

Application of Regularized Extreme Learning Machine Based on BIC Criterion and Genetic Algorithm in Iron Ore Price Forecasting

Futian Weng¹, Muzhou Hou^{1,*}, Tianle Zhang¹, Yunlei Yang¹, Zheng Wang¹, Hongli Sun¹, Hao Zhu¹ and Jianshu Luo²

¹School of Mathematics and Statistics, Central South University, 410083 Hu Nan Changsha, China

²College of Science, National University of Defense Technology, 410073 Hu Nan Changsha China

*Corresponding author

Abstract—Forecasting international iron ore is a well-known issue, BIC criterion is used to select the relevant variables of iron ore price. On the basis of the traditional extreme learning machine (ELM), the regular term is introduced to control the complexity of the model, and the genetic algorithm (GA) is used to regularize the extreme learning machine. The input-layer weight matrix and the hidden-layer threshold matrix of the (RE-ELM) model are optimized to establish a BIC-based genetic algorithm and a regularization extreme learning machine (BIC-GA-RELM) iron ore price prediction model to increase the performance of the RE-ELM model. The results show that BIC-GA-RELM model has achieved the state of art performance, then a new method is provided for iron ore price prediction.

Keywords—BIC criterion; genetic algorithm; regularized extreme learning machine; iron ore price forecast

I. INTRODUCTION

China's imports account for 65% of the total global iron ore trade. The rise or fall in iron ore prices has a huge impact on China's economic interests. It is of great value to accurately predict the direction of iron ore price changes [1-6].

Ultimately, the evolution of prices is time-sequential, and it is not only the time factor that determines prices. In fact, many external factors work together [7-14]. It is the joint action of the influencing factors that determines prices and changes. Over time, external factors and their structures are all changing, which may lead to long-term accumulation of biases in price forecasts, leading to the use of classical statistical methods to predict long-term prices. The results were not satisfactory [15-17].

In recent years, with the development of new technologies and new theories, artificial neural networks (ANN) that give intelligence principles have achieved outstanding results in price forecasting, and they have shown more excellent performance and have good robustness [18-20]. However, the traditional neural network has the disadvantages of over-fitting, difficult to determine the hidden nodes, and it is complex to search the optimal parameters. Moreover, when there are large input data and exists multiple collinearity, the training speed and efficiency of the network will decline obviously [21-25].

Based on this, this paper proposes a BIC-GA-RELM iron ore price forecasting model with BIC criterion. This model adds

a regular term to control the complexity of the learning model. The genetic algorithm is used to optimize the initial weights and thresholds of RE-ELM, and the genetic algorithm is chosen to adapt to the chromosome process. Through genetic operations such as selection, crossover and mutation of genetic algorithms, the optimal solution is selected as the input weight and threshold of RE-ELM. Then use the least squares method to calculate output weights of neurons in the hidden layer, and calculate the predicted output of iron ore price to verify the feasibility and accuracy of this model for iron ore price prediction.

II. RESEARCH METHODS

A. Performance Evaluation Criteria

Four criteria are introduced to measure the prediction accuracy. The mean square error (MSE), the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percent error (MAPE). The representations are presented as below:

$$MSE = \frac{1}{N} \sum_{i=1}^N (R_{pred} - R_{obs})^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (R_{pred} - R_{obs})^2} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |R_{pred} - R_{obs}| \quad (3)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{R_{pred} - R_{obs}}{R_{obs}} \right| \quad (4)$$

where N is the number of data for testing period, and R_{obs} , R_{pred} denote the real value and estimating value.

