ISSN: 2688-8645

Research on Fault Diagnosis Method of Wind Turbine based on CALO Algorithm for Optimizing WKLM

Cong Jiang, Chenggang Zhen

School of Control and Computer Engineering, North China Electric Power University, Baoding 071000, China.

Abstract

Aiming at the fault characteristics of generator equipment of wind turbine, this paper proposes a fault diagnosis method based on improved ant lion optimization algorithm (CALO) to optimize wavelet kernel limit learning machine (WKELM). Firstly, the feature vector of SCADA data is extracted by using the guaranteed bureau projection algorithm. Secondly, CALO algorithm USES cauchy mutation operator to improve ant lion algorithm and improve global optimization ability. Finally, CALO algorithm is used to optimize the parameters of the wavelet kernel limit learning machine, so as to improve the diagnostic accuracy and convergence speed of the algorithm. In order to verify the effectiveness of the designed diagnostic model, experimental verification is carried out with the measured operating data of the wind turbine equipment in a wind farm in northwest China. The simulation results show that the CALO-WKELM diagnostic model can effectively identify different faults of generators and meet the needs of generator fault diagnosis.

Keywords

Wind turbine, ant lion optimization, wavelet kernel extreme learning machine,locality preserving projections, fault diagnosis.

1. Introduction

As a key component of wind turbine generators, wind turbines will affect the normal operation of the entire equipment if they fail. At present, generators are developing in the direction of intelligence, and wind turbine fault diagnosis technology is one of the key technologies for wind turbines to realize intelligence [1]. Therefore, it is of practical significance to study the fault diagnosis technology of wind turbines and conduct fault early warning to avoid major faults of wind turbines. How to construct an accurate and effective generator fault diagnosis method to realize early warning of wind turbine generator faults has become the focus of research by experts and scholars.

The fault diagnosis process of wind turbine mainly includes: feature extraction and fault identification [2]. Feature extraction is a key technology, which affects the accuracy of fault diagnosis methods. The SCADA system is a data acquisition and monitoring system, which is widely used in wind turbine status monitoring. The system can monitor and record information parameters of fan failures in an all-round way, and it is convenient to collect information. Security projection is an algorithm for manifold learning, which has the non-linear characteristics of keeping the data space structure unchanged. It can be used to extract the characteristic signals of high-dimensional information in the SCADA data samples of wind turbines [3].

With the rapid development of artificial intelligence technology, various intelligent algorithms are applied to fault diagnosis. At present, domestic and foreign experts have conducted a series of studies on the fault diagnosis of wind turbines, put forward some fault classification methods,

and obtained certain research results. Literature [4] uses BP neural network to identify faults in the main equipment of wind turbines, and obtains high diagnostic accuracy, but there are still shortcomings of long neural network model training period and model instability. Literature [5] uses a fault diagnosis method combining EEMD and RVM. This method has high accuracy for early fault classification of wind turbines, but the proposed method is suitable for small data samples. In the processing of large data samples, the model processing speed is slow and the classification The accuracy is low. Compared with traditional classifiers, extreme learning machines have faster learning speed and strong generalization ability [6]. Literature [7] introduces RBF into the extreme learning machine, and proposes a rolling bearing fault diagnosis method based on energy entropy and nuclear extreme learning machine. KELM improves the nonlinear approximation ability, but the choice of parameters will affect the performance of RBF. Literature [8] proposed a fault diagnosis method for rolling bearings based on OMP and VNWOA-WKELM. Using Wavelet Kernel Extreme Learning Machine (WKELM) as the classifier, the classification effect is good. However, WKELM is greatly affected by parameters. Using WOA to optimize WKELM parameters can weaken the dependence of the algorithm on parameters and improve the generalization ability of the algorithm, but the defects of slow convergence and low accuracy are still obvious.

Based on the above shortcomings and actual diagnosis needs, using wind turbine SCADA data samples, a wind turbine fault diagnosis method based on the improved Antlion optimization algorithm to optimize the wavelet kernel extreme learning machine is proposed. First, use the security projection algorithm to preprocess the SCADA data, and extract the feature vector representing the intrinsic structure of the generator state parameters; second, use the Cauchy mutation factor to improve the Antlion algorithm (CALO) to optimize the penalty factor of the wavelet kernel extreme learning machine C and wavelet kernel function parameters a, b, c are used to obtain the optimal parameter pair and establish the optimal fault classification model. Finally, the feature vector representing the intrinsic structure extracted by the LPP algorithm is used as the input of the CALO-WKELM model to perform the fault classification of the wind turbine.

