Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

The stock trend prediction based on GA-PSO-ELM

Yingjie Wang *, Le Li and Yuqian Liu

School of artificial intelligence, Zhengzhou Railway vocational & technical college, Zhengzhou 450000, China.

Abstract: For the precision of stock trend prediction by the ELM (extreme learning machine) is low, a stock trend forecasting model based on ELM improved by genetic algorithm (GA) and particle swarm optimization (PSO) is proposed. Nine indicators constructed by stock index historical data are selected as input features of the predicting model. Then, the global search of GA and the local search of PSO are used to establish the prediction model based on GA-PSO-ELM; Finally, evaluate the prediction model with the prediction accuracy and compare the GA-PSO-ELM with ELM, OS-ELM, I-ELM. The results show that the GA-PSO-ELM model has an average prediction accuracy of 66.19%, which is more 10.22%, 15.97%, 12.36% than that of ELM, I-ELM, OS-ELM. So, the GA-PSO-ELM has higher prediction accuracy and better generalization ability in stock trend forecasting.
Keywords: Stock trend prediction, genetic algorithm, particle swarm optimization, extreme learning machine(ELM).

---

1. INTRODUCTION

Stock trend prediction is the focus of financial research, a large number of studies show that the stock market has the characteristics of nonlinear, non-stationary and high noise, the stock market is a nonlinear system, the traditional linear model does not satisfy the market research[1,2]. There are many researches on stock trend prediction, the methods mainly are neural network[3], SVM, chaos theory, fractal theory and so on, but the neural network prediction training speed is slow, easy to fall into local optimum, and the number of hidden nodes set by people greatly affected the prediction accuracy [4,5]. For support vector machine prediction methods, a large number of experiments are needed to tune the kernel function and error control parameters, which is time-consuming and uncertain in accuracy[6].

ELM (Extreme learning machine)[7] is a new type of neural network algorithm proposed by Nanyang Technology University Professor Huang et al in 2006, which is a single-hidden layer feed-forward neural networks. The extreme learning machine has been widely used in fault diagnosis, image segmentation, data mining, automatic control and other fields[8-10]. In this algorithm, the input weights were randomly selected, the output weights are determined by the least square method, which greatly improves the network training speed and generalization ability [11-12]. However, in solving the problem of gradient descent,due to the randomness of the parameters of the hidden layer neurons,

the ELM is easy falling into local optimal value. In order to solve these problems, Y. proposed DE-ELM[13], but the method also has this problems; Chen. put forward the bat algorithm
to optimize ELM, which may fall into the local optimal[14], and the ELM is easy affected by the number of hidden nodes,the accuracy of the training results is closely related with selection of hidden nodes, which is random.

2. GA-PSO-ELM

2.1 Extreme Learning Machine

Extreme learning machine (ELM) randomly initializtion input weights and bias weights, and the hidden layer neurons only calculate the output weights, which making the learning speed of ELM increase the thousands of times than the traditional BP network, support vector machine learning algorithm[15-16], and has better generalization performance. The mathematical model of ELM as shown in Fig.1:

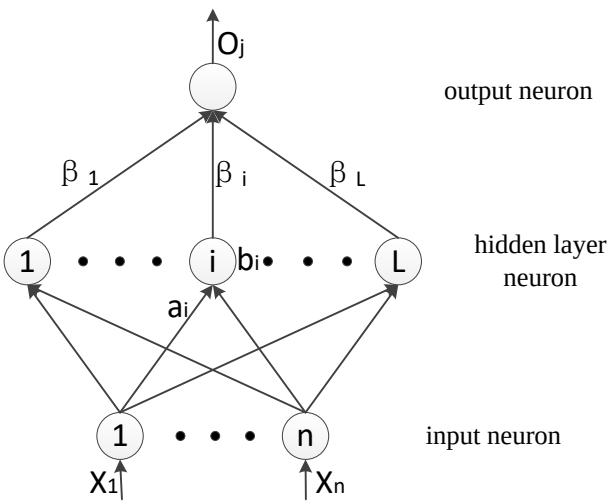

Fig.1 The mathematical model of ELM

Given any samples $\{(x_i, t_i)\}_{i=1}^{N} \subset R^n \times R^m$, where $x_i = [x_{i,1}, x_{i,2}, \cdots, x_{i,n}]^T \in R^n$,$t_i = [t_{i,1}, t_{i,2}, \cdots, t_{i,n}]^T \in R^m$, for a hidden layer network with L hidden layer nodes, the computational steps are as follows:

1) Randomly selecting the hidden layer node parameters $(a_i, b_i), i = 1, \ldots, L$, $a_i$is ith input weights of hidden layer neuron, b_i is threshold value of hidden layer neuron.