B. Base Models

1) Bayesian information criteria

The model selection for the Bayesian Information Criterion (BIC) and the AIC Criterion is similar, which was proposed by Schwarz in 1978. The basic idea of the BIC criterion is to assume that there is a uniform distribution in the candidate model, then use the sample distribution to find the posterior distribution on the model, and finally select the model with the maximum posterior probability, so we think that the subset of variables that minimize the BIC value is the most excellent [32].

$$BIC = -2 \ln g(\hat{\theta}_k | y) + k \log n \quad (5)$$

2) Extreme learning machine

Most algorithms proposed for feedforward neural networks do not consider the generalization performance when they are proposed first time. ELM aims to reach better generalization performance by reaching both the smallest training error and the smallest norm of output weights:

$$\text{Minimize: } \|\beta\|_p^{\delta_1} + C \|H\beta - T\|_q^{\delta_2} \quad (6)$$

Where $\sigma_1 > 0$, $\sigma_2 > 0$, $p, q = 0, 1, 2, \dots, +\infty$. The first term in the objective function is a regularization term which controls the complexity of the learned model.

In 2012, Huang, Zhou, et al. especially studied the stability and generalization performance of ELM with $\sigma_1 = \sigma_2 = p = q = 2$:

$$\min_{\beta \in R^{L \times m}} \frac{1}{2} \|\beta\|^2 + \frac{C}{2} \sum_{i=1}^N \|e_i\|^2 \quad (7)$$

$$\text{s.t. } h(x_i)\beta = t_i^T - e_i^T, i = 1, \dots, N$$

By substituting the constraints of (7) into its objective function, we obtain the following equivalent unconstrained optimization problem:

$$\min_{\beta \in R^{L \times m}} L_{ELM} = \frac{1}{2} \|\beta\|^2 + \frac{C}{2} \|T - H\beta\|^2 \quad (8)$$

The above problem is widely known as the ridge regression or regularized least squares. By setting the gradient of LELM with respect to β to zero, we have

$$L_{ELM} = \beta^* - CH^T(T - H\beta^*) = 0 \quad (9)$$

If H has more rows than columns ($N > L$), which is usually the case where the number of training patterns is larger than the number of the hidden neurons, we have the following closed form solution for β :

$$\beta^* = (H^T H + \frac{I}{C})^{-1} H^T T, \quad (10)$$

where I is an identity matrix of dimension L .

If the number of training patterns is less than the number of hidden neurons ($N < L$), then H will have more columns than rows, which usually gives an under-determined least squares problem. Then we get:

$$\beta^* = H^T (H^T H + \frac{I}{C})^{-1} T, \quad (11)$$

where I is an identity matrix of dimension N .

C. Comparison and Evaluation of Models

Huang et al. [26] proposed an Extreme Learning Machine (ELM) algorithm in 2006 that randomly generated the input layer weight matrix and the hidden layer threshold matrix, they also did not need to be adjusted in subsequent operations. The characteristics of ability and self-learning have greatly improved the speed of network training and generalization ability, and have been widely used [27]. However, when solving the gradient descent problem, the randomly generated network input weights and the threshold vector parameters of the hidden layer nodes cannot guarantee the ELM model is optimal, and the complexity of the model cannot be controlled. In response to these problems, Chen Shaowei [28] proposed a method to optimize ELM with bat algorithms, Wang Jie et al [29] proposed an extreme learning machine with particle swarm optimization, Zhang Weihui [30] proposed DE-ELM, but these methods are likely to fall into Local optimum.

Genetic algorithm [31] (GA) is a computational model that simulates the biological evolution process of Darwinian biological evolution theory's natural selection and genetic mechanism. It is a global search optimization method based on probability transformation rules. The solution of the problem is regarded as a population, and the result is getting better and better by continuous selection, crossover, mutation and other genetic operations. Adding regularization terms, you can control the complexity of the ELM model. At the same time, due to the possibility of collinearity between variables, it may well lead to decline the prediction accuracy of the model. Therefore, it is very important to select variables initially in the learning model. In recent years, the combination of AIC criteria with different neural networks has been widely used.

The BIC-GA-RELM iron ore price prediction model based on the BIC criterion genetic algorithm is proposed in this paper. This model adds a regular term to control the complexity of the learning model. The genetic algorithm is used to optimize the initial weights and thresholds of RE-ELM, and the

genetic algorithm is chosen to adapt to the chromosome process. Through genetic operations such as selection, crossover and mutation of genetic algorithms, the optimal solution is selected as the input weight and threshold of RE-ELM. Then use the least squares method to calculate output weights of neurons in the hidden layer, and calculate the predicted output of iron ore price to verify the feasibility and accuracy of this model for iron ore price prediction.

