Pearson Correlation Coefficient-Based Performance Enhancement of Broad Learning System for Stock Price Prediction

Guanzhi Li, Aining Zhang, Qizhi Zhang, Di Wu, and Choujun Zhan, *Member, IEEE*

***Abstract*—Accurate prediction of a stock price is a challenging task due to the complexity, chaos, and non-linearity nature of financial systems. In this brief, we proposed a multi-indicator feature selection method for stock price prediction based on Pearson correlation coefficient (PCC) and Broad Learning System (BLS), named the PCC-BLS framework. Firstly, PCC was used to select the input features from 35 features, including original stock price, technical indicators, and financial indicators. Secondly, these screened input features were used for rapid information feature extraction and training a BLS. Four stocks recorded on the Shanghai Stock Exchange or Shenzhen Stock Exchange were adopted to evaluate the performance of the proposed method. In addition, we compared the forecasting results with ten machine learning methods, including Support Vector Regression (SVR), Adaptive Boosting (Adaboost), Bootstrap aggregating (Bagging), Random Forest (RF), Gradient Boosting Decision Tree (GBDT), Multi-layer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and Broad Learning System (BLS). Among all algorithms used in this brief, the proposed model showed the best performance with the highest model fitting ability.**

***Index Terms*—Broad learning system, machine learning, Pearson correlation coefficient, time series forecasting, complex system.**

I. INTRODUCTION

FORECASTING stock prices is of significant importance for analyzing financial systems, as the stock market reflects the economic situation of a country, even the world. Accurate prediction of the stock market allows investors and other stakeholders to understand the movement of the financial market. Effective trading strategies are then adopted to achieve greater profits as well as returns with less risk [1]. In general, the stock market is often analyzed by researchers as a complex system [2]–[4]. However, the non-linearity, high noise and volatility of the stock market make the prediction of financial systems take higher risks than to other complex systems [5]. In addition, external factors such as national macro adjustments, changes from the political situation and investor psychology make the prediction of the financial system extremely difficult [6], [7]. Data-driven methods show great potential for the analysis and control of such complex systems [8], [9].

With the rapid development of data acquisition, storage and machine learning and big data techniques, researchers are beginning to use machine learning methods to try to capture nonlinear patterns of financial systems. Fischer and Krauss predicted the S&P 500 trend from the up-down signals generated by LSTM [10]. Deep Neural Networks (DNNs) also has shown remarkable results in machine learning problems [11]–[13]. A hybrid model based on Naive Bayes classifier and LSTM was used to predict the opening price of listed companies in China [14]. Thakkar *et al.* proposed a method based on Pearson Correlation Coefficient (PCC) to initialize the edge weights of neurons between the input and hidden layers of the vanilla neural network (VNN) [15]. In conclusion, machine learning methods are good candidates for volatility financial time series forecasting [16].

Recent researches indicate that feature selection is essential for improving the accuracy of stock market prediction. Several feature selection approaches are taken to effectively filter redundant or irrelevant features [17], [18]. Ni *et al.* hybridized the fractal feature selection method and Support Vector Machine (SVM) to predict the daily trend of Shanghai Stock Exchange Index [19]. Yang *et al.* filtered the training features by the Maximum Information Coefficient (MIC) and built an ensemble forecasting model for stock price movement prediction based on Support Vector Machine (SVM), RF, and Adaboost [20]. Huang and Tsai established a hybrid SOFM-SVR based on filter feature selection to improve prediction accuracy and predict the fluctuation of the Taiwan Index (FITX) [21].

This brief proposed a novel framework for forecasting the movement of complex financial systems called the Pearson correlation coefficient and Broad Learning System (PCC-BLS) framework. The proposed method was used to predict the close price of a stock listed in China for the day ahead. For verifying the performance of the proposed model, ten machine learning models were adopted for comparison. Experimental results indicate that the PCC-BLS shows more accurate results with

Manuscript received February 7, 2022; accepted March 11, 2022. Date of publication March 17, 2022; date of current version May 3, 2022. This work was supported in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2020A1515010761, and in part by the Key Areas Research and Development Program of Science and Technology Program of Guangzhou under Grant 202103010005. This brief was recommended by Associate Editor H. Li. *(Corresponding author: Choujun Zhan.)*

Guanzhi Li, Aining Zhang, Qizhi Zhang, and Choujun Zhan are with the School of Computing, South China Normal University, Guangzhou 510641, China (e-mail: zchoujun2@gmail.com).

Di Wu is with the Department of ICT and Natural Science, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: di.wu@ntnu.no).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSII.2022.3160266.

