Bayesian networks and hypothesis testing,” *IEEE Transactions on Dielectrics Electrical Insulation*, vol. 25, no. 2, pp. 494–506, 2018.
- [16] Z. Yin and J. Hou, “Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes,” *Neurocomputing*, vol. 58, no. 7, pp. 643–650, 2016.
- [17] F. Shi, H. Wang, L. Yu et al., “MATLAB intelligent algorithm 30 case analysis,” *Beijing University of Aeronautics and Astronautics*, p. 27, 2011.
- [18] Y. J. Liang, R. Chao, H. Y. Wang et al., “Research on soil moisture inversion method based on GA-BP neural network model,” *International Journal of Remote Sensing*, pp. 1–17, 2018.
- [19] N. Salman, A. Lawi, and S. Syarif, “Artificial neural network backpropagation with particle swarm optimization for crude palm oil price prediction,” *Journal of Physics Conference Series*, vol. 1114, no. 1, pp. 1–8, 2018.
- [20] Y. Kuang, R. Singh, S. Singh, and S. P. Singh, “A novel macroeconomic forecasting model based on revised multimedia assisted BP neural network model and ant Colony algorithm,” *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18749–18770, 2017.
- [21] Y. Li, C. Yan, W. Liu, and M. Li, “A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification,” *Applied Soft Computing*, vol. 70, pp. 1000–1009, 2018.
- [22] B. Li, X. D. Zhang, and H. Fang, “An improved BP-Adaboost algorithm and its application in radar multi-target classification,” *Journal of Nanjing University (Natural Science)*, vol. 53, no. 5, pp. 984–989, 2017.
- [23] C. Yan, L. Wang, W. Liu et al., “Financial early warning of nonlife insurance company based on RBF neural network optimized by genetic algorithm,” *Concurrency & Computation Practice & Experience*, vol. 30, no. 6, article no. e4343, 2017.
- [24] D. M. Gilbert, “Making sense of eukaryotic DNA replication origins,” *Science*, vol. 294, no. 5540, pp. 96–100, 2001.
- [25] B. Y. Song, J. W. Xu et al., “Multi-level inverter fault diagnosis based on WPT-PCA-NN algorithm,” *Journal of Shandong University of Science and Technology (Natural Science Edition)*, vol. 38, no. 1, pp. 111–120, 2019.