2. Antlion Optimization Algorithm and Wavelet Kernel Extreme Learning Machine

2.1. Antlion Optimization Algorithm

Ant Lion Optimizer (ALO) is a bionic optimization algorithm [9], and its algorithm idea comes from the natural behavior of ant lions preying on ants in nature. The ant colony algorithm includes three roles: ants, ant lions and elite ant lions. The position of the ant represents the trial solution, the position of the ant lion represents the local optimal solution, and the position of the elite ant lion represents the global optimal solution. Ants walk around randomly, but under the influence of the ant lion trap, they will slide to the ant lion's position, and then explore the space around the ant lion. If the ant finds a better solution than the current ant lion position, the ant lion will prey on the ant and update the position, The optimal ant lion position obtained in each iteration is the elite ant lion position, and the final iteration is completed to obtain the global optimal solution, thus ensuring the diversity of the population and the optimization performance of the algorithm. The mathematical model formula of ant random walk is as follows:

$$X(t) = [0, cs(2r(t_1) - 1), \cdots, cs(2r(t_n) - 1)] \tag{1}$$

Where cs is the cumulative sum of calculations, n is the maximum number of iterations, r(t) is a random function, and the expression is:

$$r(t)=\begin{cases}1 & \text{rand}>0.5\\ 0 & \text{rand}\le 0.5\end{cases} \tag{2}$$

Among them, t is a random walk step, and rand is a random number generated uniformly in the interval [0,1].

The random walk of the ant will be affected by the ant lion trap, sliding to the ant lion position, and moving in the hypersphere around the ant lion. The mathematical model expression is as follows:

$$\begin{cases}c_i^t = Antlion_j^t + c^t\\ d_i^t = Antlion_j^t + d^t\end{cases} \tag{3}$$

Among them, c_i^t、d_i^t are the minimum and maximum values of the t-th iteration of the i-th ant; c^t、d^t are the minimum and maximum values of the t-th iteration; $Anlion_j^t$ is the position of the j-th ant lion in the t-th iteration .

Due to the influence of the ant lion trap, the radius of motion of the ant in the hypersphere will be reduced. Therefore, as the number of iterations increases, the values of c and d will be adaptively reduced to improve the convergence speed. The mathematical model of adaptive reduction is as follows:

$$\begin{cases}c^t = \dfrac{c^t}{I}\\ d^t = \dfrac{d^t}{I}\end{cases} \tag{4}$$

Among them, I is the ratio.

In the iteration of the ant lion optimization algorithm, the elite ant lion is the best ant lion in each iteration, which affects the behavior of all ants. The Antlion optimization algorithm uses a combination of roulette selection and random walk to determine the position of the ant to reduce or even avoid the algorithm from falling into the local extreme value problem. The formula is as follows:

$$Ant_i^t = \frac{R_A^t + R_E^t}{2} \tag{5}$$

Where R_A^t is the random walk during the tth iteration of the roulette strategy; R_E^t represents the random walk of the ant under the influence of the elite ant lion at the tth iteration; Ant_i^t is the position of the i-th ant in the tth iteration .

In the iteration, if the position of the ant is better than the position of the ant lion, the ant will be eaten, and the ant lion will replace the ant position to obtain a better position. The formula is as follows:

$$AntLion_j^i = Ant_i^t \ \ if \ \ f(Ant_i^t) > f(AntLion_j^i) \tag{6}$$

2.2. Wavelet Kernel Extreme Learning Machine

Extreme Learning Machine (ELM) is an efficient and simple single hidden layer feedforward neural network algorithm [10], which solves the problem of low learning efficiency and many parameters of BP neural network algorithm. Its derivative algorithm kernel extreme learning machine introduces a kernel function, combines SVM and ELM, and improves the generalization ability and calculation speed of the model. The expression function of the ELM algorithm is as follows:

$$f(x) = h(x)\beta \tag{7}$$

In the formula, h(x) is the neuron composition vector, and β is the output weight.

