



Transformer Fault Diagnosis Based on BP Neural Network by Improved Apriori Algorithm

Chang Guoxiang¹ , Gao Qiaoli¹  , Gao Xinming² , and Cheng Juntao¹ 

¹ Heilongjiang University of Science and Technology, Heilongjiang, Harbin, China
270080877@qq.com, 763790270@qq.com, yj74615@qq.com

² Shandong Transport Vocational College, Weifang, Shandong, China
287743080@qq.com

Abstract. With the continuous expansion of power system, it is increasing that the fault rate of transformer equipment in power system. Through the fault diagnosis technology, it can be found that the transformer fault in advance, the accident rate is can taken to reduce by measures in time, so the high accuracy of transformer fault diagnosis is required. In this paper, the BP neural network based on the optimized Apriori algorithm which is used to diagnose the transformer fault. It is found that the Apriori algorithm reveals association rules by mining the high frequency term of feature data. The data is directly analyzed and inferred to achieve the purpose of simplifying data association by rough set algorithm. Apriori algorithm combines the rough set is used to accurately mine the confidence of association rules, which is used as the weight of BP neural network link, it is simplifying the complexity of data training. Through the simulation experiment, this new method is compared with the traditional BP neural network method, it has the advantages of high accuracy and fast speed of fault diagnosis, and has certain practical value.

Keywords: Apriori algorithm · BP neural network · Transformer · Fault diagnosis

1 Introduction

Power transformer is very important for power system. In recent years, with the continuous improvement of science and technology, the power system is also developing towards the direction of high voltage and ultra-high voltage. Once the transformer breaks down, it often causes serious accidents, so it is very import that the safe and stable work of power transformer to the production and people's life. The fault of power transformer can be extracted and diagnosed, and effective measures can avoid accidents, and save the time and life, which requires a more accurate and rapid transformer fault diagnosis technology [1]. Now, the three ratio method is widely used in all transformer fault diagnosis methods, which has the characteristics of simple principle and simple calculation. But in practical application, there are some problems such as incomplete coding, too absolute coding boundary and limited multi fault judgment ability. In order to improve

the accuracy of fault diagnosis, neural network technology is applied to the research of power transformer fault diagnosis, and good results are achieved. BP neural network can store and train many input-output mapping relationships without knowing how to describe them [2]. It has the ability of nonlinear mapping, self-learning and good fault tolerance. In this paper, the improved association rules and BP neural network are used to diagnose the transformer fault. The improved association rules method is to combine the rough algorithm with the association rules. First, the rough set of data is reduced, and then the data is associated. BP neural network is used to train and learn the processed data, predict the operation state of transformer, simplify the operation of transformer, and the complexity of data calculation improves the accuracy and speed of fault diagnosis [3].

2 BP Neural Network

BP neural network is a multilayer feedforward neural network which is trained by error back propagation algorithm. It is an effective classification and recognition tool, which can classify the fault types of transformer. The rule of BP neural network algorithm is to modify the weight and threshold of neural network by error back propagation, and use the method of the fastest descent, so as to reduce the square sum of the gap between the ideal value and the actual value [4]. The topological structure of BP neural network model is shown in Fig. 1, including input layer, hidden layer and output layer.

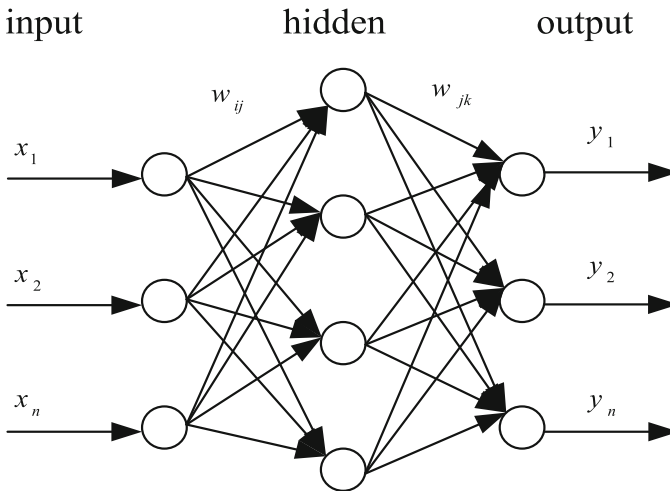


Fig. 1. BP neural network model

BP algorithm is a kind of back propagation algorithm according to the error, and its learning process is the process of constantly adjusting the weight.