D. Design of Prediction Method

Compared to the base ELM algorithm, the regularized extreme learning machine (RE-ELM) can control the complexity of learning and improve the generalization ability of the ELM model. However, the random input layer weight matrix and the hidden layer threshold matrix will cause the iron ore price change a lot by the network, and it is hard to obtain the optimal learning model for prediction. The genetic algorithm has a strong global optimization ability. Using genetic algorithm to find the optimal initial W and b for the RE-ELM model can further improve the fitting accuracy and increase the generalization performance of the model, obtain the optimal RE-ELM model.

In this model the input weights and implicit node thresholds of the RE-ELM training data are mapped to the genes of each chromosome in the genetic algorithm population; the chromosome fitness of the genetic algorithm corresponds to the training error of the RE-ELM. The problem of solving the optimal input weights and thresholds is transformed into the problem of selecting optimal chromosomes by reducing chromosome fitness. Through genetic algorithm selection, crossover, mutation and other operations, the optimal chromosome is selected as the input weight and threshold of RE-ELM after optimization. Then use the least squares method to obtain the output weights of the hidden layer neurons so as to calculate the predicted output. The regularized extreme learning machine based on genetic algorithm integrates the global search optimal ability of genetic algorithm and the strong learning and generalization ability of RE-ELM.

The experimental data (iron ore price and its various influencing factors) were divided into training set and test set, and the data was normalized to avoid large errors due to the large magnitude difference of data. This set of data has A input neurons and B hidden neurons. The activation function can select Hardlim function, Sigmoid function, Gaussian function and so on.

The learning process of bayesian genetic algorithm and regularization extreme learning machine with Bayesian Bayesian criteria :

step 1: The full-subset regression method is used to select the final variables that are eventually brought into the learning model, and the optimal variable selection is selected according to the BIC value.

Step 2: The population X is initialized to include m chromosomes, each of which includes A*B input weights and B thresholds, and uses the initial population as the first generation population.

Step 3: A genetic algorithm is used to find the optimal initial W and b of the RE-ELM algorithm. Each individual in the population contains all the weights and thresholds of an ELM network. Individuals calculate their fitness values through the fitness function and find the corresponding individuals with minimum fitness through selection, crossover and mutation operations. In this paper, the mean square error between the predicted value and the real value of the predicted sample is used as the individual fitness function.

Step 4: The optimal individual obtained by the genetic algorithm assigns RE-ELM initial weights and thresholds, and sets the number of hidden layer nodes to complete the establishment of the GA-RE-ELM model.

Step 5: Finally, using the test set sample to verify the GA-RE-ELM model and evaluate the effect.

III. EXPERIMENTAL PART

A. Data Sources

The basic data of this article is derived from (<http://www.tiekuangshi.com/tongji/ciopi.htm>) iron ore network, and choose the time period from October 18, 2013 to March 15, 2018, with a total of 1077 iron ore data. The data features include the daily opening price, highest price, lowest price, closing price, trading volume, turnover, open interest, and the settlement price of iron ore. There are 8 types of data, ie 1077 sets of data (sample number 8*1077).

B. Variable Selection

The nine variables of the opening price, highest price, lowest price, closing price, trading volume, turnover, open interest, settlement price of iron ore on days i-1, i-2, and the closing of day i price are used to be the full subset regression. According to the BIC criterion, the subset of variables that minimizes the BIC value is selected as the final variable of the subsequent model.

