论文《Prediction and Data Analysis of Price Based on PSO with Extreme Learning Machine Algorithm and Particle Swarm Optimization》

---基于粒子群优化的极限学习机算法对股票价格的预测与分析（EI 检索证明）

Proceedings

2022 2nd International Signal Processing, Communications and Engineering Management Conference

ISPCEM 2022

Montreal, Canada
25-27 November 2022

IEEE COMPUTER SOCIETY
CONFERENCE
PUBLISHING
SERVICES

ISBN-13: 978-1-6654-9271-3
BMS Part# CFP22AA9-ART

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Editorial production by Cristina Ceballos
Cover art production by Hector Torres

IEEE COMPUTER SOCIETY
CONFERENCE
PUBLISHING
SERVICES

IEEE Computer Society
***Conference Publishing Services* (CPS)**
http://www.computer.org/cps

2022 2nd International Signal Processing, Communications and Engineering Management Conference (ISPCEM)

ISPCEM 2022

Table of Contents

2022 2nd International Signal Processing, Communications and Engineering Management Conference

Prediction and Data Analysis of Price Based on PSO with Extreme Learning Machine Algorithm and Particle Swarm Optimization

Huanyun Chen*,#
Private Hualian College
Guangzhou, China
Huanyunchen@qq.com
#These authors contributed equally.

Weiming Zhao#
Private Hualian College
Guangzhou, China
ncchzwm@126.com

Abstract—**The changes in the trading market are affected by many factors, and traditional forecasting methods are more and more difficult to meet people's needs. In order to improve the accuracy of prediction, this paper proposes an extreme learning machine algorithm based on particle swarm optimization for prediction and analysis. First, the fluctuation data of previous years are collected, and the closing price of the stock is determined as an experiment and simulated to establish a learning sample, and then the learning sample is input. Based on the online extreme learning machine model optimized by particle swarm optimization, the prediction model is established. The results show that the optimized extreme learning machine algorithm, compared with models such as ELM, PSO-BP and DE-ELM, has better surface prediction accuracy and fitting effect, and can more accurately describe the change trend, which is more efficient than other methods.**

Keywords—extreme learning machines, particle swarm optimization, forecasting models

I. Introduction

With the development of the national economy and the improvement of people's living standards, stocks are becoming more and more important as a means of financial management. Compared to other financial instruments such as funds, the high risk and high return characteristics of equities make people choose them more carefully and cautiously, and investors will analyse and forecast equities through a number of data on the stock market to make investment plans. Stock prices are influenced by a variety of factors and are highly variable and complex, and are closely related to the financial environment.

Extreme learning machines (ELM) are a class of machine learning methods based on feed-forward neural networks, which are widely used in various prediction problems[2] . Compared with traditional BP neural networks, the weights and thresholds in ELM algorithms can be generated randomly and do not need to be adjusted at the runtime of the algorithm, which makes learning and reaching a steady state fast and greatly reduces the training time. In recent years there have been many uses of extreme learning machines to learn algorithms for stock price prediction and analysis. Bingzhong Xiong [3] based a sparse Bayesian extreme learning machine (SBELM) that has the advantage of automatically selecting the number of nodes in the hidden layer. Xiaochun Liang and Xiaoyun Chen [4] based their extreme learning machine forecasting model (SVDELM) on singular value decomposition, which is more effective than the traditional extreme learning machine (ELM) in improving multivariate time series forecasting accuracy.

In this paper, in order to improve the accuracy of stock price prediction, based on the stock price data of Pudong Development Bank and Ping An Bank, this paper proposes a stock price prediction model based on particle swarm algorithm optimized learning machine, which has higher surface prediction accuracy and better fitting effect compared with ELM, PSO-BP and DE-ELM.

II. Data Acquisition

The Chinese stock market has grown rapidly in recent years, mainly in Shanghai and Shenzhen, and as a result the two regions have produced the most highly regarded SSE and SZSE indices respectively. Here we choose the more comprehensive CSI 300 index data as the data base for our study, and Table I shows the data structure of one of the stocks. However, of the thousands of companies that have been listed, some have been delisted and some have been suspended. The stocks selected need to be not only representative of the market, but also stable and long-lasting.

In Table I, trade_date is the date in the format YYYYMMDD, open is the opening price, close is the closing price, low is the lowest value of the stock for the day, and high is the peak. pre_close is the previous day's closing price, pct_chg is the percentage change of the current period's data over the previous period's data , vol is the volume indicator.