The relationship between input layer vector and hidden layer vector is as follows:

$$e_j = f\left(\sum_{i=0}^n w_{ij}x_i\right) \quad (1)$$

1) The value of the output layer is:

$$e_k = f_x\left(\sum_{j=0}^n w_{jk} o_j\right) \quad (2)$$

2) The actual output value is:

$$o_k = f_x(e_k) \quad (3)$$

3) Mapping relation:

$$f(x) = 1/(1 - e^{-x}) \quad (4)$$

There is an error:

$$E_P = \frac{1}{2} \sum_k (e_{pk} - o_{pk})^2 \quad (5)$$

The square difference formula is:

$$E = \frac{1}{2P} \sum_P \sum_K (e_{pk} - o_{pk})^2 \quad (6)$$

3 Association Rules and Apriori Algorithm

3.1 Association Rules

Association rule is an important data mining method and an unsupervised machine learning method. It is used to establish the association between projects, study the correlation between projects, and give opinions according to the correlation analysis. Apriori algorithm is the most widely used one.

Association rule definition: set sets $I = \{i_1, i_2, i_3, \dots, i_n\}$, the data set is: $D = \{T_1, T_2, \dots, T_m\}$.

The subset of itemset I is the corresponding transaction set t , and A is a itemset, if $A \in T$, then $A \Rightarrow B$, $A \in I$, $B \in I$, $A \cap B = \Phi$, Indicates that if item set A appears in a transaction, y will also appear with a certain probability.

N is the number of transactions. In association rules, support and confidence are the two most commonly used standards. If $\text{count}(A \subseteq T)$ is the number of transactions containing x in transaction set D , then the support of item set A is:

$$\text{Support}(A) = \frac{\text{count}(A \in T)}{N} \quad (7)$$

$$\text{Support}(A \Rightarrow B) = \frac{\text{count}(A \cup B)}{N} \quad (8)$$

The confidence formula is:

$$\text{Confidence}(A \Rightarrow B) = \frac{\text{Support}(A \Rightarrow B)}{\text{Support}(A)} \quad (9)$$

3.2 Apriori Algorithm

If an itemset is frequent, then all subsets of it are frequent. Conversely, if an item set is infrequent and all of its supersets are infrequent, then the entire set of supersets containing infrequent sets can be immediately cut off. This strategy is called support-based pruning.

The basic working idea of the prior algorithm is to retrieve the entire database, calculate the number of each item, and generate the item with the minimum support, which is represented by L_1 . On this basis, the second frequent item is distributed L_2 , and continue to execute until the K frequent item set cannot be found. Prior algorithm flow:

Step 1: It is got all the frequent itemsets in the transaction database are implemented through iterative operations, that is, the itemsets whose support is no less than the threshold set by the user [5];

Step 2: It is to use frequent item sets to construct rules that satisfy the minimum trust of users.

The specific method is: first of all, it is been found out the frequent 1-term set, and record it as L_1 ; then use L_1 to generate candidate item set C_2 , determine the items in C_2 , and mine L_2 , i.e. frequent 2-term set; keep cycling until no more frequent k -term set can be found [6]. That is to say, when generating a k -i candidate, if a subset of the candidate is not in $(k-1) - i$ (which has been determined to be frequent), then the candidate will be deleted without taking it out and judging the support [7]. However, the traditional Apriori algorithm has two defects:

- 1) A great number of candidate sets are generated in the process of algorithm, which takes up a large amount of memory;
- 2) Every calculation needs to scan the whole database, so the algorithm has a high time and complexity.