2) Calculating the hidden layer node output matrix $H = g(W_i, b_i, x_i)$ by formula(1):

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_n) \end{bmatrix} = \begin{bmatrix} g(W_1 \cdot x_1 + b_1) & \cdots & g(W_L \cdot x_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(W_1 \cdot x_n + b_1) & \cdots & g(W_L \cdot x_n + b_L) \end{bmatrix}_{n \times L} \tag{1}$$

Where $W_i = [a_{i,1}, a_{i,2}, \cdots, a_{i,n}]^T$is the input weight vector, $g(x)$ is the hidden layer node activation function, is nonlinear functions, and it can be Hardlim, Sigmoid, Gaussian and so on.

3) Calculating the output layer weights β , when the hidden layer node parameters (a_i, b_i) is determined randomly,the hidden layer output matrix H is determined, and the β can be calculated by:

$$\beta = H^{+} \cdot T, \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times m}, T = \begin{bmatrix} t_1^T \\ \vdots \\ t_n^T \end{bmatrix}_{n\times m} \tag{2}$$

Where H^{+} is the Left pseudo inverse matrix of H,T is the target output,and $T = \{t_j\}_{j=1}^{n}$.

4) Calculating the output value O_j。 When the error ($|O_j - T_j|$) less than a preset constant ε,the ELM can close training samples. The formulate as follows:

$$O_j = \sum_{i=1}^{L} \beta_i \, g(W_i, b_i, x_i), |O_j - T_j| \le \varepsilon, j = 1, \cdots, n \tag{3}$$

Where T_j and O_j is the actual and predicted output value of jth group data respectively.

5) Getting output error by:

$$E_{(W_i, b_i)} = \sqrt{\sum_{j=1}^{n} (O_j - T_j)^2} \tag{4}$$

2.2 GA-PSO-ELM Algorithm

In order to improve the prediction accuracy of limit learning machine, this paper proposed an improved extreme learning machine based on genetic algorithm and particle swarm optimization(GA-PSO-ELM).In this algorithm, mapping the input weights and hidden layer nodes threshold of the ELM to each chromosome genes of population GA, and selecting the best chromosome constitute the elite group by the global search ability of GA algorithm; and then selecting the best chromosome by the local search ability of PSO as the input weights and thresholds of ELM; finally calculating the ELM hidden layer neuron output weights by the least square method, and getting the predictive output. This algorithm transformed the problem of seeking weight and threshold into the problem of finding the optimal chromosome, which made full use of the local search ability of GA and global search ability of PSO, and combined with the strong learning ability of ELM, achieved a good prediction effect.

GA-PSO-ELM learning process:

Suppose the training data N has A afferent neurons and B hidden layer neurons,selecting the activation function Hardlim, Sigmoid, Gaussian and other activation function. The steps as follows:

Step1: Initial population

Initial population X into N subgroups randomly,denoted by $GA_i, i = 1,2,\cdots,N$. Each subgroup includes m chromosomes, and each chromosome x_i includes A input weights and B thresholds. The initial population is considered as the first generation, and GA_i can be get by:

$$GA_i = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1A} & b_{11} & \cdots & b_{1B} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mA} & b_{m1} & \cdots & b_{mB} \end{bmatrix} \tag{5}$$

Where, a_{kg} is input weight,b_{kh} is hidden layer neuron threshold,a_{kg} and b_{kh} of the initial population are obtained randomly, where, $k = 1,2,\cdots,m; g = 1,2,\cdots,A; h = 1,2,\cdots,B$.

