International Core Journal of Engineering Volume 5 Issue 10, 2019
ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

Optimizing Depth Discrimination Restricted Boltzmann Machine Based on Genetic Algorithm for Fault Diagnosis of Transformer

Xiaoshan Li [a], Yongli Zhu [b]

School of Control and Computer Engineering, North China Power University, Baoding Hebei 071003, China.

[a]2915743179@qq.com, [b]yonglipw@163.com

Abstract

Aiming at the low accuracy of the traditional transformer fault diagnosis method and the problem that the discriminant restricted Boltzmann machine (DRMB) is easy to fall into local optimum and difficult to guarantee the network optimization, a new model for transformer fault diagnosis based on genetic algorithm optimization for depth discriminant restricted Boltzmann machine (GA-DDRBM) is proposed. The improved model superimposed the multi-layer DRBM model to form a depth DRBM model (DDRBM). The model used a genetic algorithm to perform global optimization, determined the optimal initial parameter values, and further trained the DDRBM model in the local solution space. Finally, a small amount of tagged data was used to fine-tune the model to obtain the optimal model parameters and complete the fault diagnosis. The experimental results show that the improved model has better stability and higher recognition accuracy than the simple DDRBM model. Compared with the traditional BP algorithm and SVM algorithm, the accuracy rate is also greatly improved.

Keywords

Fault diagnosis, Genetic algorithm, DRBM, Power transformer.

1. Introduction

The power transformer is one of the most important hub devices in the power system. The safety of its operating state is directly related to the safety and stability of the entire power system. However, due to the long-term continuous operation of the transformer and the influence and damage of the external environment, transformer failure occurs [1]. Therefore, it is of great significance to study the fault diagnosis of transformers and find faults in time.

Dissolved Gas Analysis (DGA) technology in transformer oil is one of the most effective methods for diagnosing latent faults inside transformers. DGA-based transformer fault diagnosis methods include three-ratio method, Rogers ratio method, and improved three-ratio method [2]. However, the disadvantage of these methods is that there may be a problem that the gas ratio is not comprehensive and the ratio falls below the critical point, causing the diagnosis to fail. At present, there are many intelligent diagnostic methods combined with the three-ratio method, such as BP neural network and SVM. However, due to the difficulty in SVM kernel function selection [3], BP neural network is easy to fall into local extremum, and the convergence speed is slow [4], the fault diagnosis accuracy can not meet the requirements. And these intelligent diagnostic methods are shallow machine learning methods, which can only process a small amount of data and cannot be processed for a large number of unlabeled samples.

Discriminant Restricted Boltzmann Machine (DRBM) is a typical deep network, which is widely used due to its good feature extraction ability and classification ability. Literature [5] uses the DRBM

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

model to classify medical images. Compared with traditional support vector machines and neural network classifiers, the classification accuracy is significantly improved. The literature [6] uses recurrent neural networks and restricted Boltzmann machines. The combined method predicts the short-term load of the power system, and its accuracy rate is also significantly improved. The literature [7] improves the classification-limited Boltzmann machine and increases the number of neurons in the model to make the classification effect better. In [8], Gaussian noise is added to the restricted Boltzmann model and used for transformer fault diagnosis. However, this model also has shortcomings. Because the random assignment of initial parameter values makes the model easy to fall into local optimum.

Aiming at these problems, a transformer fault diagnosis model based on genetic algorithm optimization depth discriminant restricted Boltzmann machine is proposed. Firstly, the multi-layer DRBM models are superimposed to form a depth discriminant-restricted Boltzmann machine. Then the genetic algorithm is used to determine the optimal initial parameter values, and the DDRBM model is further trained in the local solution space. Finally, a small amount of tagged data is used to fine-tune the model parameters. And the model was validated using DGA data. The model can effectively avoid falling into local optimum, and has stronger feature extraction ability and faster convergence speed, which can effectively diagnose transformer faults.

2. Depth Discriminant Restricted Boltzmann Machine Model Based on Genetic Algorithm Optimization

2.1 Discriminant Restricted Boltzmann Machine.

Discriminant Restricted Boltzmann Machine (DRBM) is an algorithm based on Restricted Boltzmann Machine (RBM) that can be used as an independent nonlinear classifier. Compared with RBM, its structure adds a classification unit [9] in addition to visible and hidden units. Its structure is shown in Fig. 1.