4 Rough Set

4.1 Definition of Rough Set

- 1) Information system: In general, a knowledge expression system or the information system can be expressed as, in which, it is a domain, which is a collection of all samples; it is an attribute collection, in which the subset is a conditional attribute set, which reflects the characteristics of the object, D is a decision attribute set, which reflects the class of the object; it is an attribute set, which represents the value range of attribute R ; it is an information function, which is used to determine each one in U Property values of objects, that is, any.
- 2) Indiscernible relation: when two objects are described by the same attribute, they are classified into the same category in the system. Their relation is called indiscernible relation, that is, for any attribute subset, if the object, if and only then, and is indiscernible relation, the indiscernible relation is called equivalent relation for short.

- 3) Upper approximation set and lower approximation set: the lower approximation set is defined as the set composed of all objects that must belong to A in U according to the existing knowledge R , i.e. in the formula, the equivalence relation R contains the equivalence class of relation A ; the upper approximation set is defined as the set composed of objects that must belong to and may belong to X in U according to the existing knowledge R .

The main characteristics of rough set are: it does not need to provide subjective evaluation of knowledge or data, only based on the observation data, it can delete redundant information, compare incomplete knowledge roughness, defining dependencies and importance between attributes.

Main research fields: attribute reduction, rule acquisition, and intelligent algorithm based on rough set.

4.2 The Principle of Rough Sets

The function of rough set theory is very powerful, especially for the processing of uncertain information, it has its own congenital advantages. There are two differences between rough set theory and other uncertain and imprecise theories: first, it is not needed to provide additional hints beyond the data set required for the problem; second, it has strong applicability and is easy to apply to any field. Because the purpose and starting point of rough set theory is to analyze and reason data directly. The fuzzy theory and evidence theory Compared with the methods of dealing with uncertainty problems, (The most obvious advantage of this algorithm is that it does not need any information other than data as supplement and support, and it has strong complementarity with the theories dealing with other uncertainty problems (especially fuzzy theory) [8]. In this paper, the optimized Apriori algorithm is applied to transformer fault diagnosis.

4.3 Advantages of Rough Set

The main feature of rough set: the rough set is that it does not need to provide subjective evaluation of knowledge or data, but can remove redundant information, compare the incomplete knowledge roughness and define the relevance and importance of the attributes according to the observed data.

- 1) It can handle all the data which is included incomplete data and multivariable data;
- 2) It can deal with the uncertainty and ambiguity of data, including the situation of certainty and uncertainty;
- 3) It can get the minimum expression of knowledge and different particle levels of knowledge;
- 4) It can reveal the simple concept and easy operation mode from the data;
- 5) It can generate rules with high accuracy and easy to verify. It is especially suitable for intelligent control and can generate rules automatically.

4.4 The Function of the Rough Set

Rough set theory is greatly used in data mining, involving medical research, market analysis, business risk prediction, meteorology, speech recognition, engineering design, etc. In many data mining systems, the role of rough set theory mainly focuses on the following aspects:

1) Data reduction

Rough set algorithm has a strong ability in data reduction. In the preprocessing stage of data mining system, redundant information (object, attribute value, etc.) is deleted by rough set theory. In the data mining system, which can greatly improve the running speed of the system. [2].

2) Rule extraction

The feature of rough set theory algorithm is simple and direct, and its association rules are one-to-one correspondence. The general steps of generating rules in rough set method are:

- (1) It can be got a reduction of condition attributes and deleting redundant attributes;
- (2) It can be deleted redundant attribute values of each rule;
- (3) It can merged the remaining rules.

3) Incremental algorithm

In the face of large-scale and high-dimensional data in data mining, it is a research hotspot to find an effective incremental algorithm.

4) Integration with other methods

The combination of rough set theory and neural network method can show their advantages to the greatest extent and improve the speed and accuracy of data operation.