TABLE I. SUBSETS OF OPTIMAL PREDICTOR VARIABLES FOR DIFFERENT MODEL SIZES

model	OP	HP	Lp	V	T	OI	SP	Ycp	Bycp
1							*		
2	*						*		
3	*	*	*						
4	*	*	*				*		
5	*	*	*				*	*	
6	*	*	*	*			*	*	
7	*	*	*	*	*		*	*	
8	*	*	*	*	*	*	*	*	
9	*	*	*	*	*	*	*	*	*

Note OP: Opening price, HP: Highest price, LP: Lowest price, V: Volume, T: Turnover, OI: Open interest, SP: Settlement price, Ycp: Yesterday's closing price, Bycp: The day before yesterday closing price.

Table.1 shows the output of the optimal predictor subset for different model sizes. The asterisk indicates that the corresponding variables in the column are contained in the corresponding model in the row. Since the iron ore price forecasting model has 9 related variables, the optimal 9-variable model is output. The 9 variables were used to screen the results and 9 models were fitted to the optimal 9 variable

models. Table 2 shows that the BIC value of model 5 is the lowest, so we choose the variable subset corresponding to this model as the variables of the follow-up learning model. They are the opening price, the highest price, the lowest price, the settlement price, and the closing of the previous day of iron ore price.

TABLE II. BIC VALUES FOR DIFFERENT VARIABLE SUBSETS

Model	1	2	3	4	5
BIC	-7180.312	-7579.4	-7676.008	-7683.124	-7683.16
Model	6	7	8	9	
BIC	-7679.58	-7673.14	-7666.192	-7659.211	

C. Sample Selection and Parameter Selection

The samples are divided into two parts of the training set and test set. Among them, the sample data of the previous 1026 days is used as a training set, and the remaining 50 days of sample data are used as a test set. According to the variable selection in the previous paragraph, we use the five indicators of the opening price, the highest price, the lowest price, the settlement price, and the closing price of the i -1 day as the input of the model, and the closing price of the i-day as the forecast output of the model.

In the following experiment, the input is normalized to [-1,1], weights and thresholds are in [-1,1], output is normalized to [-1,1], and the excitation function of hidden layer neurons Use the Sigmoid function. The BIC-GA-RELM model completely preserves the adjustable parameters in the GA and ELM algorithms. The selection of these parameters is based on the relevant literature and is adjusted through multiple experiments. The recommended values are shown in Table 3.

TABLE III. GA-RE-ELM PARAMETERS SETTING TABLE

No.	parameter	value
1	Population size	30
2	Max number of iterations	150
3	Crossover probability	10
4	Variation probability	0.75
5	Objective function	0.01
6	Termination condition	Max number of iterations

As a kind of single hidden layer feedforward neural network, how to determine the number of neurons in the hidden layer in the ELM algorithm is critical. If the number of neurons in the hidden layer is too small, the network will be under-fitted, and if the number of neurons is too large, the network will be over-fitted, which will result in a good fit of the training set and a poor fitting effect on the verification set. Therefore, it is very important to determine the number of hidden layer nodes in the ELM network. This paper tries to find the optimal number of nodes by continuously experimenting and changing the number of hidden layer nodes in the network. Select the range of nodes in the hidden layer of the ELM network as [1,200] and calculate the mean squared error of the corresponding test set. The result is shown in Fig.1.