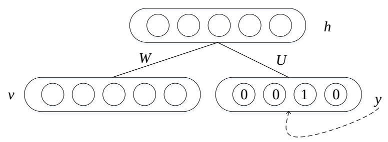

Fig.1 Structure of DRBM

Taking a layer of DRBM structure as an example, the DRBM energy is defined as:

$$E(y,v,h) = -h^T Wv - c^T v - d^T h - e^T y^\wedge - h^T Uy^\wedge \tag{1}$$

Where Θ={W, c, d, e, U} are parameters of the DRBM; W represents the connection weight between the input unit and the implicit unit; U represents the connection weight between the classification unit and the implicit unit; c indicates the offset of the input unit; d represents the offset of the implicit unit; e represents the offset of the classification unit.

Then the distribution determined by this model can be expressed as:

$$p(y,v,h;\Theta) = \frac{\exp^{-E(y,v,h)}}{Z(\Theta)} \tag{2}$$

It can be seen from the structure diagram of Fig. 1 that the joint probability distribution of hidden units can be derived from the state of the visible unit:

$$P(h|y,v) = sigm(d + Uy^\wedge + Wv) \tag{3}$$

Where sigm(x)=1/(1+e^{-x}). The conditional probability distribution of the taxon can be derived from the state of the hidden cell:

$$P(y|h) = \frac{e^{d_y + \sum U_{jy} h_j}}{\sum_{y^*} e^{d_{y^*} + \sum_j U_{jy^*} h_j}} \tag{4}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

Because DRBM is a classification model based on RBM construction, the contrast divergence algorithm (CD) can still be used for the tuning training problem of DRBM. Find the parameters by CD algorithm:

$$\Delta w_{ij} = \varepsilon(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recom}) \tag{5}$$

$$\Delta c_j = \varepsilon(\langle h_j \rangle_{data} - \langle h_j \rangle_{recom}) \tag{6}$$

$$\Delta b_i = \varepsilon(\langle x_i \rangle_{data} - \langle x_i \rangle_{recom}) \tag{7}$$

$$\Delta d_i = \varepsilon(\langle y_i \rangle_{data} - \langle y_i \rangle_{recom}) \tag{8}$$

Where ε is the learning rate, $<\bullet>_{data}$ represents the divergence of the data distribution, and $<\bullet>_{recom}$ represents the divergence of the model distribution. $<v_i h_j>_{data}$ represents the expected value at which the visual unit v_i and the implicit unit h_j are simultaneously activated when training the sample data.

2.2 Genetic Algorithm.

Genetic Algorithm (GA) is a global optimization search algorithm based on population genetics. The algorithm evolves and iteratively generates individuals with better fitness through selection, crossover and mutation operations [10]. The algorithm has strong robustness and adaptability, and has global optimization and search-independent gradients. Good performance is good for finding the optimal solution to the optimization problem.

2.3 GA-DDRBM Model Construction.

For the single-layer structure of the traditional DRBM model, the DDRBM model is a hierarchical model of multi-layer structure. The model superimposes the multi-layer DRBM model, which has higher accuracy than the single-layer DRBM model. Its structure is shown in Fig.2.

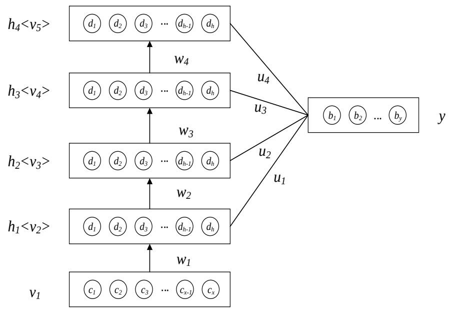

Fig.2 Structure of DDRBM

The entire model has four layers of DRBM, each layer of DRBM contains x input nodes, h hidden layer nodes and y classification nodes. c is the visible unit offset, d is the hidden unit offset, and b is the classification unit offset.

In order to make the model have higher classification accuracy and avoid the model falling into local optimum, this paper uses genetic algorithm to optimize the DDRBM model and construct GA-DDRBM model. The main steps are as follows:

1) Determine the encoding method.