5 Improved Association Rule Algorithm

5.1 Combination of Rough Set and Association Rules

According to the research on the characteristics of traditional Apriori algorithm, it is found that the traditional Apriori algorithm has its own defects, this paper combines the rough set algorithm with the Apriori algorithm to get an improved Apriori algorithm, in order to obtain a better speed effect of transformer fault detection. The basic idea of the improved Apriori algorithm is: firstly, the rough set algorithm is used to process the original data to achieve the purpose of data reduction, and then the Apriori algorithm is used to analyze the reduced data and generate frequent item sets and association rules. The transformer fault classification is realized. The specific process is as follows

1) Select attribute to generate decision table

Select the appropriate commonality as the commonality and decision attribute of the project, and generate the decision table.

2) Attribute reduction

The method based on discernibility matrix is applied to attribute reduction, conditional attribute reduction and kernel calculation. When there are multiple kernel values, the minimum reduction set containing kernel values is selected. After reduction, we get the reduction set of decision table, and use the reduction set to transform the decision table into a new decision table with Boolean generality suitable for Apriori analysis.

3) Use set operation instead of scanning attribute set

Apriori algorithm is used to calculate and generate the frequent set of one item on the new decision table. Delete the items whose support is less than the threshold value, expressed as C_1 . The frequent item set is generated on the basis of C'_1 , $C'_i = C_{i-1} * C_i$, i is greater than 1, less than or equal to the length of the reduction set, and the items with support less than the threshold in C'_i are deleted.

- 4) For decision attribute set D'_j , $D'_j = D_{j-1} * D_j$ j is greater than 1 and less than or equal to the length of decision attribute set calculated according to D_1 ;
- 5) C'_i and D'_j generate frequent itemsets with length of $i + j$ through ptwojoin operation with attribute 1, expressed as L_{i+j} ;
- 6) in L_{i+j} , frequent item sets with support and trust greater than or equal to threshold can be put into rule set R.

In the above calculation steps, C'_i , D'_j and L_{i+j} can be used for each calculation according to attribute 1 to avoid rescanning the decision table (Fig. 2).

6 Example Analysis

The gas composition in the transformer oil is captured, and the gas detection data are obtained. After the gas composition analysis, five characteristic gases CH_4 , H_2 , C_2H_2 , C_2H_4 and C_2H_6 , which can be obtained. The content percentage of these five gases is taken as the input quantity, and the output results are the five states of normal, low energy discharge, high energy discharge, medium and low temperature overheating and high temperature overheating. The collected data is shown in Table 2. The improved Apriori algorithm is used to analyze and simplify the association rules between features, and the BP neural network method is selected for training and experiment. The simulation results in this paper are compared with the traditional algorithm diagnosis results to verify the effectiveness of the proposed method. The input of the neural network takes the percentage of the volume of petroleum dissolved in the five gases during synthesis, so as to determine that the input layer node is $a = 5$. The operation states of the above