Step2: Calculate the chromosome fitness

Calculating the chromosomes fitness of each subgroup, getting the hidden layer neuron output weights β by formula(2),then the fitness function fit is:

$$\text{fit} = \sqrt{\sum_{j=1}^{n}(O_j - T_j)^2} = \sqrt{\sum_{j=1}^{N}\sum_{i=1}^{B}(\beta_i g(W_i, b_i, x_i) - T_i)^2} \tag{6}$$

Step3: Selection

The tournament selection method is used to select the optimal chromosome, and the equation is as follows:

$$x_i^* = \begin{cases} x_j & (f_{x_i} < f_{x_j}) \\ x_i & (f_{x_i} > f_{x_j}) \end{cases} \tag{7}$$

Where,f_{x_i} is fitness value of x_i ,$i, j = 1,2,\cdots,N$,x_i^* is next generation population, selecting the highest fitness between two chromosomes enter into the next generation every time. The convergence rate of this method is fast, and there is a large probability to ensure that the optimal individual is chosen, and the worst individual is eliminated.

Step4: Crossover

Whether the selected individuals are crossed is determined by cross probability p_i, the individuals will be crossed if the random number is greater than p_i. In this paper, the p_i is 0.8,the crossed formula is:

$$\begin{cases} x_i = \text{rand}*x_{i+1}^* + (1\text{-rand})x_i^* \\ x_{i+1} = \text{rand}*x_i^* + (1\text{-rand})x_{i+1}^* \end{cases} \tag{8}$$

Where $rand$ is random numbers subject to uniform distribution from 0 to 1. Examining the two crossed chromosomes, if they are not viable, they will be recrossed.

Step5: Mutation

Inconsistent variation was used, according to mutation rate p_0 deal each dimension of all individuals, the formula is as follows:

$$x_i = \begin{cases} x_i + (X_{max}^i - x_i)*(\text{rand}*(1\text{-}t/T))^b & p_0 > 0.5 \\ x_i - (x_i - X_{min}^i)*(\text{rand}*(1\text{-}t/T))^b & p_0 \le 0.5 \end{cases} \tag{9}$$

Where,X_{max}^i is upper bound of x_i,X_{min}^i is lower bound of x_i,t is current iterations,T is maximum iterations,b is a parameter that determine the degree of non conformance variability, generally is from 2 to 5,the smaller the value is, the larger the variance is. Setting b is 2 in paper,making the range of variation greater, using global search capability of genetic algorithm.

Step6: initialize elite groups

Calculating the fitness value of the compiled individual to find the current best individual; the optimal individuals of each subgroup constitute the elite group (the particles of PSO),initializing particle speed, in this paper, the initialization rate of each dimension of a particle swarm is set from 0 to V_{max} , decreasing the velocity dimension,greatly reducing the degree of freedom of particles, and this is easy to achieve the purpose of local search.

Step7: Select input weights and thresholds

The standard particle swarm algorithm (PSO) is used for local search of the initialized elite, selecting the optimal chromosome if the stopping criterion is satisfied after some algebra; if it is not satisfied, exchanging the chromosome limit values random selected from elites with that from subgroup, and recycling for step 2, until the stopping criterion is satisfied, and the optimal chromosome is selected as the input weight and threshold.

3. STOCK TREND FORECAST

3.1 Data Processing

1) Original data

This paper takes Shanghai stock index from 2000.1.4 to 2015.12.30 as the original data, including each trading day's ceiling price, bottom price, opening price and closing price and other trading information, and the data are divided into training set and test set according to the time, as shown in Table 1. All transaction data is obtained free from the great wisdom 365 database on the Internet, software details as see on http://www.gw.com.cn/download.shtml.

Table 1. Experimental data set

data set	training set	test set
1	2000/1/4-2006/12/29	2007/1/4-2007/12/28
2	2008/1/2-2014/12/31	2015/1/5-2015/12/30

2) Input data

A lot of research in the financial field proved that the technical indicators are widely used in trend prediction and has good prediction effect. Based on this, this article construct feature index from the original data, as shown in table 2, which are used for prediction.

