Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Transformer Fault Diagnosis Based on RapidMiner and Modified ELM Algorithm

To cite this article: Ji-you Zhong *et al* 2020 *J. Phys.: Conf. Ser.* **1585** 012030

 IOP Publishing
1585 (2020) 012030 doi:10.1088/1742-6596/1585/1/012030

Transformer Fault Diagnosis Based on RapidMiner and Modified ELM Algorithm

ZHONG Ji-you[1], WEI Jin-xiao[1,2], WU Xiao-rong[1], and TANG Hao[1,2]

[1]POWERCHINA Sichuan Electric Power Engineering Co., Ltd, Chengdu 610094, China
[2]School of Electrical Engineering and Information,Sichang University ,Chengdu 610065,China

815847817@qq.com

Abstract. For transformer fault diagnosis, the three ratio method lacks of encoding, and artificial intelligence methods lack of anti-interference ability.Thus, a new method of transformer fault diagnosis based on RapidMiner and modified particle swarm optimization Extreme Learning Machine (RM-MPSO-ELM) is proposed. Firstly, RapidMiner picks out the most relevant input variables of the transformer fault. then, using the modified particle swarm algorithm to optimize the parameters for Extreme Learning Machine. Finally, using the ELM to identify the potential of transformer fault, the diagnostic performance of IEC three ratio method, support vector machine (SVM) method and different combinations of ELM algorithm are also compared. The results show that the proposed method achieves higher diagnosis precision.

1. Introduction

Dissolved Gas-in-oil Analysis(DGA)[1] is widely used in potential fault diagnosis of power transformers. However, the analysis of DGA by different methods will produce deviation of different degrees and directions. Various diagnostic methods of artificial intelligence will affect the diagnosis results due to different input variables. Therefore, the selection of input variables and the anti-interference ability of the diagnosis algorithm are particularly important.

The power transformer is critical to the reliability of power system. However, many factors can cause transformer insulation to deteriorate, and affecting normal power supply. DGA analysis is one of the topics that have received widespread attention in transformer fault diagnosis. In order to solve such problems, many artificial intelligence technologies have been applied to this field, including artificial neural networks [2], support vector machines[3], fuzzy logic algorithms[4], type II fuzzy logic algorithms[5], and the others[6~7].

This paper synthesizes the above research results, selects the main influencing factors of transformer faults, combines particle swarm optimization to optimize ELM parameters, and propose a RM-MPSO-ELM diagnosis method to obtain a more stable and accurate fault diagnosis effect.

2. Extraction of main factors

Using RapidMiner data mining software[8] and the PCA [9] operator to build a model, and ranking it based on the impact weights, then the main impact factors are extracted.

2.1. Data set and attribute selection

Journal of Physics: Conference Series **1585** (2020) 012030 doi:10.1088/1742-6596/1585/1/012030

This article uses the DGA analysis results of 223 oil-immersed transformers in a certain place, and 150 groups are randomly selected to characterize Selection and model training, another 73 groups were tested.
Integrated into the input space vector *H* can be represented as a 223 × 14 matrix:

$$H=\begin{bmatrix} CH_4/H_2, C_2H_2/C_2H_4, C_2H_4/C_2H_6, C_2H_2/CH_4, \\ C_2H_2/H_2, CO_2/CO, C_2H_4/CH_4, \%CH_4, \%C_2H_2, \\ \%C_2H_4, \%H_2, \%C_2H_6, \%CO, \%CO_2 \end{bmatrix}_{223\times 14} \tag{1}$$

Wherein,$\%CH_4=CH_4/(CH_4+C_2H_2+C_2H_4)$，$\%C_2H_2=C_2H_2/(CH_4+C_2H_2+C_2H_4)$，$\%C_2H_4=C_2H_4/(CH_4+C_2H_2+C_2H_4)$，$\%H_2=H_2/(H_2+C_2H_6+CO+CO_2)$，$\%C_2H_6=C_2H_6/(C_2H_6+CH_4+C_2H_2+C_2H_4)$，$\%CO=CO/(CO+CO_2+C_2H_6+CH_4+C_2H_2+C_2H_4)$，$\%CO_2=CO_2/(CO+CO_2+C_2H_6+CH_4+C_2H_2+C_2H_4)$。