The goal of ELM is to minimize the output weight and goal. The following formula can be constructed:

$$\min : \|\beta\|^2 \ \text{ and } \sum_{i=1}^{N} \|\beta \cdot h(x_i) - t_i\| \tag{8}$$

Can be converted into:

$$\min : L = \frac{1}{2}\|\beta\|^2 + \frac{1}{2}C\sum_{i=1}^{N}\|\xi_i\|^2 \tag{9}$$

According to the KKT condition, the objective function after introducing the Lagrange multiplier is transformed into:

$$L = \frac{1}{2}\|\beta\|^2 + \frac{1}{2}C\sum_{i=1}^{N}\|\xi_i\|^2 - \sum_{i=1}^{N}\alpha_i(h(x_i)\beta - t_i + \xi_i) \tag{10}$$

Where: x_i is the training sample; t_i is the target output value, C is the penalty factor; ξ_i is the training error; α_i is the Lagrangian operator. The ELM output function can be expressed as:

$$f(x) = h(x)\beta = h(x)H^T(\frac{I}{C} + HH^T)^{-1}T \tag{11}$$

Where: I, H, T are the identity matrix, the hidden layer output matrix, and the output layer matrix respectively. The kernel function is introduced into ELM, and the kernel extreme learning machine algorithm is constructed. The output function equation of the KELM algorithm is as follows:

$$f(x) = h(x)H^T(\frac{I}{C} + HH^T)^{-1}T = \begin{bmatrix} K(x, x_1) \\ \cdots \\ K(x, x_N) \end{bmatrix} (\frac{I}{C} + \Omega)^{-1}T \tag{12}$$

The wavelet kernel function has good nonlinear classification ability, can achieve approximation to other functions, and has strong generalization ability. The wavelet function satisfies the Mercer condition and can be used as the kernel function of the kernel extreme learning machine. The expression of the wavelet kernel function is as follows:

$$K(x, x') = \cos\left[a\frac{\left\|x_i - x_i'\right\|}{b}\right]\exp\left[-\frac{\left\|x_i - x_i'\right\|^2}{c}\right] \tag{13}$$

In the formula: a is the coefficient factor; b is the scale factor; c is the translation factor.

3. Fault Diagnosis Strategy

3.1. Improved Antlion Optimization Algorithm

The Antlion optimization algorithm is a relatively new swarm intelligence algorithm with the advantages of fewer parameters and easy implementation. It is used in many fields. For example, the Antlion optimization algorithm based on Levi mutation is applied to the parameter identification of silicon single crystal; The lion optimization algorithm is applied to noise influence, etc., and the advantages are obvious, but the shortcomings still exist. Based on this, the Cauchy mutation operator optimization Antlion optimization algorithm [11] is proposed. This paper applies it to the generator fault identification and obtains good optimization results.

The Cauchy distribution is as follows:

$$s = \frac{t}{\pi(x^2 + t^2)} \qquad -\infty < x < +\infty \tag{14}$$

Using the characteristics of the Cauchy distribution, the Antlion optimization algorithm is improved to enhance the global optimization ability, and at the same time, random quantities that obey the Cauchy distribution are introduced into the elite Antlion. The mathematical model is as follows:

$$x' = x + \eta\ Cauchy(0,1) \tag{15}$$

Among them: x' is the updated position of x, Cauchy (0,1) is the Cauchy distribution, and η is the variation intensity parameter of the Cauchy distribution.

Introducing the Cauchy mutation operator, the ant lion group update position formula is as follows:

$$\begin{cases} A_0' = A_0 + A_0 * Cauchy(0,1) \\ AL_0' = AL_0 + AL_0 * Cauchy(0,1) \end{cases} \tag{16}$$

Where: A0 and AL0 are the positions of the original ants and ant lions, and A_0', AL_0' is the new position after improvement.

3.2. WKELM Parameter Selection

The extreme learning machine algorithm has fast learning speed, good generalization performance and few parameters, so the selection of the penalty factor C and the kernel parameters a, b, c will affect the final recognition result of the wavelet kernel extreme learning machine. Only when the parameters are selected reasonably, WKELM can achieve better classification accuracy, and random selection of parameters will reduce the classification accuracy. Based on this, in the fault diagnosis of wind turbine generators, this paper uses the Antlion optimization algorithm to optimize the parameter selection of WKELM, and select reasonable parameters. Therefore, the correct rate of generator fault WKELM classification is regarded as a function of C and wavelet kernel function parameters a, b, c, and the mathematical model of the optimization process is:

$$\begin{cases} \max\ fitness = acc(C,a,b,c) \\ st: C \in (l_C, u_C) \\ a,b,c \in (l_{a,b,c}, u_{a,b,c}) \end{cases} \tag{17}$$

In the formula: $fitness$ is the fitness value, acc is the classification accuracy rate, and u,l represents the upper and lower limits of the 4 types of parameters.