Digital Object Identifier 10.1109/TCSII.2022.3160266

TABLE I
INPUT FEATURES FOR MACHINE LEARNING MODELS

Symbols	Original feature
$H_s(t)$	Highest price at time t
$L_s(t)$	Lowest price at time t
$O_s(t)$	Opening price at time t
[1] $CHg_s(t)$	$\frac{C_s(t)-C_s(t-1)}{C_s(t-1)} \times 100$
$\%T_s(t)$	Turnover rate at time t
$V_s(t)$	Volume at time t
$T_s(t)$	Turnover at time t
$C_{se}(t)$	SSEC or SZI closing price at time t
$H_{se}(t)$	SSEC or SZI highest price at time t
$L_{se}(t)$	SSEC or SZI lowest price at time t
$O_{se}(t)$	SSEC or SZI opening price at time t
$PC_{se}(t)$	SSEC or SZI previous closing price at time t
$CHg_{se}(t)$	$C_{se}(t) - C_{se}(t-1)$
$CHg_{se}(t)$	$\frac{C_{se}(t)-C_{se}(t-1)}{C_{se}(t-1)} \times 100$
$V_{se}(t)$	SSEC or SZI Volume of the listed company at time t
$T_{se}(t)$	SSEC or SZI Turnover of the listed company at time t
	Finance indicator
PE_s	Price-to-Earning Ratio Trailing Twelve Months
PB_s	Price-to-Book Ratio most recent quarter
PS_s	Price-to-sales Ratio Trailing Twelve Months
PC_s	Price Cash Flow Ratio Trailing Twelve Months
	Technical indicator
$MA12_s$	$\frac{1}{12}\sum_{i=0}^{n-1} C_s(t-i)$, where $n = 12$
$MA24_s$	$\frac{1}{24}\sum_{i=0}^{n-1} C_s(t-i)$, where $n = 24$
$BIAS12_s$	$\frac{C_s(t)-MA12_s}{MA12_s} \times 100$
$BIAS24_s$	$\frac{C_s(t)-MA24_s}{MA24_s} \times 100$
$EMA12_s$	$\boldsymbol{\alpha} C_s(t) + (1-\boldsymbol{\alpha})EMA12_s(t-1)$, where $\boldsymbol{\alpha} = \frac{2}{13}$
$EMA26_s$	$\boldsymbol{\alpha} C_s(t) + (1-\boldsymbol{\alpha})EMA26_s(t-1)$, where $\boldsymbol{\alpha} = \frac{2}{27}$
$MOM12_s$	$C_s(t) - C_s(t-n+1)$, where $n = 12$
$MOM24_s$	$C_s(t) - C_s(t-n+1)$, where $n = 24$
[2] $WR\%10_s$	$\frac{HH_{t\_t-n+1}-C_s(t)}{HH_{t\_t-n+1}-LL_{t\_t-n+1}} \times 100$, where $n = 10$
$WR\%20_s$	$\frac{HH_{t\_t-n+1}-C_s(t)}{HH_{t\_t-n+1}-LL_{t\_t-n+1}} \times 100$, where $n = 20$
DIF_s	$EMA12_s$-$EMA26_s$
[3] RSI_s	100-$\frac{100}{1+sum\_up(n)/sum\_dw(n)}$, where $n = 14$
[4] CII_s	$\frac{M_s(t)-MA_s(t)}{0.015D_s(t)}$
$\%K_s$	$\frac{C_s(t)-LL_{t\_t-n+1}}{HH_{t\_t-n+1}-LL_{t\_t-n+1}} \times 100$, where $n = 3$
$\%D_s$	$\frac{\sum_{i=0}^{n-1} K_s(t-i)}{n}$, where $n = 3$

[1] $C_s(t)$ represents the closing price at time t.
[2] HH_{t_t-n+1} and LL_{t_t-n+1} represent the highest high and lowest low prices of the listed stock in interval $[t-n+1, t]$.
[3] $sum\_up(n)$ represents the sum of the upward price change over the last n days, $Sum\_dw(n)$ represents the sum of the downward price change over the last n days.
[4] $M_s(t) = (C_s(t)+H_s(t)+L_s(t)/3$, $MA_s(t) = \sum_{i=0}^{n-1} M_s(t-i)/n$, $D_s(t) = \sum_{i=0}^{n-1} |M_s(t-i) - MA_s(t)|/n$

the highest model fitting ability in terms of stock prediction. The main contributions of this brief are as follows.

- A large-scale dataset, containing multiple factors that affect the closing price, has been collected. The collection of features involves technical and financial indicators in addition to the original features (shown in Table I).
- PCC is adopted for feature selection analysis to improve BLS. The PCC-BLS model is proposed for forecasting the closing price of listed companies in China.
- Ten machine learning methods, including ensemble learning, RF, standard BLS and etc., have been adopted to predict the closing prices of the adopted stock price and compare the performance with the proposed model.