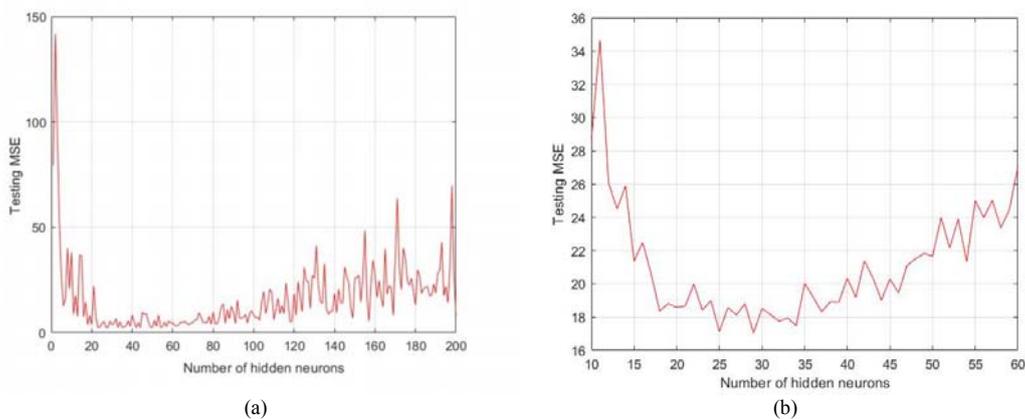


FIGURE I. ELM NETWORK HIDDEN NODE NUMBER TEST PROCESS

It can be seen from Figure 1 that when the number of hidden layer nodes in the network is too small (under-fitting), or too much (over-fitting), the mean square error of the test samples is large, which is consistent with the previous analysis. In order to more accurately select the optimal number of hidden layer nodes, the scope of the hidden layer nodes is further selected as [10, 60]. Perform 100 tests on each node and obtain the mean square error of the test samples, as shown in Figure 1(b). The experimental results show that when the number of hidden layer nodes is 29, the mean square error of the test samples is the smallest, so the final determination of the PM2.5 prediction model in the hidden layer node parameters is set to 29.

D. Results and Analysis

In this paper, the original limit learning machine, the regularized limit learning machine of the BIC rule, and the regularized limit learning machine optimized by the genetic algorithm of the BIC rule are used to predict the price of iron ore data samples. After many experiments, the implicitness of the original ELM model is determined. The number of layer neurons is 29. The “Sigmoid” function is selected for the hidden layer excitation function; the regularization coefficient C is 660 in the regularization limit learning machine model, the number of hidden layer nodes is set to 29, and the hidden layer excitation function selects “Sigmoid” function. The parameters of the regularized extreme learning machine optimized by the genetic algorithm of Bayesian information criterion are shown in Table 3.

The parameters in Table 3 are used as the parameters of the GA-ELM model and are brought into the training sample. It can be seen that the BIC-GA-RELM prediction model can be obtained through 150 evolutionary calculations by optimizing the RE-ELM evolution results by using the genetic algorithm in Fig.2. A stable fitness value for optimal fitness.

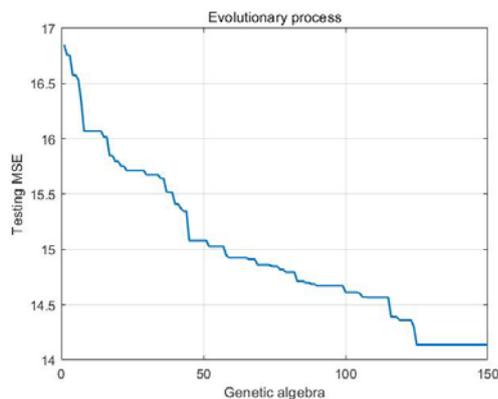


FIGURE II. ERROR EVOLUTION CURVE OF GA

As can be seen from Table 4, the extreme learning machine has strong generalization ability, and the three models have better prediction results than the training set on the test set. The predictive mean-square error on the BIC-GA-RELM model training set decreased by 78.43% and 50.10%, respectively, compared with ELM and BIC-RELM, and the root mean square error decreased by 53.56% and 29.36%, respectively, the average percentage error decreased by 61.61%, 36.76%, the

average absolute error decreased by 44.90% and 14.72%, respectively. Compared with ELM and BIC-RELM, the mean square error of predictions on the BIC-GA-RELM model test set was reduced by 83.86% and 45.43%, respectively, and the root mean square error was reduced by 59.82% and 26.13%, respectively, the average percentage error was reduced by 67.86 and 28.95%, the average absolute error decreased by 59.29%, 8.6%.