3.3. CALO Optimizes WKELM Steps

The CALO algorithm is used to optimize the parameters of WKELM. The experimental steps are as follows:

Step 1 Normalize the SCADA data to eliminate the difference between different dimensions, and divide the data into training set and test set according to a 3:2 ratio.

Step 2 Set the parameters of the CALO algorithm, including the maximum number of iterations, population size and the number of hidden layer nodes. And initialize the value of the wavelet kernel extreme learning machine [C, a, b, c].

Step 3 Initialize the population. The restriction conditions of formula (17) are used to initialize the ant and ant lion populations randomly.

Step 4 Calculate the fitness value of each ant and ant lion individual, then merge the populations and sort them according to the fitness value from largest to smallest. The first N fitness values are assigned to the ant lion population, and the last N are assigned to the ant population.

Step 5 The N ant lion populations are searched in parallel, sorted by fitness again, and the elite ant lions are updated according to formula (6).

Step 6 In the iterative update of position, there is almost no difference between the elite ant lions of two adjacent generations. It is determined that the algorithm has fallen into a local optimum and needs to jump out of operation.

Step 7 Jump out of the local optimum: copy the ant lion colony to the original population size, introduce the Cauchy mutation factor, update the ant lion position using formula (16), and find the elite ant lion.

ISSN: 2688-8645

Step 8 Determine whether the CALO algorithm currently reaches the maximum number of iterations. If it reaches, it jumps out of the algorithm loop, and outputs the optimal fitness value corresponding to the elite ant lion and the parameter [C,a,b,c] to be optimized by WKELM; if not, it returns to step 4 to continue iterating.

Step 7: Select the WKELM parameters corresponding to the individual position of the elite ant lion to train the data and classify the test.

3.4. Establishment of Diagnostic Model

Wind turbines are the key equipment of wind turbines, and their operating status affects the efficiency of wind power generation. Because wind turbines are in a harsh environment, as time goes by, the equipment enters a period of high failures. According to actual needs, this paper designs the CALO-WKELM generator fault diagnosis model. Take common generator faults as an example, including three types of faults such as generator bearing temperature overrun, bent shaft, and rotor mass imbalance, and design the diagnosis model [12]. The structure of WKELM fault diagnosis model is shown in Figure 1.

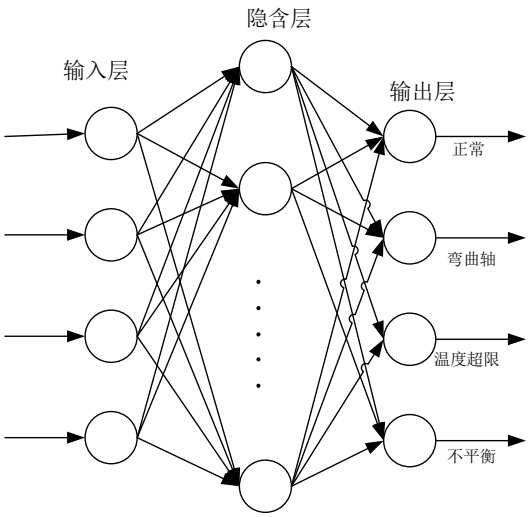

Fig 1. Structure diagram of WKELM fault diagnosis model

The wavelet kernel extreme learning machine contains 4 important parameter variables, namely C, a, b, and c. The choice of parameter values directly affects the classification performance of the algorithm. Therefore, this paper uses the Antlion optimization algorithm to find the four parameters of WKELM. Excellent to ensure the performance of the algorithm. Based on this, this paper proposes a generator fault diagnosis strategy based on CALO algorithm to optimize WKELM. The fault diagnosis process is shown in Figure 2.