2.2. Model selection

RapidMiner is the world's leading data mining platform, which provides a large number of modeling methods or learners.The "weight by PCA" and "select by weight" operators are selected as the main parts of the extraction model with prominent influence factors.
For the processing of missing values and extreme data, the following principles should be followed: ① set to five if the given gas attribute value is not available; ② set to five if the given gas attribute value is uncertain; ③ if the attribute value is infinite, set to Twenty; ④ Set to 0.5 when the attribute value is less than one.
After extraction, the first eight attribute values H1 ~ H8 are selected as: $\%C_2H_2$，$\%CH_4$，$\%C_2H_6$，$\%H_2$，C_2H_2/CH_4，C_2H_2/H_2，C_2H_4/C_2H_6，CH_4/H_2。

3. Fault classification using MPSO-ELM

In this paper, the ELM algorithm is used for fault classification, and the modified ELM algorithm is used to improve its generalization ability.

3.1. Modified ELM algorithm

Particle swarm optimization is a bionic algorithm in which all particles have their own adaptive values. During the iterative process, they constantly update themselves through the two indicators of individual extreme value *Pbest* and global extreme value *gbest*. Usually used for optimization of related algorithm parameters.
The particle is represented by *i*, and its velocity and position update equations are as follows:

$$V_{id}^{k+1}=\omega v_{id}^{k}+c_1 random_1^k(pbest_{id}^k-x_{id}^k)+c_2 random_2^k(gbest_{id}^k-x_{id}^k) \tag{2}$$

$$x_{id}^{k+1}=x_{id}^k+v_{id}^k \tag{3}$$

Wherein, ω is the inertia weight, ranging 0.4 ~ 1.2; c_1, c_2 is the acceleration factor, in the range of 0 to 2; x_i is position information; v_i is velocity information; Random is a random number between 0 and 1.
The basic PSO algorithm lacks consideration of local or global. In view of this, an improved particle swarm algorithm is proposed to optimize the optimal position of each particle (pbesti) are given corresponding weight coefficients, and the size of the weight coefficient is determined by the size of its adaptive value, then the global best position (gbestg) can be continuously modified by the sum of the product of each particle and its weight.
The weight coefficient of each particle is expressed by α_i, *i* = 1,2,3, ..., *m*, *j* = 1,2, ..., *N*, then:

$$\begin{cases} pbest_{gj} = \sum_{i=1}^{m} \alpha_{ij} pbest_{ij} \\ \alpha_{gj} = \dfrac{pbest_{ij}}{\sum_{i=1}^{m} pbest_{ij}} \end{cases} \quad (4)$$

The larger the fitness value of a particle, the larger its weight coefficient, and the greater its impact on the global particle swarm. The addition of a weight coefficient enables all particles to be taken into account, which improves the algorithm's ability of global optimization.

In addition, the inertia weight controls the balance of the entire search process. It is a decisive parameter for particles to maintain inertia and expand the search range. Too large or too small will affect the search process. If the value of ω is large, the global convergence will be increased, and the corresponding calculation amount will be increased, and the convergence speed will be slowed down; otherwise, the value of ω is small, the convergence speed will be accelerated, and the calculation amount will be reduced, but it may fall into a local optimum . Therefore, based on the mechanism of the ω parameter, we propose to apply the $y = \sqrt{x}, x \in [0,1]$ function to simulate the change of the value of the inertia weight ω, and define a new inertia weight ω' instead of ω in Equation 5:

$$\omega' = \omega_{\max} - \sqrt{\frac{d}{d_{\max}}}(\omega_{\max} - \omega_{\min}) \quad (5)$$

Wherein, $\omega_{\max}$ is a maximum inertia weight, $\omega_{\min}$ is the minimum inertia weight, d is defined as the current iteration, $d_{\max}$ is the maximum number of iterations.

That is to make the initial value of ω larger, and slowly decrease ω as the iteration progresses to increase the speed of convergence. In this way, the processing can make the PSO search more accurately near the optimal value, thereby improving its optimization accuracy.