Since the model parameter θ=(W, b, c, d, U) is complex and has a large range of variation, floating-point number coding is used. This floating-point number is also the true value of the parameter, and no corresponding decoding calculation is needed.

2) Initialize the genetic algorithm parameters.

The initial population is generated and the fitness function is determined. The objective function of the model is used as the fitness function of the GA algorithm, as shown in equation (9).

$$L(y, v; \Theta) = -\sum_{i=1}^{|Dt|} \log\ p(y_i, x_i; \Theta) \tag{9}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

Where Dt represents the training set, x_i is the input data, and y_i is the classification vector corresponding to x_i.

3) Genetic iteration.

Select the appropriate selection, crossover, mutation operator, multiple iterations, and find the best initial parameter values for the DDRBM model.

4) Model layer by layer training.

The model training adopts gradient descent and contrast divergence (CD) algorithm, and trains from the bottom layer of the model layer by layer, and updates the parameters layer by layer according to joint probability and conditional probability distribution.

5) Fine tuning of model parameters.

Based on the BP neural network, a small amount of tagged data is used to fine-tune the model parameters, and the global optimal parameters of the model are found, so that the model classification effect is optimal.

3. Transformer Fault Diagnosis Based on GA-DDRBM

In order to solve the problem that the DRBM model is easy to fall into local optimum, the GA-DDRBM model is proposed for transformer fault diagnosis. The sample data used in this paper is the transformer oil chromatographic data, including H_2, CH_4, C_2H_2, C_2H_4, and C_2H_6. The transformer fault diagnosis is performed by analyzing their contents and their changes.

Figure 3 is a flow chart of transformer fault diagnosis based on the deep DRBM model optimized by genetic algorithm.

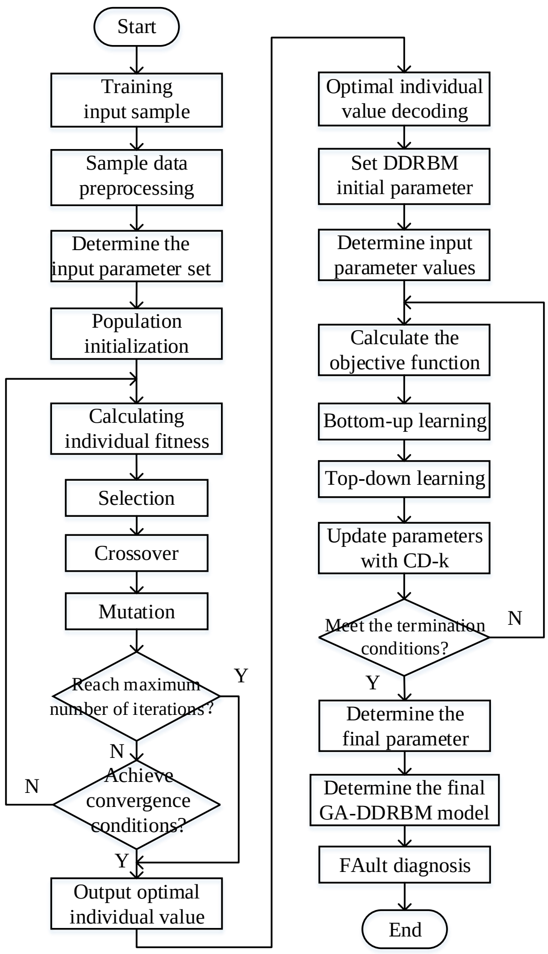

Fig.3 Fault diagnosis flow chart

The main steps of the algorithm are as follows:

1) After collecting DGA data and standardizing it, it is divided into two categories, one is training samples and the other is test samples.

2) Initialize the parameters in the genetic algorithm, including population size, crossover rate, mutation rate, etc., determine the initial population and fitness function of the genetic algorithm, and form a new generation population of chromosomes that meet the fitness function requirements.

3) Processing the newly generated population through selection, crossover, and mutation in the genetic algorithm, and continuing to generate new populations.

4) Repeat step 3) to allow the population to evolve until the specified number of iterations is reached. The chromosome at this time is the best chromosome and is decoded as the initial parameter value of DDRBM.