The rest of this brief is organized as follows: In Section II, the proposed Pearson correlation coefficient (PCC) and Broad

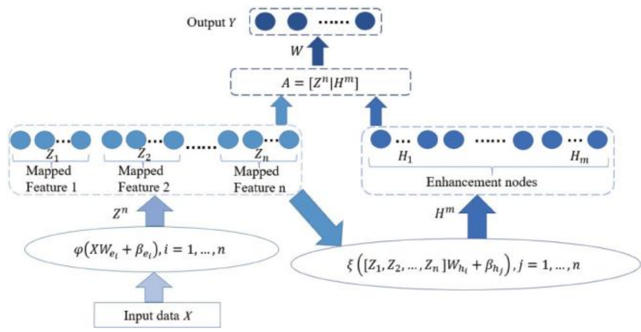


Fig. 1. The illustrative structure of a typical Broad Learning System.

Learning System (BLS) framework are introduced in detail. Then, the experimental design and results are given in Section III. Conclusion and future work are presented in Section IV.

II. METHODOLOGY

A. Broad Learning System

Broad Learning System (BLS), based on the random vector functional link neural network (RVFLNN), is a flat network proposed by Chen and Liu [22]. The framework of a typical BLS is shown in Fig. 1. The input features are first transformed into mapped features by mapping functions and stored in the feature nodes, which would be extended to enhancement nodes through nonlinear activation functions and random weights. In BLS, all the outputs of the enhancement and feature nodes are connected to the output layer, and the connection weights are derived by ridge regression of the pseudoinverse.

Considering the time series forecasting problem, the training data set is represented by $\{(\boldsymbol{X}, \boldsymbol{Y}) \mid \boldsymbol{X} \in \mathbb{R}^{N\times D}, \boldsymbol{Y} \in \mathbb{R}^{N\times 1}\}$. The mapping function first transforms the training samples into mapping feature spaces in BLS. For n feature mappings, each generates k nodes and this procedure can be expressed as the following equation:

$$Z_i = \varphi\left(XW_{e_i} + \beta_{e_i}\right), \qquad i = 1, \ldots, n, \tag{1}$$

where weights W_{e_i} and bias term β_{e_i} are randomly generated matrices with the proper dimensions. $\varphi(\cdot)$ is usually the transformation of the activation function. The feature space of training samples are defined as $Z^n = [Z_1, Z_2, \ldots, Z_n]$ and the ith group of enhancement nodes is denoted as follow:

$$H_j = \xi(Z^n W_{h_j} + \beta_{h_j}), \qquad j = 1, 2, \ldots, m, \tag{2}$$

where weights W_{h_j} and bias term β_{h_j} are also generated randomly, and $\xi(\cdot)$ represents the activation function. And we denote the outputs of the enhancement layer by $\boldsymbol{H}^m$, $[\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_m]$. Therefore, the output Y of BLS can be represented as:

$$\begin{aligned} Y &= [Z_1, Z_2, \ldots, Z_n \mid H_1, H_2, \ldots, H_m]W^m \\ &= \left[Z^n \mid H^m\right] W^m \\ &= AW^m, \end{aligned} \tag{3}$$

where $A = [Z^n \mid H^m]$. W^m is the output weight connecting the feature and enhancement nodes to the output layer. W^m can be rapidly approximated by the ridge regression.

$$W^m = \left(\lambda I + A^T A\right)^{-1} A^T Y, \tag{4}$$

where λ denotes a regularisation parameter to balance the effect of the error term.

BLS allows output weights to be obtained quickly by pseudoinversion, which not only ensures efficient modeling but also avoids structural redundancy. Therefore, the BLS provides an alternative approach to a deep learning framework, which is typically subject to time-consuming training of numerous parameters in a layer.

B. Pearson Correlation Coefficient

Pearson Correlation Coefficient, introduced by Pearson in 1895, is used to measure the strength and direction of the relationship between two variables [23]. The formula for deriving the PCC between variables x_i (stock features) and y (stock price)is shown below:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{5}$$

The greater the absolute value of the $r_{x,y}$ is, the stronger the correlation. The value of PCC ranges from + 1 to −1. Here, $r_{x,y} = 1$ represents a completely positive linear correlation, while $r_{x,y} = -1$ stands for a completely negative linear corre- lation. On the other hand, $r_{x,y} = 0$ means no linear correlation. In this brief, we calculated the PCC between the features and the closing price of a stock. Features with an absolute value of PCC greater than a threshold of 0.5 were used as input data.

III. Experiment Design and Results

A. Data Description

In this brief, we selected four stocks listed on the Shenzhen Stock Exchange or Shanghai Stock Exchange as illustrative examples. Research data was collected from September 2010 to September 2021, including approximately 2500 trading days. The entire data set was divided into two parts. 80% of the dataset was used as the training dataset, while the remaining 20% was acted as the test dataset.