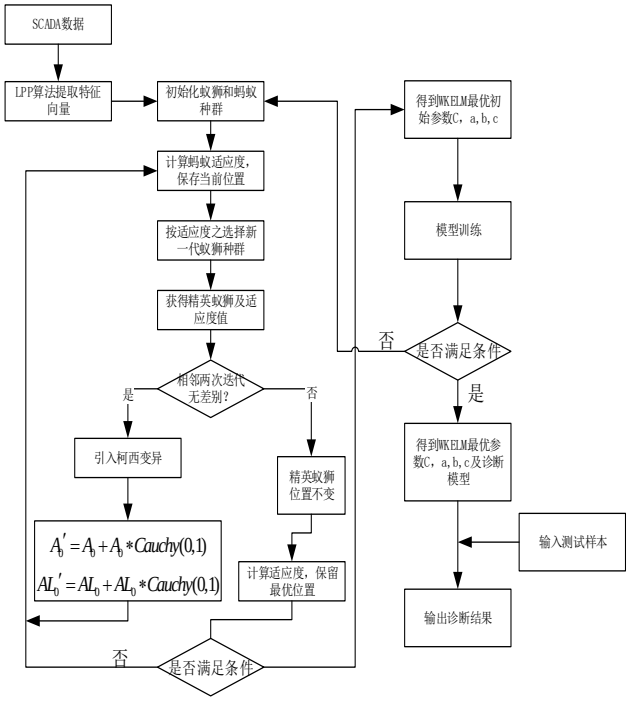

Fig 2. Fault diagnosis flow chart based on CALO - WKELM

4. Simulation and Case Analysis

4.1. Simulation Verification

In order to verify that the Antlion optimization algorithm designed in this paper optimizes the performance of the wavelet kernel extreme learning machine, three types of data in the UCI standard database are used as simulation experimental data to verify the performance of the algorithm. The experimental data information is shown in Table 1.

Table 1. Experimental data set

Data set	Attributes	Number of categories	Total number of samples	Training set	Test set
Iris	4	3	150	90	60
Wine	13	3	180	108	72
Seeds	7	3	210	126	84

This paper combines the nuclear extreme learning machine (KELM), wavelet nuclear extreme learning machine (WKELM), particle swarm optimized nuclear extreme learning machine (PSO-KELM), whale optimized nuclear extreme learning machine (WOA-KELM) with the design of this article Antlion's optimized wavelet kernel extreme learning machine (CALO-WKELM) was compared on three types of experimental data to verify the superiority of CALO-WKELM.

For comparison, the unified population size of all intelligent algorithms is set to 100, the number of iterations is 50, and the number of trials is 20. At the same time, in order to achieve the optimal performance of each algorithm, the number of nodes in the collective hidden layer of the collective algorithm is determined by the data set. The simulation results are shown by the mean (mean), standard deviation (std), mean training time (time) and the number of hidden

layer nodes (L) of the classification test accuracy rate. The results are shown in Table 2, Table 3 and Table 4.

Table 2. Comparison of classification results in Iris

algorithm	mean	Std	time	L
KELM	0.9428	0.0435	0.0736	22
WKELM	0.9546	0.0325	0.1896	22
PSO-KELM	0.9736	0.0296	6.3872	20
WOA-KELM	0.9796	0.0282	9.6028	23
CALO-WKELM	0.9801	0.0146	6.0018	20

Table 3. Comparison of classification results in Wine

algorithm	mean	Std	time	L
KELM	0.8884	0.2096	0.0432	12
WKELM	0.8926	0.2153	0.0490	12
PSO-KELM	0.9162	0.0201	1.9852	10
WOA-KELM	0.9214	0.0276	3.0765	13
CALO-WKELM	0.9364	0.0231	1.6649	10

Table 4. Comparison of classification results in Seeds

algorithm	mean	Std	time	L
KELM	0.8792	0.1122	0.0736	20
WKELM	0.8810	0.1436	0.1896	20
PSO-KELM	0.9012	0.0196	2.3692	15
WOA-KELM	0.9108	0.0147	4.5893	20
CALO-WKELM	0.9226	0.0158	3.0276	16

From Table 1, Table 2, and Table 3, it can be seen that the kernel extreme learning machine with optimized parameters of the algorithm has a significant improvement in classification accuracy compared with the unoptimized kernel extreme learning machine, which reflects the advantages of the machine learning algorithm. Compared with the PSO-KELM algorithm and the WOA-KELM algorithm, the CALO-WKELM algorithm shows higher classification accuracy on the three types of data sets in the UCI database, and shows excellent optimization capabilities. Through 20 test results and average verification, the results show that the Antlion optimization algorithm has advantages in the parameter optimization of the wavelet kernel extreme learning machine. The designed CALO-WKELM algorithm has a higher classification accuracy in data classification. In summary, the wavelet kernel extreme learning machine optimized by the Antlion optimization algorithm is a relatively stable algorithm with high classification accuracy, and has the potential to improve the accuracy of wind turbine fault diagnosis.