In summary, we decided to choose to use the data of Pudong Development Bank and Ping An Bank for a total of 588 trading days over a five-year period from 18 January 2020 to 18 July 2022 to construct the dataset. The data was obtained through python's tushare financial data interface package with five basic data of the selected stocks: opening price, closing price, high price, low price and volume [5-6]. Some of the data for Pudong Development Bank are shown in Table II.

III. Extreme Learning Machine Algorithm Based on Particle Swarm Optimization

A. Extreme Learning Machine Algorithms

Extreme learning machine algorithms can be discussed as learning strategies or neural network constructions. The standard extreme learning machine algorithm uses a single-layer feed-forward neural network structure, consisting of an input layer, an implicit layer and an output layer, where the output function of the implicit layer has the following

978-1-6654-9271-3/22/$31.00 ©2022 IEEE
DOI 10.1109/ISPCEM57418.2022.00056

definition.

TABLE I. PING AN BANK OF CHINA STOCK PRICE CASE DATA

trade_date	open	high	low	close	pre_close	change	pct_chg	vol	amount
20220715	13.48	13.56	13.22	13.24	13.37	-0.13	-0.9723	1598637.4	2139915.274
20220714	13.7	13.71	13.29	13.37	13.97	-0.6	-4.2949	2632120.87	3534578.702
20220713	14.42	14.47	13.89	13.97	14.42	-0.45	-3.1207	2175996.06	3056913.118
20220712	14.41	14.73	14.38	14.42	14.46	-0.04	-0.2766	742882.42	1078431.99
20220711	14.5	14.57	14.36	14.46	14.54	-0.08	-0.5502	611691.94	883613.754
...	...	...	...	...	...	...	...	...	...
20200124	14.66	15.08	14.5	14.64	14.65	-0.01	-0.07	2591292.12	3838734.922
20200123	14.36	14.9	14.33	14.65	14.44	0.21	1.45	2388791.7	3492462.005
20200122	14.6	14.94	14.43	14.44	14.8	-0.36	-2.43	2073867.14	3032799.07
20200119	14.8	15.13	14.68	14.8	14.72	0.08	0.54	2571146.69	3832857.44
20200118	14.4	14.72	14.28	14.72	14.23	0.49	3.44	2148026.8	3120455.742

TABLE II. SAMPLE DATA FOR SHANGHAI PUDONG DEVELOPMENT BANK TOCK PRICES

trade_date	open	high	low	close	pre_close	change	pct_chg	vol	amount
20220715	7.76	7.8	7.65	7.67	7.79	-0.12	-1.5404	471276.32	364457.391
20220714	7.84	7.86	7.75	7.79	7.85	-0.06	-0.7643	405658.78	315441.409
20220713	7.93	7.94	7.85	7.85	7.91	-0.06	-0.7585	305065.23	240374.854
20220712	7.86	7.94	7.86	7.91	7.87	0.04	0.5083	198897.98	157433.752
20220711	7.87	7.9	7.86	7.87	7.88	-0.01	-0.1269	161954.72	127636.557
...	...	...	...	...	...	...	...	...	...
20200124	12.97	14	12.89	13.57	12.9	0.67	5.19	3796536.29	5099364.114
20200123	12.75	12.92	12.68	12.9	12.77	0.13	1.02	2178274.57	2791828.653
20200122	12.93	13.15	12.71	12.77	13.24	-0.47	-3.55	2725138.01	3514775.144
20200119	13.33	13.61	13.18	13.24	13.24	0	0	1716488.71	2300220.633
20200118	13.14	13.25	13.05	13.24	13.1	0.14	1.07	1116080.76	1465159.094

$$f_L = \sum_{i=1}^{l} \beta_i h_i(x) = h(x)\beta, \quad (1)$$

In equation (1), x is the output of the neural network, L is the total number of samples, β is the output weight, and $h(x)$ is the feature mapping or excitation function, which can map the data of the input layer from the original space to the feature space of the extreme learning machine algorithm.

$$\mathbf{h(x)} = \mathbf{G(a_i, b_i, x)}, \quad (2)$$

In equation (2), a_i and b_i are the parameters of the feature mapping, which are used as node parameters in the extreme learning machine algorithm, where a_i is the input weight. The connection between each layer is done using the feature mapping function, and the information from the input layer is processed through the implicit layer and passed to the output layer, which then derives the calculated values according to the mapping function.

During the training process of the extreme learning machine, the weights and biases of the implicit layer are often randomly generated or artificially given, and do not need to be updated. The training process is completed by calculating the weights of the output layer.

The steps are as follows.