TABLE IV. PREDICTION RESULTS COMPARISON OF DIFFERENT MODELS OF TRAINING SETS

		ELM	BIC-RELM	BIC-GA-RELM
MSE	Tra	51.1029	22.0915	11.0231
	Test	24.7793	7.3306	4.0003
RMSE	Tra	7.1486	4.7002	3.3201
	Test	4.9779	2.7075	2.0001
MAPE	Tra	0.0112	0.0068	0.0043
	Test	0.0084	0.0038	0.0027
MAE	Tra	5.4111	3.4961	2.9816
	Test	4.4585	1.9860	1.815

Note: Tra means Training

IV. CONCLUSIONS

In summary, we combines the BIC criterion, genetic algorithm and regularized extreme learning machine, proposes a BIC-GA-RELM model, and applies it to the prediction of iron ore price. In the selection of ELM hidden layer nodes, the number of better hidden layer nodes is determined by repeating the experiment several times. In the learning model, according to the size of the BIC value, select the optimal subset of variables to determine the iron ore price forecasting indicators, the regularization term can control the complexity of the ELM learning model, and then use the genetic algorithm for the BIC-RELM model. The input layer weight matrix and the hidden layer threshold matrix are optimized to reduce the influence of the randomness of the input layer weight matrix and the hidden layer threshold matrix on the prediction accuracy, and improve the prediction accuracy. Therefore, the introduction of BIC criteria, genetic algorithms, and regularization terms can control the complexity of the model and achieve a global optimum.

ACKNOWLEDGEMENT

This research was funded by the National Natural Science Foundation of China under Grants 61375063, 61773404, 11301549 and 11271378, and in part supported by the Graduate Student Innovation Foundation of Central South University (2018zzts322).

REFERENCES

- [1] Shao L, Xu Z, Zhang S. A Research of the Factors that Influence the Iron Ore Price under New Market Patterns [J]. Management Review, 2018.
- [2] Wårell L. An analysis of iron ore prices during the latest commodity boom [J]. Mineral Economics, 2018(2):1-14.