3.2. ELM optimization by improved PSO algorithm

Aiming at the problem that the input weight w_i and hidden layer threshold b_i of the ELM algorithm are randomly determined, the global search capability of the improved PSO algorithm is used to select two optimal initial input weight and hidden layer threshold of the ELM. The specific steps are as follows:

1) Determine the topology and training samples, and select appropriate acceleration factors c_1 and c_2, inertia weight ω, maximum running time T, maximum number of iterations D, and population size M.

2) Calculate the particle iteration value, output weight and hidden layer threshold for the first iteration, and get the corresponding output weight matrix. Calculate the fitness of the individual individuals of the initial population by the transformer training, that is:

$$\sigma = \sqrt{\sum r_i / (n-1)} \quad (6)$$

Where r is the deviation between a set of measured values and the expected value, and the current fitness of the particle swarm algorithm is set to $F_i = \sigma$. First, compare the size with *pbest* in each iteration. If $F_i < pbest$, use F_i instead of *pbest*; otherwise . Then compare the size F_i and $gbest$, if $F_i < gbest$, replace gbest of F_i , make it the latest best position, while updating the weights, speed and position of the particle, otherwise unchanged. At the same time, when: ① the number of runs reaches the maximum iterations; ② the run time reaches the longest run time; ③ *gbest* reaches stability, one of these three conditions exits the program and returns the current optimal individual and its fitness.

4. Fault classification based on RM-MPSO-ELM

The RM-MPSO-ELM-based fault classification process is divided into three parts, namely the data preparation stage, the training stage, and the test stage. The specific execution steps are:

Step 1 Data preparation stage: Collect transformer failure samples and combine the prominent influence factor extraction in the second section to list the eight characteristic attribute values; normalize the characteristic attribute data set into two parts: training samples and test samples.

Step 2 Training stage: Determine the input weights by the MPSO algorithm, calculate the output weight matrix *H*, and input the training sample data into the ELM for training until the training error is less than 1%, when covariance is not met <1%, increase training times and adjust input weights. Then, save the trained ELM model.

Step 3 Test stage: input the test sample data into the ELM after training, output the model test results, and then output the diagnosis results in combination with Table 1 below.

The detailed fault diagnosis model is shown in Figure 1, which defines the covariance:

$$\text{MSE} = \frac{1}{n}\sum_{q=1}^{n}(T_q - OA_q)^2 \tag{7}$$

In the formula, *n* is the number of samples, T_q is the target value, and OA_q is the actual output value.

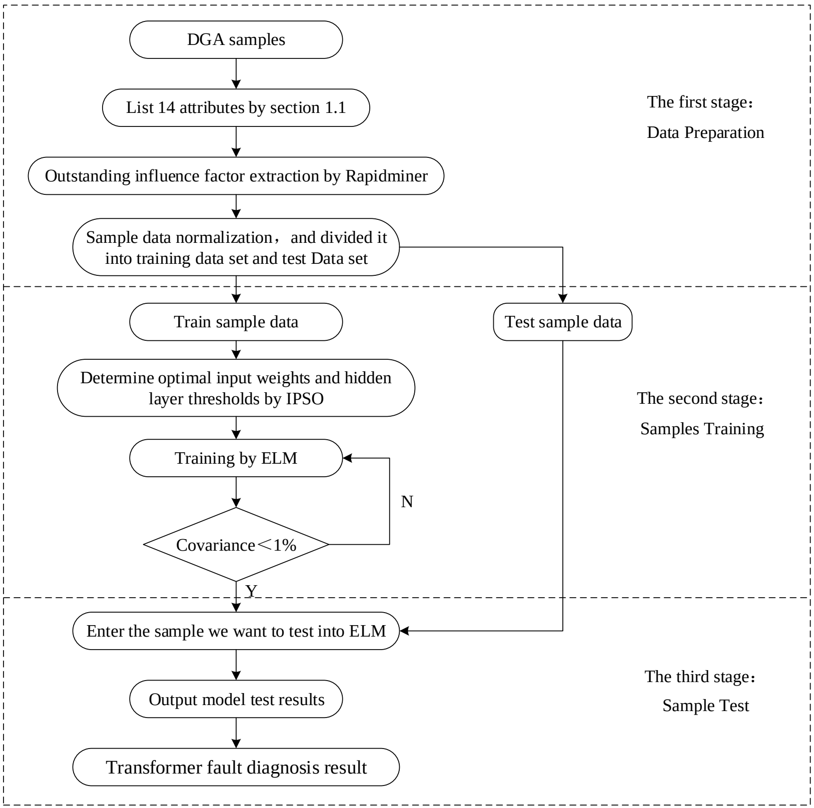

Figure 1 Fault diagnosis model

With reference to China's DL / T722-2000 "Guidelines" and the relevant provisions of IEC 60599, combined with the actual situation of transformer fault diagnosis, this article from the perspective of practicability and operability, this paper divides the transformer state into normal state and six fault states.