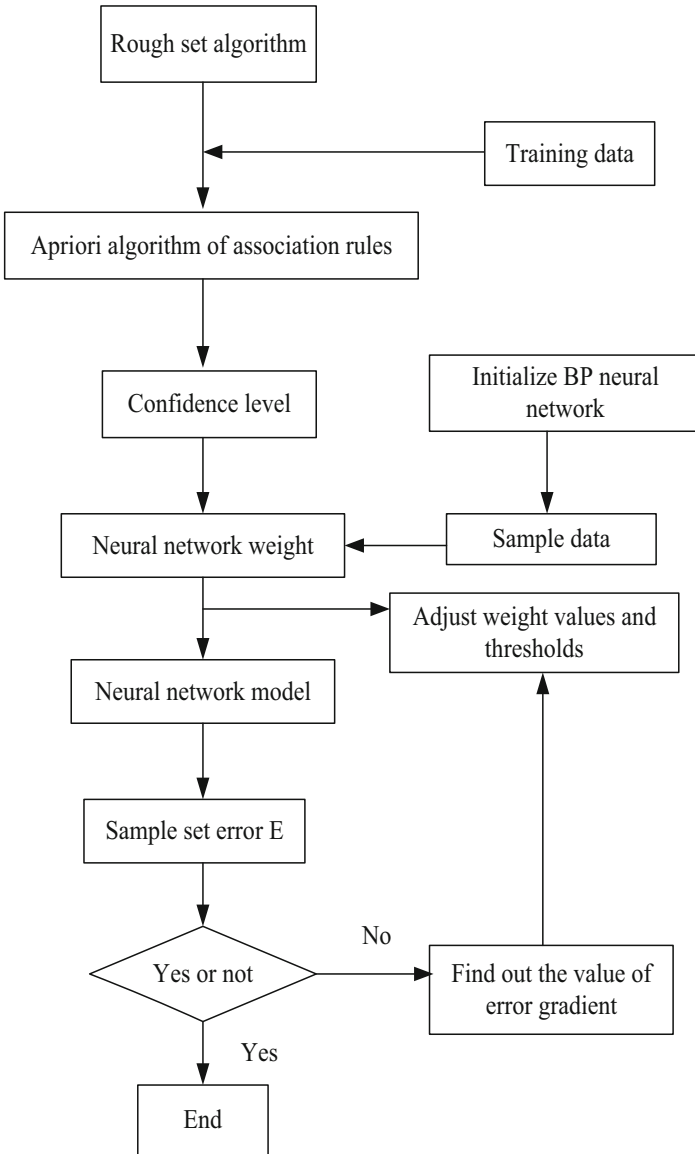


Fig. 2. Fault diagnosis flow chart of BP neural network based on improved Apriori algorithm

five transformers are taken as the output of neural network, the output layer node number $b = 5$ is determined. In general, the number of hidden layer neural nodes is selected according to the empirical formula:

$$h = \sqrt{(a + b)} + \alpha \tag{10}$$

Where α is any number between 1 and 10. According to experience, $\alpha = 2$, so $t h = 5$.

In order to solve the problem of complex calculation and slow speed, this paper uses a BP neural network based on optimal association rules to diagnose the transformer fault. Let the input value be expressed as A, and the transformer state as B, that is, since the confidence value represents the association rule $A \Rightarrow B$, It contains the conditional probability that B exists when A exists, and this condition reflects the mapping relationship between the content of dissolved gas in transformer oil and the state of transformer. According to the comparison of the original weights obtained by the traditional association rules, the weight matrix obtained by the optimized association rules is more suitable for the initial weights of BP neural network. Therefore, in this paper, the improved Apriori algorithm is used to process the data, and the initial weights between the input layer and the hidden layer of BP neural network, and between the hidden layer and the output layer, ω_n is obtained.

$$w_n = \frac{Confidence_n}{Confidence_1 + Confidence_2 + \dots + Confidence_5} \quad (11)$$

There: w_n is the weight of the transformer when the content of n-dissolved gas exceeds the standard; Confidence is the confidence of the transformer when the content of n-dissolved gas exceeds the standard.

To verify the effectiveness of this method for sample data processing and classification, 150 groups of DGA data that have been identified fault types are collected for analysis, and divided into training samples and test samples in a ratio of 2:1. Among them, 100 groups of training sample data are used to optimize the Apriori algorithm to optimize the weight and threshold parameters of BP neural network, and 50 groups of test sample data are used to verify the diagnosis accuracy of the optimized parameter model. See Table 2 for the specific distribution of transformer fault sample data. Table 3 shows the DGA data of 10 groups of test samples. In this paper, The oil content and operation state data of 150 transformers in a certain area of Southwest China are selected as the test data set: 5 fault characteristic gases were marked respectively. H_2 exceeds the standard, CH_4 exceeds the standard, C_2H_2 exceeds the standard, C_2H_4 exceeds the standard, C_2H_6 exceeds the standard, five operation states of transformer are marked respectively, the statistical conditions of various gases exceeding the standard and transformer operation are shown in Table 1. The gas frequency for standard or transformer operating conditions is shown in brackets:

It is used to test the transformer fault diagnosis performance of BP neural network based on the improved Apriori algorithm, Matlab is used to write the fault diagnosis program, and the improved Analysis of gas content in transformer oil by Apriori algorithm, so as to obtain the confidence degree of the input quantity and the output state. As the weight coefficient of BP neural network, the average detection accuracy of each gas is counted. The initial weights are trained by BP neural network. The diagnosis speed and accuracy results are compared and shown in Table 4.