5. Conclusion

This paper studies the Antlion optimization algorithm, the wavelet kernel extreme learning machine, and the guaranteed projection algorithm, and proposes a wind turbine fault classification and recognition method based on the Cauchy mutation operator to improve the Antlion optimization algorithm to optimize the wavelet kernel extreme learning machine. The

security bureau projection algorithm is used to extract the feature vector of the high-dimensional information of the SCADA data to weaken the influence of irrelevant features and nonlinearity on the generator fault identification; secondly, it is proposed to use the Cauchy mutation operator to improve the Antlion optimization algorithm to classify the wavelet kernel extreme learning machine The parameters [C,a,b,c] of the generator are optimized to construct a recognition classifier for fault diagnosis of wind turbines. The experimental results show that (1) the proposed method can effectively dig out the information in the SCADA data, effectively identify the fault of the wind turbine and have high diagnostic accuracy. (2) Compared with other models in the article, CALO has better optimization ability, which can effectively solve the problem of poor parameter selection in WKELM, greatly improve the diagnosis accuracy and speed of the fault diagnosis model, and achieve the wind turbine fault diagnosis the goal of.

References

[1] Wu Jianbo, Wang Chunyan, Hong Huajun, et al. Fault diagnosis of marine diesel engine based on extreme learning machine[J]. Computer Engineering and Applications, 2019, 55(15)：147-152..

[2] Wong Tiantian, Wang Yan, Ji Zhicheng, et al..Fault Diagnosis of Rolling Bearing Based on Improved Extreme Learning Machine [J]. Journal of System Simulation,2018,30(11):4413-4420.

[3] Liu Shuai, Liu Changliang, Zeng Huaqingl．Research on fault warning for wind turbine gearbox based on kernel extreme learning machine[J]. CHINA MEASUREMENT & TEST,2019,45(02):121-127.

[4] Dong Kaisong,,Li Taotao,Yin Haolin.Fault Analysis and Intelligent Diagnosis of Wind Turbine Generator Set [J]. High Voltage Apparatus,2016,52(10):176-181.

[5] Niu Shengyu Zhang Xinyan Yang Lulu,et al.Research on fault diagnosis method of wind turbine based on EEMD-RVM[J]. Electrical Measurement ＆ Instrumentation , 2018,55(19):1-6.

[6] Wang Chunming，Zhu Yongli.Transformer fault diagnosis based on deep de-noising extreme learning machine[J]. Electrical Measurement ＆ Instrumentation , 2019,56(15):144-147.

[7] Qin Bo Wang Zuda Sun Guodong,et al.Application of VMD and hierarchical extreme learning machine in rolling bearing fault diagnosis[J]. China Measurement & Test ,2017,43(5):91-95.

[8] Xu Jiya, Wang Yan, Ji Zhicheng..Fault Diagnosis Method of Rolling Bearing Based on WKELM Optimized by Whale Optimization Algorithm[J]. Journal of System Simulation , 2017, 29(9): 2189-2197.

[9] Mirjalili S, Mirjalili M, Lewis A. Grey Wolf Optimizer[J]. Advances in Engineering Software (S0965-9978), 2014, 69(3): 46-61.

[10] HUANG G B, ZHU Q Y，SIEW C K．Extreme learning machine: A new learning scheme of feed forward neural net- works [C]/ /2004 IEEE International Joint Conference on Neural Networks.IEEE,2004.

[11] Yu J F,Liu S,Han F F,et al..Ant lion optimization algorithm based on cauchy variation [J]. microelectronics & computer,2019,36(6):46-54.

[12] Yuan Jinsha，Zhang Liwei, Wang Yu,et al..Study of Transformers Fault Diagnosis Based on Extreme Learning Machine[J]. Electrical Measurement & Instrumentation , 2013,50(576):21-26.