In the parameter selection, the population size of the particle swarm algorithm is 30, the acceleration factors c_1 and c_2 are both selected as 1.5, the maximum number of iterations is gradually increased from 30 to 800, the number of input layer nodes is selected, and the number of output layer nodes is selected as 7. The number of hidden neurons can be determined according to $L=\sqrt{j+k}+a$,where *L* is the number of hidden neurons, *j* is the number of input layer nodes,*k* is the number of output layer nodes,*a* is the population size of the particle swarm algorithm. The final number of nodes in the hidden layer is determined by trial and error. Figure 2 shows the relationship between the number of neurons and classification accuracy:

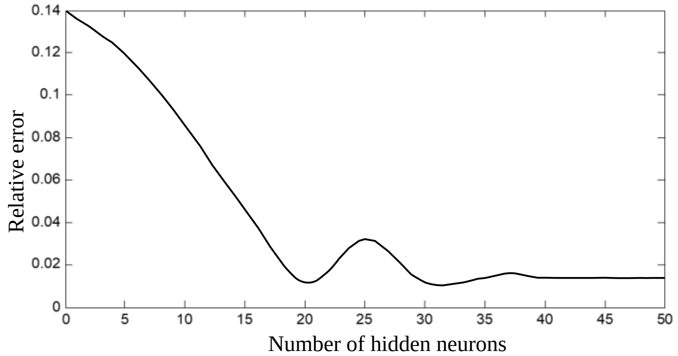

Figure 2 Relationship between neuron number and classification accuracy

With the increase of the number of hidden neurons, the classification accuracy increases first. When it increases to 32, the accuracy is relatively high, but it continues to increase, and its accuracy improvement is not obvious. Therefore, the topology is 8-32-7.

5. Case analysis

In this paper, simulation experiments are performed in MATLAB 7.0. 150 cases were randomly selected as training samples, and the rest 73 cases were used as test samples.

5.1. Case one

Construct a transformer fault diagnosis model based on RM-MPSO-ELM, MPSO-ELM, PSO-ELM, SVM, and IEC three-ratio method. Except for the RM-MPSO-ELM method, the remaining characteristic parameters are selected as CH_4, C_2H_4, C_2H_2, C_2H_6, H_2, these five parameters are closely related to the gas volume fraction of the transformer state, and then input the same sample data for training, and finally get the test results shown in Table 1.

Table 1 Comparison of diagnostic results of various methods

Fault type	sample number	Fault Diagnosis accuracy /%				
		IEC	SVM	PSO-ELM	MPSO-ELM	RM-MPSO-ELM
Normal	11	27.3	72.7	72.7	100.0	100.0
Low overheating	7	57.1	71.4	71.4	85.7	100.0
Moderate overheating	12	66.7	66.7	75.0	83.3	91.7
Severe overheating	15	73.3	80.0	86.7	86.7	93.3
Partial discharge	5	60.0	80.0	80.0	80.0	100.0
Low energy discharge	9	88.9	88.9	88.9	88.9	88.9
High energy discharge	14	64.3	78.6	85.7	85.7	92.9
Total	73	63.0	76.7	80.8	87.7	94.5

The comparison results show that the RM-MPSO-ELM diagnostic model constructed in this paper has higher diagnostic accuracy of 94.5%.