(1) Let there be n neurons input terminals in the input layer, m neurons in the output layer, and l neurons in the hidden layer. Denote w as the connection weights between the input layer and the hidden layer, and w correspondingly as a matrix of $l \times n$. The elements of the weight matrix are generally random numbers between [-1,1].

$$\mathbf{w} = \begin{bmatrix} \mathbf{w_{11}} & \cdots & \boldsymbol{\beta}_{\mathbf{1n}} \\ \vdots & \ddots & \vdots \\ \mathbf{w_{l1}} & \cdots & \boldsymbol{\beta}_{\mathbf{ln}} \end{bmatrix}_{\mathbf{l \times n}}, \quad (3)$$

(2) Noting that b is the bias of the hidden layer, corresponding to the $l \times l \times l$ neuron nodes of the hidden layer, the size of the bias vector b is the $l \times 1$ column vector. The elements of the bias vector are generally random numbers between [0,1].

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \end{bmatrix}_{l \times 1}, \quad (4)$$

(3) Noting β as the connection weights of the hidden layer and the output layer, the size is a matrix of $l \times m$. When the neuron nodes of the output layer are $m = 1$, the training yields a single-output limit learning machine model. With the initial implied layer weights $\boldsymbol{\omega}$ and bias b, the output layer weight matrix β can be solved for to obtain the trained limit learning

machine model [7].

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{l1} & \beta_{l2} & \cdots & \beta_{lm} \end{bmatrix}, \tag{5}$$

(4) Feature mapping through the activation function *G(x)* gives the output value of the limit learning machine as

$$f_L(x) = \sum_{i=1}^{L} \beta_i G(w_i \bullet x_i + b_i) = \sum_{i=1}^{L} \beta_i h(x_i), \tag{6}$$

(5) Let T be the desired output value, so that it is only necessary to find the appropriate β such that the value of the error function is minimum or close to 0.

$$\|H\beta - T\|^2 = 0, \tag{7}$$

The inverse of the above equation is solved for the weights β of the output layer to obtain a trained extreme learning machine model.

B. Particle swarm optimization

This paper combines the update of particle positions and the limit learning machine algorithm, resulting in an optimised stock price prediction method.

Vim(t) represents the initial position of the *i-th* particle, H_{j1} is regarded as the output of the hidden layer node j_1, I_h is coming from the input layer node *h*; α and $\beta$are the particle learning rate; $r_1$and r_2 are 2 random values, generally taken between 0 and 1; *t* is the number of current training iterations; [] denotes the rounding function.

Initialize the particle swarm, including the particle size *N* and the position vector and velocity vector of each particle, the individual extreme value and global optimum value of each particle, iteration error accuracy as ε, constant coefficients c_1 and c_2, maximum inertia weight as η_{max}, minimum inertia weight as η_{min}, maximum velocity as $v_{max}$and maximum number of iterations, etc.

The updated position of the *i-th* particle is as follows.

$$v_{im}(t+1) = \eta v_{im}(t) + c_1 r_1 (p_{im} - x_{im}(t)) + c_2 r_2 \left(p_{gm} - x_{im}(t)\right) m = 1,2,\cdots n, \tag{8}$$

$$x_{im}(t+1) = \begin{cases} x_{im}(t) + v_{im}(t+1) + \alpha\delta_k H_{j_1}, \\ \quad if m \in (0, n_o n_h] \\ x_{im}(t) + v_{im}(t+1) + \alpha\delta_{j_2} I_h, \\ \quad if m \in (n_o n_h, n_o n_h + n_h n_i] \\ x_{im}(t) + v_{im}(t+1) + \beta\delta_{k_1}, \\ \quad if m \in (n_o n_h + n_h n_i, n_o n_h + n_h n_i + n_o] \\ x_{im}(t) + v_{im}(t+1) + \beta\delta_{j_3}, \\ \quad if m \in (n_o n_h + n_h n_i + n_o, n_o n_h + n_h n_i + n_h] \end{cases} \tag{9}$$

Where, $k = \left[\frac{m}{n_h}\right]$, $j_1 = m - (k-1)n_h$, $j_2 = (m - l_1)/n_h$, $h = m - l_1 - (j_2 - 1)n_i$, $k_1 = m - l_2$, $j_3 = m - l_3$.

During the training process, the velocity of each particle is constantly updated and it is determined whether the updated velocity is greater than the maximum velocity v_{max}, if it is greater than the maximum velocity v_{max}, the updated velocity takes the value of the maximum velocity v_{max}, otherwise, it remains unchanged. Similarly, the position of each particle is updated.