5) Train the model using the k-step CD algorithm and perform k-step Gibbs sampling. Calculate the following formula according to the probability distribution of the DRBM model:

$$h_0 = sigm(d + U_0 \hat{y}_0) \tag{10}$$

$$h_1 = P(h_0 | y_0, v_0) \tag{11}$$

$$y_1 = P(y_0 | h_0) \tag{12}$$

$$v_1 = P(v_0 | h_0) \tag{13}$$

The exact value of the DDRBM parameter is available:

$$\theta = \theta - \lambda(E_\theta^{'}(y_0, v_0, h_0) - E_\theta^{'}(y_1, v_1, h_1)) \tag{14}$$

Where λ is the learning rate.

6) Fine-tuning the model with a small number of labeled samples. The process uses BP neural network to further modify the model parameters to obtain a more accurate diagnostic model.

7) Using the processed DGA data as the model input, the GA-DDRBM model is used for transformer fault diagnosis.

4. Experimental results and analysis

4.1 Data Analysis

The oil chromatographic data used in this paper is the online monitoring of dissolved gas data in oil before and after the transformer faults collected in multiple field projects. Some of these data are unlabeled data for genetic algorithm and deep discriminative restricted Boltzmann machine training. Learning, a total of 1500 groups; the other part is a small amount of tagged data, used for fine-tuning the model parameters of the BP network.

Since the raw data of each gas content value has an order of magnitude difference, in order to avoid the influence on the calculation accuracy, these values need to be converted into dimensionless data, that is, the data is normalized so that the data are at 0-1. Within the scope, the normalization is as follows:

$$X_{new} = \frac{X - X_{mean}}{X_{\max} - X_{\min}} \tag{15}$$

Where X_{new} represents the normalized gas content value, X_{max} and X_{min} are the highest and lowest values of the gas content, respectively, X_{mean} represents the mean of the gas content, and X represents the actual value of the gas content.

The sample data is labeled as six states of the transformer, namely low energy discharge D1, high energy discharge D2, medium and low temperature overheat T12, high temperature overheat T3, partial discharge PD and normal.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

4.2 Parameter Setting and Analysis

The performance of the GA-DDRBM model was tested by Matlab. The running platform was Lenovo computer, processor i5-3317, running memory was 6G, operating system was win10, and programming environment was Matlab R2016a.

The parameters for setting the model training are as follows: the crossover rate of the genetic algorithm is set to 0.6, and the mutation rate is generally small, set to 0.2, and the initial population size is set to 100. Because the model input data is the processed transformer oil chromatographic data, including H_2, CH_4, C_2H_2, C_2H_4, C_2H_6 five kinds of gases; the output data is the transformer state, including low energy discharge D1, high energy discharge D2, medium and low temperature overheating T12, high temperature Overheated T3, partial discharge PD and normal six states, so the input layer of the DDRBM model is set to 5, the target layer is set to 6. And the network parameters W, U are initialized to values that are randomly small and obey Gaussian distribution. c, d, e are initialized to 0, the target layer uses the Softmax unit, the initialization learning rate is set to 0.1, and the training number is set to 100.

Experimental analysis of the model hidden layer number, hidden layer nodes and genetic algorithm evolution algebra, determine its optimal value, and prepare for transformer fault diagnosis.

4.2.1 Number of hidden layers

Since the number of DRBM layers will affect the accuracy of transformer fault diagnosis, multiple experiments are carried out to test the accuracy of transformer fault diagnosis with the change of the number of DRBM layers. The number of DRBM layers is changed from 1 to 10.

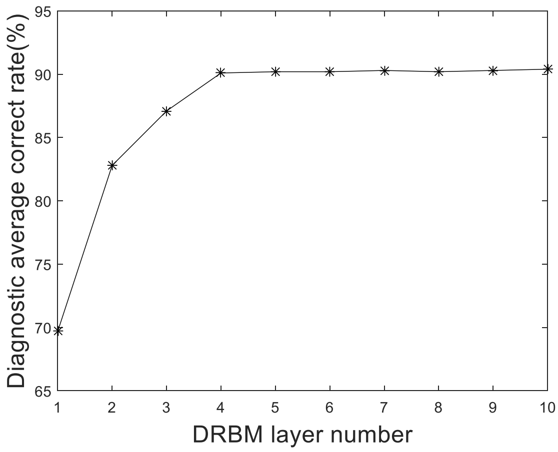

Fig.4 The relationship between the number of DRBM layers and the diagnostic average correct rate

As can be seen from the figure, as the number of DRBM layers increases, the diagnostic accuracy also increases, and the stability is basically achieved when the number of layers is 4. But because computer performance consumption increases as the number of layers increases. Therefore, the number of DRBM layers in this experiment was set to 4.