5.2. Case two

In order to further verify the anti-interference ability of RapidMiner's prominent influence factor extraction for transformer fault diagnosis, a RM-MPSO-ELM model and a MPSO-ELM model were used for comparison tests. In addition, to the volume fraction of these five gases, add interference 5%, 10%, 15%, keep all common parameters the same, the test results are shown in Table 2, MPSOELM diagnosis after adding interference The overall accuracy rate of the model decreased from 87.7% to 78.1%, 65.8%, and 56.2%, which decreased by 9.6%, 21.9%, and 31.5% respectively; while the RM-MPSO-ELM diagnostic model was affected by noise; but until it was added 15% interference, the accuracy rate only dropped by 15.1%, which is enough to prove that RapidMiner's prominent influence factor extraction has improved the role and advantages of fault diagnosis anti-interference ability.

Table 2 Comparison of diagnostic results after noise

Fault tape		Test Results		Add noise by 5%		Add noise by 10%		Add noise by 15%	
		NO RM	RM	NO RM	RM	NO RM	RM	NO RM	RM
Correct rate/%	Normal	100.0	100.0	81.8	100.0	72.7	100.0	72.7	81.8
	Low overheating	85.7	100.0	71.4	100.0	57.1	100.0	57.1	85.7
	Moderate overheating	83.3	91.7	75.0	91.7	66.7	91.7	58.3	83.3
	Severe overheating	86.7	93.3	80.0	93.3	73.3	86.7	60.0	80.0
	Partial discharge	80.0	100.0	80.0	80.0	40.0	80.0	40.0	80.0
	Low energy discharge	88.9	88.9	88.9	88.9	77.8	88.9	55.6	88.9
	High energy discharge	85.7	92.9	71.4	92.9	57.1	71.4	42.9	64.3
Total		87.7	94.5	78.1	93.2	65.8	87.7	56.2	79.5

6. Conclusion

The author applies the data mining platform RapidMiner to transformer fault diagnosis, and combines a particle swarm algorithm and ELM algorithm to propose a new diagnostic strategy, which increases the extraction steps of prominent influencing factors and improves its diagnostic anti-interference ability. The selection of relevant parameters of ELM is optimized and the diagnostic accuracy is improved.

7. References

[1] GB/T 7252-2001 Guide to the analysis and the diagnosis of gases dissolved in transformer oil [S].
[2] Saeed Madani, Seyed. Electric Load Forecasting Using an Artificial Neural Network[J]. IEEE Transactions on Power Systems, 2013, 6(2):442-449.
[3] Zhang Y, Wei H, Liao R, et al. A New Support Vector Machine Model Based on Improved Imperialist Competitive Algorithm for Fault Diagnosis of Oil-immersed Transformers[J]. Journal of Electrical Engineering & Technology, 2017, 12.
[4] A. Soltani, T. Battikh, I. Jabri,等. A new expert system based on fuzzy logic and image processing algorithms for early glaucoma diagnosis[J]. Biomedical Signal Processing & Control, 2018, 40:366-377.
[5] Moore, T.S, Campbell, J.W, Hui Feng. A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms[J]. IEEE Transactions on Geoscience & Remote Sensing, 39(8):0-1776.
[6] K Ramesh and Y Mallikharjuna Rao, "POWER TRANSFORMER PROTECTION USING WAVELETS," International Journal of Electrical and Electronic Engineering & Telecommunications, Vol. 3, No. 4, pp. 14-21, October 2014
[7] Pyae P. Phyo, Chawalit Jeenanunta, and Kiyota Hashimoto, "Electricity Load Forecasting in Thailand Using Deep Learning Models," International Journal of Electrical and Electronic Engineering & Telecommunications, Vol. 8, No. 4, pp. 221-225, July 2019. Doi: 10.18178/ijeetc.8.4.221-225
[8] Surhone L M, Tennoe M T, Henssonow S F. RapidMiner[M]. Betascript Publishing, 2010.
[9] Wu H， Li Y. The Impacts of Female Executives on Firm Performances: Based on Principle Component Analysis (PCA) and Data Envelopment Analysis (DEA)[M]// Proceedings of the Tenth International Conference on Management Science and Engineering Management. Springer Singapore，2017.

Acknowledgments
I would like to express my deepest gratitude to my leader, Mo Jun, a senior engineer who is a respected, responsible and resourceful scholar who provided me with valuable advice at each stage of writing this article. Without his enlightening guidance, impressive kindness and patience, I could not complete my dissertation. His sharp and powerful academic observation is not only inspiring to this article, but also to my future research.