Calculate the fitness value for each particle.

$$f_i = \frac{1}{n_t} \sum_{q=1}^{n_t} \left(O_{iq} - T_{iq}\right)^2, \tag{10}$$

Where, n_t is the number of training samples; O_{iq} and $T_{iq}$are the actual and desired outputs of the network under the network weights and thresholds determined by the position of the training sample *q* at particle *i*, respectively.

The global minimum adaptation value $f_g = min\{f_1, f_2, \ldots, f_N\}$ of the particle swarm is calculated; if the current number of iterations reaches the maximum number of iterations or $f_g < \varepsilon$,the iteration is stopped; otherwise, the individual extreme value P_i and the global extreme P_g position of each particle are calculated, and the velocity and position of the particles continue to be updated. Finally, the network weights and thresholds determined by the global pole P_g position are output.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Evaluation indicators

In order to better evaluate the model prediction effect, the mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) were selected as evaluation criteria [8], and each evaluation formula was as follows.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i|, \tag{11}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left|\frac{\hat{y}_i - y_i}{y_i}\right|, \tag{12}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}, \tag{13}$$

Where: *N* is the number of samples, y_i is the actual value, $\hat{y}_i$ is the predicted value, the smaller the value of MAE, MAPE, RMSE, the better the prediction accuracy of the model.

B. Prediction results and comparative analysis

Based on the stock price data of China Pudong Development Bank and Ping An Bank, the accuracy of the prediction model of the extreme learning machine was tested, and the ratio of the training set to the test set in the sample data was 7:3. The number of input neurons is 5, the number of hidden neurons is 5, the number of output neurons is 1, and the sig-moid function is used as the activation function of the hidden layer neurons.

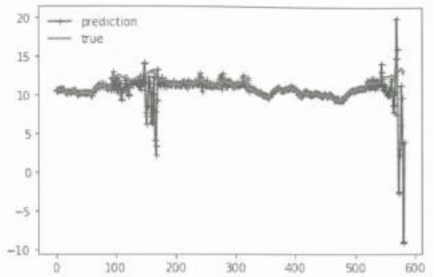

Fig. 1. Comparison of ELM Forecast Results for Pudong Development Bank Stock

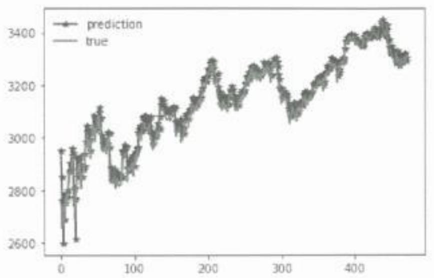

Fig. 2. Comparison of ELM forecast results for Ping An Bank shares

The extreme learning machine stock price prediction model based on particle swarm optimisation was then applied to the sample data to compare the true and predicted values, setting some of the parameters of particle swarm optimisation as follows: the population size was set to 20, the maximum number of iterations k was 300 and the learning factor $c_1 = c_2 = 2$.

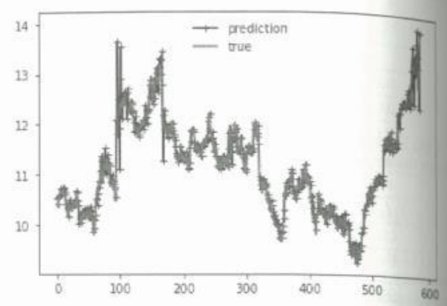

Fig. 3. Comparison of PSO-ELM Forecast Results for Pudong Development Bank Stock

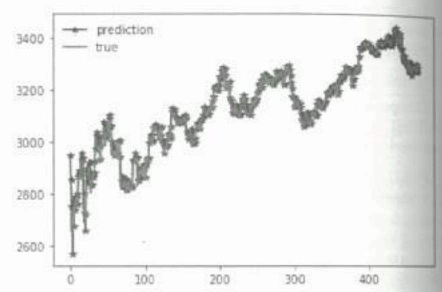

Fig. 4. Comparison of PSO-ELM forecast results for Ping An Bank shares

For a more intuitive comparison of the predictions, the values of the evaluation indicators for each model are shown in Table III.