Table2. Description of characteristic index and its formula

Characteristic index	description	formula
AR	Popularity index , reflecting the popularity of the market	$\sum_{t}^{n+t}(H_t - O_t) / \sum_{t}^{n+t}(O_t - L_t)$
VR	The volume ratio, measure the heat of the stock price in terms of volume	$\frac{AVS_n - 1/2CVS_n}{BVS_n + 1/2CVS_n}$
ROC	Rate of change, analyze the trend of the stock price and its willingness to change	$\frac{C_t - C_{t-n}}{C_{t-n}} \times 100$
RSI	The relative strength index, shows the market strength according to the stock price decline	$100 - \frac{100}{1 + \frac{(\sum_{i=0}^{n-1} Up_{t-i} / n)}{(\sum_{i=0}^{n-1} Dw_{t-i} / n)}}$
K %	Random index k%, represents the relative height of the stock price over	$\frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} \times 100$

	a particular period	
PSY	Psychological line index, the mood index of investor's fluctuation of stock market	$\frac{D_n}{n}\times 100$
TRIX	The three index smoothed average line, reflects the long-term volatility of the stock price	$(TR_t - TR_{t-1}) / TR_{t-1} \times 100$
MTM	Momentum index , measures the momentum of stock price volatility and reveals the law of stock price reversal	$C_t - C_{t-n}$
MFI	Money flow indicators, measure momentum in trading volume and interest in investment	$MF = M_t \times A_n;$ $MFI = 100 - \frac{100}{1 + PMF / NMF}$

Where H_t L_t C_t is the highest, lowest, and closing prices of t days respectively;HH_{t-n} LL_{t-n} is respectively highest and lowest price in n days before ith day; $M_t = (H_t + L_t + C_t)/3$; TR is true volatility index,$TR = max(|H_t - L_t|, |H_t - C_{t-1}|, |C_{t-1} - L_t|)$;D_n is number of rise days in n days;AVS_n is volume of n rise days, BVS_n is the volume of n fall days,CVS_n is volume of the n neither up nor down days;A_n is the n days' turnover,MF is the money flow volume,PMF and NMF is positive and negative monetary flow respectively, if day's MF> yesterday's MF, the day's MF is today's PMF,and today's NMF is 0;on the contrary, the day's MF is today's NMF,and today's PMF is 0;

Before the network training, preprocessing the data, then use the formula (10) to standardize the data, and make the processed data as input.

$$x_i = \frac{x_i - \overline{x_i}}{\sigma_i} \tag{10}$$

Where $\overline{x_i}$ is the average value of ith index,σ_i is standard deviation of ith index.

3) Output data

Stock trend prediction is a classification problem, that is, if the test samples are classified as "1" ,it forecast to rise, if classified as "-1" ,it forecast to fall. Therefore, labeling the sample output data by the rules are as follows:

a) if the day's closing price is higher than the second day's in the sample data, it is marked "1";

b) if the day's closing price is lower than or equal to the second day's in the sample data, it is marked "-1";

Taking the processed data as the output of the GA-PSO-ELM, training the network, and then classifying the test set by trained network, the classification results are the prediction results, analyzing the prediction results and test set results and calculating prediction accuracy.

3.2 Forecast Process

Stock trend prediction based on GA-PSO-ELM firstly preprocessing the original data to determine the input and output data, then building and training the prediction model, classifying test set using the trained network, the specific steps are as follows:

Step1. Acquire source data and construct technical index;

Step2. Determine the training set and test set, the input and output data, and standardize the input data, classify the output data;

Step3. Train the prediction model based on GA-PSO-ELM, and get the input weights and thresholds;

Step4. Classify the test set, and get the predicted value and prediction accuracy;

The overall flow chart of the prediction system is shown in Fig. 2:

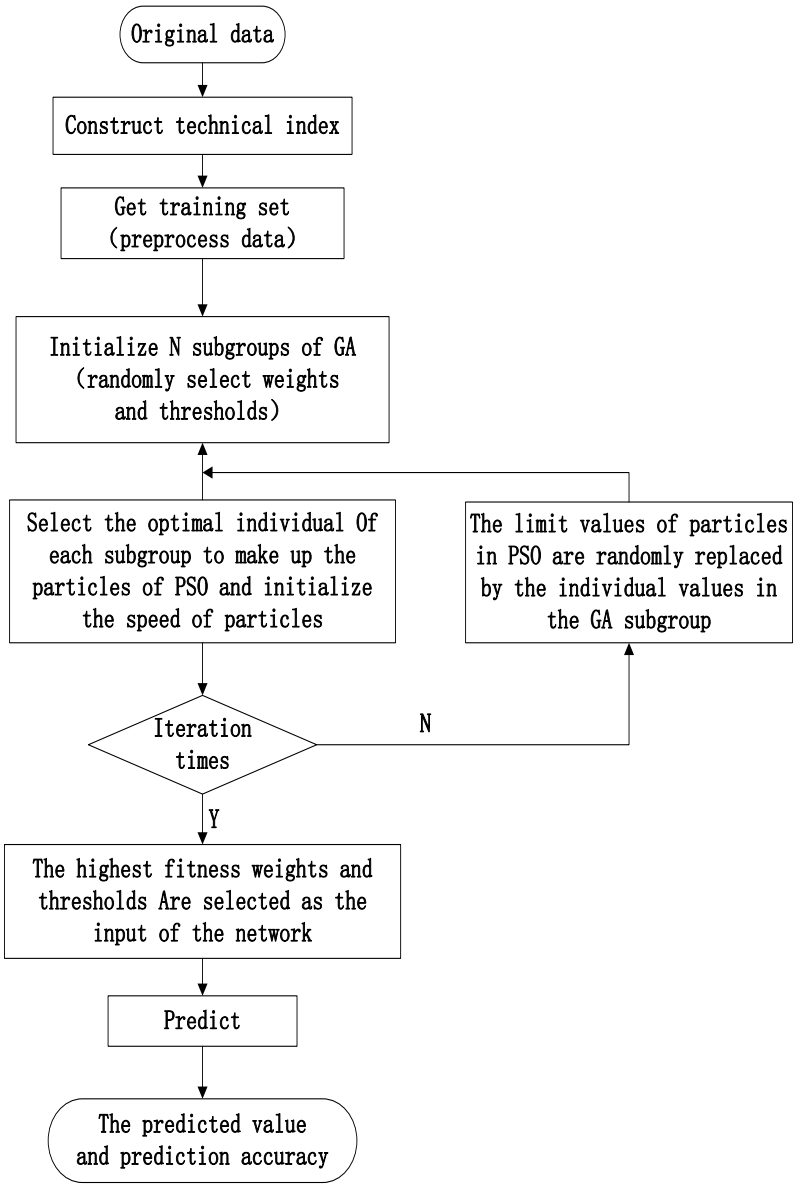

Fig.2 The overall flow chart of forecasting system

3.3 Result Analysis

1)Forecast result

Making the feature index in table 2 as input, using GA-PSO-ELM model to train the training set. The test data of data set 1 is 242 and set 2 is 243, Fig.3 and Fig.4 respectively shows the predicted value and the true value of set 1 and set 2.

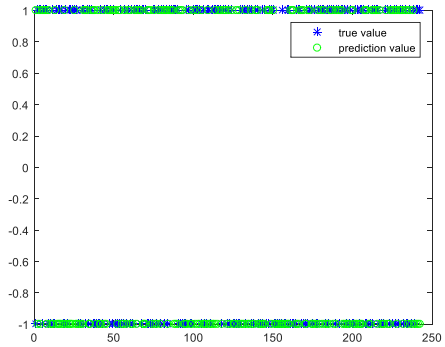

Fig.3 Predicted and true values of data set 1

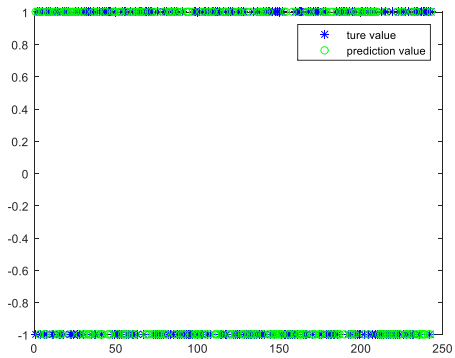

Fig.4 Predicted and true values of data set 2

Results: Fig.3 and Fig.4 clearly show the classification results, the accurate prediction of data set 1 is 169, and the accuracy rate is 69.83% , the accurate prediction of data set 2 is 152, and the accuracy rate is 62.55%. It proved the validity and accuracy of the stock trend prediction based on GA-PSO-ELM.