4.2.2 Number of hidden layer nodes

The number of DRBM hidden layer nodes also has a significant impact on the network classification effect, but there is no clear theoretical basis for its setting, so this paper uses experimental methods to determine the number of hidden layer nodes. The experimental results are shown in Figure 5.

For convenience, assume that the number of nodes for each hidden layer is the same. In the experiment, the number of hidden layers is set to 4, the number of hidden layer nodes is between 50 and 150, and the step size is set to 10. As can be seen from Figure 5, when the number of hidden layer nodes is set to 80, the classification accuracy is the highest.

4.2.3 Genetic algorithm evolutionary algebra

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

In view of the shortcomings of genetic algorithm, the experiment performed 5 experiments, taking the average value of 5 experiments as the average fitness, taking the optimal value of 5 experiments as the best fitness, and experimenting with the evolution of 120 generations as the termination condition. The experimental results are shown in Figure 6.

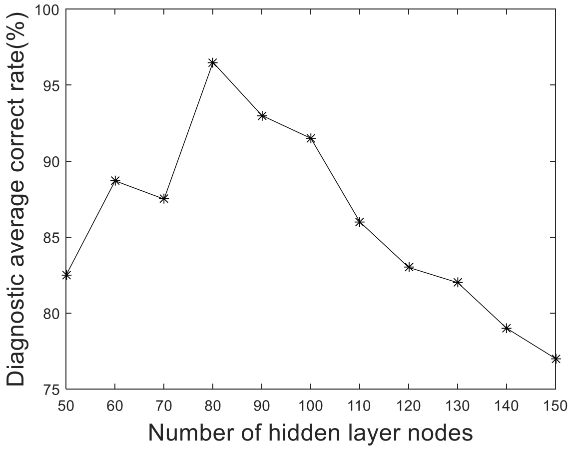

Fig.5 The relationship between the number of hidden layers nodesand the diagnostic average correct rate

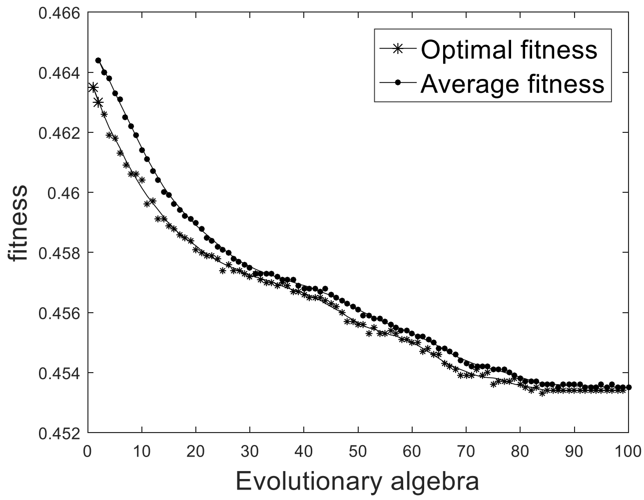

Fig.6 Simulation of fitness curves

It can be seen from the figure that with the increase of evolutionary algebra, the fitness is getting better and better, and the average fitness and the best fitness tend to be stable until about 100 generations, and the difference between the two is getting smaller and smaller.

4.3 Comparison of different algorithm precision

In order to further prove the effectiveness of the proposed method, the DGA data collected by the engineering site were used to test and compare the GA-DDRBM, DDRBM, DRBM, SVM and BP algorithms. The data set is DGA data, with 360 data as the training set and 483 data as the test set. The parameters required for the DDRBM and DRBM models are the same as the model in this paper, including the number of layers, the number of hidden layer nodes and the initial learning rate; the SVM kernel function uses the radial basis kernel function, the parameters are selected in the reference [11], and the regularization coefficient is set. For 2048, the kernel function parameter is set to 0.03; the BP neural network training parameter reference [12], the maximum number of training iterations is set to 1500, and the learning rate is set to 0.01. The test results are shown in Table 1.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