Table 1. Fault status table

No.	Fault Type	Fault output status
1	Normal	Y ₁
2	Low temperature overheating	Y ₂
3	High temperature overheating	Y ₃
4	Low energy discharge	Y ₄
5	High-energy discharge	Y ₅

Table 2. Sample data of transformer fault

Fault Type	Training Sample	Test Sample
Normal	20	10
Low temperature overheating	18	9
High temperature overheating	22	11
Low energy discharge	18	9
High-energy discharge	22	11
All	100	50

It can be seen from Table 4 that the average detection accuracy of the improved Apriori algorithm for transformer fault diagnosis is 86.3%, the operation speed is 1.365 s, the average detection accuracy of BP neural network is 77.6%, and the operation speed is 0.892 s. The improved Apriori algorithm greatly reduces the detection error rate, improves the detection speed, overcomes some problems existing in the current transformer fault diagnosis process, and reduces the harm caused by the transformer fault. This method has certain advantages.

Table 3. Sample data

No.	H ₂	CH ₄	C ₂ H ₂	C ₂ H ₄	C ₂ H ₆	Fault condition
1	90.7	23	9.2	9.0	5.1	Y ₅
2	172	26	14.5	11.6	6.0	Y ₅
3	159	25.3	15.6	12.4	5.4	Y ₅
4	59.0	45.0	23.4	20.7	11.2	Y ₄
5	86.0	110	86.0	45.9	66.8	Y ₄
6	93.0	58.2	0.1	37.2	45	Y ₂
7	68.5	58.9	1.9	59.7	70.3	Y ₂
8	53.6	70.2	4.5	17.0	158.0	Y ₂
9	183	291.0	0.33	341.9	109.8	Y ₃
10	242.8	76.2	12.5	173.2	76.5	Y ₃

Table 4. Analysis of diagnosis results

Methods	Classification accuracy	Running time
BP neural network	77.6%	1.365
BP neural network with improved association rules	86.3%	0.892

7 Conclusion

Traditional BP neural network algorithm has a wide range of applications, but there are also some defects and deficiencies. After the simulation experiment analysis, the traditional BP neural network and the BP neural network optimization association rule algorithm diagnosis results are compared, it is found that the optimization association rule BP neural network algorithm can effectively improve the accuracy of transformer fault diagnosis, shorten the running time, this optimization algorithm has a certain guiding value for the research of transformer fault diagnosis technology.

References

1. Long, J., et al.: Research on transformer fault diagnosis based on BP neural network improved by association rules. In: 2019 2nd International Conference on Electrical Materials and Power Equipment (ICEMPE) (2019)
2. Lv, C.T.: Implementation and application of clustering algorithm based on k-frequent association rules in embedded system. Yunnan university (2013)
3. Liu, Y., Lipo, W., Zhao, L., Yu, Z., (eds.) *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*. Springer Science and Business Media LLC, Heidelberg (2020). <https://doi.org/10.1007/978-3-030-32456-8>
4. Zheng, W., Shiguan, G.: Research on group purchase recommendation method based on apriori-bp algorithm. *Wirel. Internet Technol.* **15**(23), 82-84 + 92 (2008)
5. Xiang, J.: Association rule mining of glacier catalog data based on rough set and Apriori algorithm. *Mod. Comput. (professional edition)* **2009**(11), 34–37 (2009)
6. Kai, S., Peng, C.: Association rule mining and screening of experimental data on glucose tolerance based on APRIORI algorithm and interes. *Med. Inf.* **22**(08), 1397–1399 (2009)
7. Wang, M., Jiang, Y., Wang, Y., Minlett, Z.T., Zhou, Z.: Association rule mining algorithm based on rough set and single transaction item combination. *Comput. Sci.* **38**(11), 234–238 (2011)
8. Li, D.: Application of association rule mining in cross-selling of Banks. Jiangsu University of Science and Technology (2016)