TABLE III. COMPARISON OF PREDICTED RESULTS

Indicators	Pudong Development Bank Shares			Ping An Bank Shares		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE
PSO-ELM	0.1925	0.0102	0.2471	2.5774	0.0562	4.1132
ELM	0.4602	0.0363	0.3869	13.5441	0.5853	16.6141
PSO-BP	0.2420	0.0197	0.2978	4.8453	0.0892	6.3243
DE-ELM	0.2088	0.0172	0.2673	4.1874	0.0814	5.7617

The two stock prices of Pudong Development Bank and Ping An Bank in the Shanghai and Shenzhen stock markets were used in the PSO-ELM simulation test and compared experimentally with ELM, PSO-BP and DE-ELM. Fig.1 shows the ELM prediction results of Pudong Development Bank stock compared with the true value, Fig.2 shows the ELM prediction results of Ping An Bank stock compared with the true value, Fig.3 shows the PSO-ELM prediction results of Pudong Development Bank stock compared with the true value, Fig.4 shows the PSO-ELM prediction results of Ping An Bank stock compared with the true value. The stock price prediction results of all models are shown in Table III. From Table III, it can be seen that PSO-ELM has the highest stock price prediction accuracy and the lowest prediction error compared to ELM, PSO-BP and DE-ELM models [9]. The experimental results once again validate the effectiveness and superiority of PSO-ELM for stock price modelling and forecasting models.

V. CONCLUSION

Aiming at the dynamic and non-linear change characteristics of stock prices, an extreme learning machine algorithm based on particle swarm optimization was proposed to predict stock prices online with the goal of improving stock price prediction accuracy and achieving fast and accurate online stock price prediction, and multiple stock prices were used as simulation objects. The experimental results show that the extreme learning machine prediction model based on particle swarm optimization can better solve the problems of slow calculation speed and low prediction accuracy of stock price prediction models, can accurately describe the trend of stock price changes, and the prediction results can meet the practical application requirements, which can provide useful reference for stock investors.

REFERENCES

[1] Zhang T., Tang Z. P., Wu J. Chuan. A stock index forecasting method based on the optimized KELM model[J]. Statistics and Decision Making, 2021, 37(13):148-150.

[2] Chen H Y, Liu Y. Stock price forecasting model based on online kernel limit learning machine[J]. Journal of Inner Mongolia Normal University (Chinese version of natural science),2015,44(02):174-178.

[3] Xiong Bingzhong. Stock price prediction based on sparse Bayesian limit learning machine algorithm[J]. Journal of Jiaxing College, 2018,30(05):106-113.

[4] Liang S.C.,Chen X.Y. A multivariate time series forecasting model for extreme learning machines based on singular value decomposition[J]. Journal of Fuzhou University (Natural Science Edition), 2017, 45(01):37-43.

[5] Ye R. Time series prediction algorithm based on migration learning and multi-task learning[D]. Nanjing University of Aeronautics and Astronautics, 2019.

[6] Wang Dong. Improvement and application research of online kernel limit learning machine[D]. Anhui University of Finance and Economics, 2015.

[7] Lu Y Zhang H. Online stock price prediction based on improved limit learning machine algorithm[J]. Computer Engineering and Applications, 2014,50(20):139-143.

[8] Wei Shouqi. Research on stock trend prediction method based on multi-technical indicators and pattern trajectory quantification[D]. Northwestern University,2017.

[9] Li D,Zhang WY. Research on stock price forecasting based on FAM-ELM[J]. Computer Simulation,2014,31(08):209-212+316.

9 781665 492713
ISBN: 978-1-6654-9271-3

华南师范大学图书馆

Library of South China Normal University

广州石牌TEL: 020-85218396

检索证明

编号：202305306087

委托检索单位：私立华联学院

委托人：赵伟明

检索课题：文献被《工程索引》（Engineering Village Compendex）的收录情况

检索数据库：《工程索引》（Engineering Village Compendex）

检出记录数：1

检索结果：

对《工程索引》的检索结果表明，委托人提供的1篇文章（发表于2022年）被《工程索引》收录（详见附录）。

华南师范大学图书馆信息咨询部

2023年5月30日

附录：

第 1 篇

标题：Prediction and Data Analysis of Price Based on PSO with Extreme Learning Machine Algorithm and Particle Swarm Optimization

作者：Chen, Huanyun[1];Zhao, Weiming[1];

来源出版物：Proceedings - 2022 2nd International Signal Processing, Communications and Engineering Management Conference, ISPCEM 2022 页:249-253 出版年:2022

DOI：10.1109/ISPCEM57418.2022.00056

入藏号：20231513871187

文献类型：Conference article (CA)

地址：Private Hualian College, Guangzhou, China[1];

通讯作者：Chen, Huanyun;Zhao, Weiming;

以上结果均由委托人提交确认！