Tab.1 Accuracy comparison of fault diagnosis（%）

algorithm	D1	D2	T12	T3	PD	normal
GA-DDRBM	91.96	90.33	89.64	91.68	93.16	90.16
DDRBM	90.68	91.04	88.78	90.93	92.03	89.64
DRBM	88.35	85.71	88.00	90.63	81.98	86.33
SVM	82.16	73.32	83.32	90.09	81.73	84.56
BP	78.32	68.54	76.33	86.78	80.07	76.82

5. Conclusion

In this paper, a deep DRBM model based on genetic algorithm optimization is proposed and applied to transformer fault diagnosis. This method combines genetic algorithm and discriminative restricted Boltzmann machine, avoiding the problem that the simple DDRBM model is easy to fall into the local minimum value and the directivity problem caused by random initial assignment. At the same time, the genetic algorithm and DDRBM are played. Their respective advantages have effectively improved the performance of transformer fault diagnosis. Experiments show that the GA-DDRBM model has higher accuracy and faster convergence rate for transformer fault diagnosis than the simple DDRBM model. Compared with the traditional BP algorithm and SVM algorithm, the accuracy rate is also greatly improved.

Acknowledgements

This project was supported partially by National Natural Science Foundation of China (5167702).

References

[1] Shi Xin, Zhu Yongli, Sa Churila, et al. Power transformer fault classifying model based on deep belief network[J]. Power System Protection and Control, 2016, 44(1): 71-76.
[2] Zhao Feng, Li Shuo. Fault Diagnosis of Traction Transformer Based on DGA and Improved Grey Correlation Analysis Model[J]. High Voltage Apparatus, 2015, 51(1): 41-45
[3] Zhang Weihua, Yuan Jinsha, Zhang Tiefeng, Zhang Ke. Improved three-ratio method for transformer Fault Diagnosis using B-spline theory[J]. Proceedings of the CSEE, 2014, 34(24): 4129-4136.
[4] Shi Liping, Yu Pengxi, et al. Application of fuzzy support vector machine in transformer fault diagnosis[J]. Electrical Measurement & Instrumentation, 2015, 52(8): 115-119.
[5] Gong Maofa, Liu Yanni, Wang Laihe, et al. Fault diagnosis of power transformers based on chaos particle swarm optimization BP neural network[J]. Electrical Measurement & Instrumentation, 2016, 53(15): 13-16, 32
[6] Long Mingzhu, Li Zhuo, Xie Xiang, Li Guolin, Wang Zhihua. Adaptive Image Enhancement Based on Guide Image and Fraction-Power Transformation for Wireless Capsule Endoscopy[J]. IEEE transactions on biomedical circuits and systems,2018,12(5).
[7] LI Ruochen, ZHU Fan, ZHU Yongli, ZHAI Yujia. Short-term power load forecasting using recurrent neural network with restricted Boltzmann machine[J]. Power System Protection and Control, 2018,46(17):83-88.
[8] YIN Jing, LI Wei-wei, YANG De-hong, YAN He. Improved Model Based on Classification Restricted Boltzmann Machine[J].Journal of Chinese Computer Systems, 2018,39(07):1415-1419.
[9] WANG Lu-Kun, ZHAO Xiao-Ying, ZHANG Jian, Tian Chun-Peng, Su Hao, TANG Gong-You. Fault Diagnosis for Transformers Based on the Restricted Boltzmann Machine with Gaussian Noises[J]. Periodical of Ocean University of China,2018,48(06):114-119.
[10] Liaoyi Ning, Jian Liu, Chenghao Quan, Fang Shen, Yingwei Song, Yan Liu. Application of the Time-varying Outage Model of electric transmission and transformation equipment in Condition-based Maintenance[J]. IEEE PES Innovative Smart Grid Technologies, 2012, 36(3): 32-39.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201909_5(10).0017

[11] Sun Haifeng, Shen Ying, Wang Ya'nan, Prediction of contact resistance based on optimized BP neural network of genetic algorithm[J]. Electrical Measurement & Instrumentation,2019,56(05):77-83.
[12] Cao Jiuwen, Zhang Kai, Yong Hongwei, Lai Xiaoping, Chen Badong,Lin Zhiping. Extreme Learning Machine With Affine Transformation Inputs in an Activation Function.[J]. IEEE transactions on neural networks and learning systems, 2019,30(7).