2) Contrast experiment

This paper compare the GA-PSO-ELM with ELM, OS-ELM, I-ELM in stock trend forecast to prove the validity of GA-PSO-ELM. But the number of hidden layer nodes has a certain influence on the prediction accuracy, in order to better analyze the accuracy of the prediction accuracy, take the number of hidden layer nodes from 1 to 50, plus 1 each time. Analyze the prediction accuracy of various models and the average prediction accuracy of the hidden layer nodes from 1 to 50. Fig.5 and Fig.6 respectively shows the predictions of various models for data set 1 and data set 2. Table 3 shows the average prediction accuracy of various models for data set 1 and data set 2.

Results: compared to other algorithms. As shown in Fig.5 and Fig.6, the prediction accuracy of GA-PSO-ELM is much higher, it indicates that the input weights and thresholds through the global search of GA and local search of PSO make the minimum error of ELM network, so that the prediction error decreases a lot; on the other hand, the improved ELM network has high accuracy for data 1 and data 2, and has a good generalization performance; the improved prediction accuracy of ELM although affected by the number of hidden nodes, but it is stabilized at more than 60%, it prove the prediction model of GA-PSO-ELM is relatively stable and effective.

As shown in table 3, the average accuracy base on GA-PSO-ELM of data set 1 reached 69.83%, more than ELM, OS-ELM, i-ELM 11.43%, 18.59%, 14.56%, the average accuracy base on GA-PSO-ELM of data set 2 reached 62.55%, more than ELM, OS-ELM, i-ELM 8.9%, 13.35%, 10.16%, it also prove the GA-PSO-ELM compared to other methods has high accuracy on stock trend prediction.

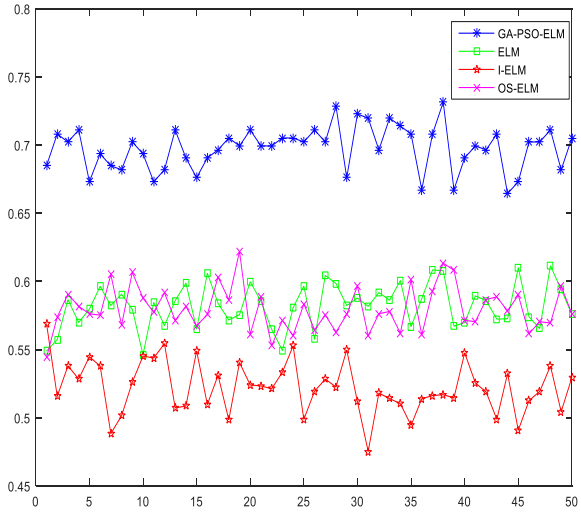

Fig.5 Prediction accuracy on various model of data set 1

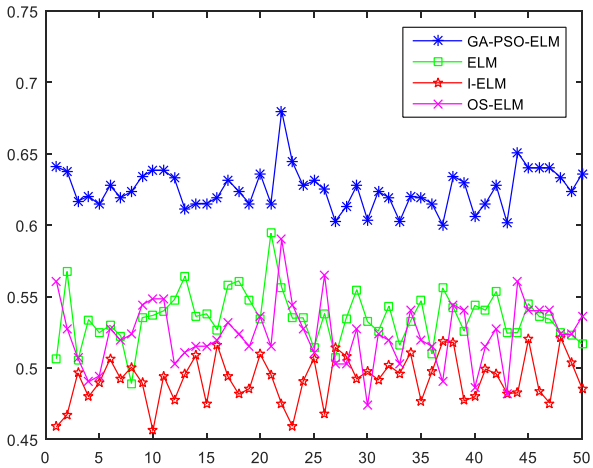

Fig. 6 Prediction accuracy on various model of data set 2

Table 3 Mean prediction accuracy on various model

data set	GA-PSO-ELM	ELM	I-ELM	OS-ELM
data set 1	0.6983	0.5840	0.5124	0.5527
data set 2	0.6255	0.5355	0.4920	0.5239

4. CONCLUSION

The stock market is a dynamic, complex nonlinear system, and the traditional linear model cannot describe the stock market comprehensively, lead to a certain gap with high prediction accuracy. This

paper presents a stock trend forecasting model of based on GA-PSO-ELM, which takes the standardized characteristic indexes as input of the network, improves ELM with GA and PSO algorithm, predicts the stock trend of the two data sets using the trained GA-PSO-ELM network, and compared with other models.