- [3] Shao Liiuo, Xu Zihua, Zhang Shiyi. Research on the Influencing Factors of Iron Ore Price under the New Market Pattern [J]. *Management Review*, 2018(2):13-24.
- [4] Zhong Meirui, Pei Jieyu, Huang Jianbai, et al. Non-linear effects of non-ferrous metal price fluctuations based on MSVAR model [J]. *Chinese Journal of Management Science*, 2016, V24 (4):45-53.
- [5] Li Hua, Dong Yuanyuan, Wang Bin. The Influencing Factors and Empirical Analysis of the Price Changes of Imported Iron Ore in China [J]. *Statistics and Decision*, 2013(10): 124-127.
- [6] Ye Haiyan, Li Jin. Analysis of the Influencing Factors of Imported Iron Ore Prices in China [J]. *Exploration of Economic Issues*, 2012(10):119-122.
- [7] Yang Z, Li C, Lian L, et al. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods [J]. *Applied Energy*, 2017, 190:291-305.
- [8] Hou A, Suardi S. A nonparametric GARCH model of crude oil price return volatility [J]. *Energy Economics*, 2012, 34(2):618-626.
- [9] Li Z, Sun J, Wang S. An information diffusion-based model of oil futures price [J]. *Energy Economics*, 2013, 36(3):518-525.
- [10] Morana C. A semiparametric approach to short-term oil price forecasting [J]. *Energy Economics*, 2004, 23(3):325-338.
- [11] 11. Gülen S G. Efficiency in the crude oil futures market [J]. *Journal of Energy Finance & Development*, 1998, 3(1):13-21.
- [12] Binner J M, Kendall G, Chen S H, et al. A COMPARISON OF VAR AND NEURAL NETWORKS WITH GENETIC ALGORITHM IN FORECASTING PRICE OF OIL[C]International Conference on Artificial Intelligence, Ic-Ai '03, June 23 - 26, 2003, Las Vegas, Nevada, Usa, Volume. DBLP, 2004:62-67.
- [13] Lin A. Prediction of international crude oil futures price based on GM (1,1)[C] IEEE International Conference on Grey Systems and Intelligent Services. IEEE, 2009:692-696.
- [14] Shi M J. The Application of ARIMA Model in Investing Forecast in Fixed Assets of Shanghai [J]. *Application of Statistics & Management*, 2005.
- [15] Zhu Xi, Quan Jiye. Research on Price Mechanism and Regulation Strategies of Iron Ore International Trade [J]. *China Soft Science*, 2012(3):49-59.
- [16] Zhu H, Hang-Fei H E, Wang C, et al. Decomposition laws of tungsten prices fluctuation since 1900 and its applications [J]. *Transactions of Nonferrous Metals Society of China (English Edition)*, 2013, 23(9):2807-2816.
- [17] Zhu Xi, Zou Ping. The Long-term Law of the Iron and Steel Industry Technology Platform to Promote International Transfer of Industry [J]. *Scientific Research Management*, 2016, 37(5):1-9.
- [18] Kulkarni S, Haidar I. Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices[J]. *Papers*, 2009.
- [19] Alessandro Lanza, Matteo Manera, Giovannini M. Modeling and forecasting cointegrated relationships among heavy oil and product prices[J]. *Energy Economics*, 2005, 27(6):831-848.
- [20] Murat A, Tokat E. Forecasting oil price movements with crack spread futures[J]. *Energy Economics*, 2009, 31(1):85-90.
- [21] Lean Yu, Shouyang Wang, Kin Keung Lai. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm [J]. *Energy Economics*, 2008, 30(5):2623-2635.
- [22] Abdullah S N, Zeng X. Machine learning approach for crude oil price prediction with Artificial Neural Networks-Quantitative (ANN-Q) model[C] International Joint Conference on Neural Networks. IEEE, 2010:1-8.
- [23] Tang L, Yu L, Wang S, et al. A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting [J]. *Applied Energy*, 2012, 93(1):432-443.
- [24] Kamyar Movagharnjad, Bahman Mehdizadeh, Morteza Banihashemi, et al. Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network[J]. *Energy*, 2011, 36(7):3979-3984.
- [25] Chiroma H, Abdulkareem S, Herawan T. Evolutionary Neural Network model for West Texas Intermediate crude oil price prediction[J]. *Applied Energy*, 2015, 142:266-273.
- [26] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications[J]. *Neurocomputing*, 2006, 70(1):489-501.
- [27] Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification.[J]. *IEEE Transactions on Systems Man & Cybernetics Part B*, 2012, 42(2):513-529.
- [28] Chen Shaowei, Liu Guangfeng, Ye Shuai, et al. Research on Fault Diagnosis of Analog Circuits Based on Bat Algorithm Optimization ELM[J]. *Electronic Measurement Technology*, 2015, 38(2):138-141.
- [29] Wang Jiewen, Li Henan. Review of Particle Swarm Optimization Algorithms [J]. 2009, 30(2):22-27.
- [30] Zhang Weihui, Huang Nantian, Yang Jincheng, et al. Power Quality Disturbance Signal Classification Based on Generalized S-Transform and DE-ELM[J]. *Electric Measurement & Instrumentation*, 2016, 53(20):50-55.
- [31] Zhou Ming, Sun Shudong. Principle and Application of Genetic Algorithm: THEORY AND APPLICATIONS [M]. National Defense Industry Press, 1999.
- [32] Burnham, Kenneth P, Anderson, David R. Multimodel Inference: Understanding AIC and BIC in Model Selection.[J]. *Sociological Methods & Research*, 2004, 33(33):261-304.
- [33] Mei Yi, Sun Quanlong, Yu Lihua, et al. Grain Size Prediction of Aluminum Alloy Die Castings Based on GA-ELM[J]. *Acta Metallurgica Sinica*, 2017, 53(9): 1125-1132
- [34] Zhang Xiayang, Huang Qiqing, Yin Zhiping, et al. Identification of Flight Load Parameters Based on GA-ELM[J]. *Aeronautical Engineering Progress*, 2014(4): 497-501.
- [35] Huang G, Huang G B, Song S, et al. Trends in extreme learning machines: A review[J]. *Neural Networks the Official Journal of the International Neural Network Society*, 2015, 61(C):32.