The experimental results show that the average prediction accuracy on GA-PSO-ELM model for data set 1 is about 69%, for data 2 is about 62%, it confirms the GA-PSO-ELM stock trend forecast is feasible and effective; the simulation results of GA-PSO-ELM and other methods are compared and confirmed GA-PSO-ELM has higher prediction accuracy, better generalization ability, and is relatively stable. The next step can be combined with feature selection to achieve higher prediction accuracy.

ACKNOWLEDGEMENTS

This work was supported by the research on talent training mode reconstruction of artificial intelligence technology service specialty under the background of industrial transformation (2020JG68).

REFERENCES

[1] J.Patel, S.Shah and Thakkar P. "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," Expert Systems with Applications, vol.42, no.1, pp. 259-268, 2015.

[2] S.Lahmiri. "A variational mode decompoisition approach for analysis and forecasting of economic and financial time series," Expert Systems with Applications, vol.55, pp.268-273, 2016.

[3] J Wang. "Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks," Neurocomputing, vol.156, pp.68-78, 2015;

[4] L Huibin , L Dandan ,S Haiyan. "Chaotic time series prediction based on PSO-BP neural network,"Computer Engineering and Applications, vol.51, no.2, pp.224-229, 2015.

[5]Guohong Y, Dezhi Y, Peili C. Study of stock price forecasting based on combination of ARIMA and RBF neural network. Computer engineering and applications, vol.49, no.18, pp. 245-248, 2013.

[6]Thenmozhi M , Chand GS. "Forecasting stock returns based on information transmission across global markets using support vector machines," Neural Computing and Applications, vol.27, no.4, pp.805-824, 2016.

[7] Guang-Bin H, Qin-Yu Z and Siew C.K. "Extreme learning machine: a new learning scheme of feedforward neural networks," Proceedings of International Joint Conference on Neural Networks, vol.2, no.2, pp.985-990, 2004.

[8]Guorui F, Lan Y and Xinpeng Z. "Dynamic adjustment of hidden node parameters for extreme learning machine," IEEE Transactions On Cybernetics, vol.5, no.2, pp.279-288, 2015.

[9]Shaobo L, Xia L and Jian F. "Is extreme learning machine feasible? a theoretical assessment(Part II)," IEEE Transactions On Neural Networks And Learning Systems, vol.26, no.1, pp.21-34, 2015.

[10]Sun X , Xu J and Jiang C. "Extreme learning machine for multi-label classification," Entropy, vol.18, no.6, pp.225, 2016.

[11]Alim S, Peijun D and Sicong L. "Ensemble extreme learning machines for hyperspectral image classification," IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, vol.7, no.4, pp.1060-1068, 2014.

[12]S Li, ZH You and H Guo. "Inverse-free extreme learning machine with optimal information updating," IEEE Transactions on Cybernetics, vol.46, no.5, pp. 1229-1241,2016.

[13]Yakoub B, Naif A and Farid M. "Differential evolution extreme learning machine for the classification of hyperspectral images," IEEE Geoscience And Remote Sensing Letters, vol.11, no.6, pp.1066-1070, 2014.

[14]Shaowei C,Guangfeng L and Shuai Z. "Research on fault diagnosis of analog circuits based on bat-ELM," Electron Measure Technology, vol.38, no.1, pp.138-141, 2015.

[15]H Zhong, C Miao and Z Shen. “Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings”. Neurocomputing, vol.128, no.5, pp.285-295, 2014.

[16] Salam AEU, Tola M and Mary S. “Application of SVM and ELM methods to predict location and magnitude leakage of pipelines on water distribution network,” International Journal of Advanced Computer Research, vol.19, no.5, pp. 2277-7970, 2015.