Based on the knowledge of the stock market, we obtained data sets from multiple sources to improve the forecasting ability [24]. Historical data and financial indicators of each stock were fetched from BaoStock's API interface. As the stock market index reflects the closing price of the company to some extent, we collected the Shenzhen Securities Component Index (SZI) and Shanghai Securities Composite Index (SSEC) from the NetEase website. Finally, we derived technical indicators and integrated historical data into the machine learning models. The input variables used in this brief, including their formula or description, are summarized in Table I. As the range of different input features varies inconsistently, these datasets were scaled into the range of [0, 1] through the Min-Max normalization technique. x_{norm} is given by the following equation:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{6}$$

TABLE II
SZ.000776: The Input Features of the PCC-BLS Model

Features	Symbols
Original feature	$H_s(t),L_s(t),O_s(t),V_s(t),T_s(t),C_{sc}(t)$ $H_{sc}(t),L_{sc}(t),O_{sc}(t),PC_{sc}(t),V_{sc}(t),T_{sc}(t)$
Finance indicator	None
Technical indicator	$MA12_s,MA24_s,BIAS12_s,BIAS24_s$ $EMA12_s,EMA26_s$

TABLE III
SH.600019: The Input Features of the PCC-BLS Model

Features	Symbols
Original feature	$H_s(t),L_s(t),O_s(t),V_s(t),T_s(t),C_{sc}(t)$ $H_{sh}(t),L_{sc}(t),O_{sc}(t),PC_{sc}(t),T_{sc}(t)$
Finance indicator	None
Technical indicator	$MA12_s,MA24_s,BIAS12_s,BIAS24_s$ $EMA12_s,EMA26_s$

where x_{min} and x_{max} are the minimum and maximum value in the training dataset, respectively.

B. Experimental Setup

In order to evaluate the forecasting performance of the proposed PCC-BLS, ten machine learning models have been adopted for comparison. We employed SVR, Adaboost, Bagging, RF, GBDT, MLP, CNN, LSTM, GRU, and BLS approaches in the comparisons. Datasets of listed companies with experimental stocks, containing GF Securities (sz.000776), BAOSTEEL (sh.600019), China Unicom (sh.600050), and Haitong Securities (sh.600837), span from September 2010 to September 2021.

To evaluate the forecasting performance of each mode, we used five evaluation criteria in the experiments: mean absolute error (MAE), mean square error (MSE), coefficient of determination (R^2), mean absolute percentage error ($MAPE$), and adjusted coefficient of determination (R^2_{adj}). Here, we assume the dataset has n samples and these criteria can be defined as follows:

$$\begin{aligned} MSE &= \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2, \\ MAE &= \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|, \\ MAPE &= \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|, \\ R^2 &= 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \\ R^2_{adj} &= 1 - \frac{(1 - R^2)(n - 1)}{n - k - 1}, \end{aligned} \tag{7}$$

where y_i and $\hat{y}_i$ represent the actual and predicted value, respectively. $\bar{y}$ is the average of the close price of the n samples, and k denotes the number of features.

C. Correlation Analysis

Tables II, III, IV, and V show the input features of the PCC-BLS for four experimental stocks. Here, we chose features with an absolute value of PCC greater than 0.5 as input

TABLE IV
SH.600050: THE INPUT FEATURES OF THE PCC-BLS MODEL

Features	Symbols
Original feature	$H_s(t)$,$L_s(t)$,$O_s(t)$,$T_s(t)$,$C_{sc}(t)$,$H_{sc}(t)$,$L_{sc}(t)$ $O_{sc}(t)$,$PC_{sc}(t)$
Finance indicator	PB_s
Technical indicator	$MA12_s$,$MA24_s$,$BIAS12_s$,$BIAS24_s$ $EMA12_s$,$EMA26_s$

TABLE V
SH.600837: THE INPUT FEATURES OF THE PCC-BLS MODEL

Features	Symbols
Original feature	$H_s(t)$,$L_s(t)$,$O_s(t)$,$C_{sc}(t)$,$H_{sc}(t)$ $L_{sc}(t)$,$O_{sc}(t)$,$PC_{sc}(t)$,$V_{sc}(t)$,$T_{sc}(t)$
Finance indicator	None
Technical indicator	$MA12_s$,$MA24_s$,$BIAS12_s$,$BIAS24_s$ $EMA12_s$,$EMA26_s$

TABLE VI
SZ.000776: THE EVALUATION VALUE OF DIFFERENT MODELS

model	MSE	MAE	MAPE	R^2	R^2_{adj}
Adaboost	0.220	0.308	1.997	0.934	0.929
Bagging	0.191	0.288	1.851	0.943	0.938
CNN	0.411	0.449	2.889	0.877	0.867
GBDT	0.228	0.291	1.848	0.932	0.927
GRU	0.344	0.389	2.499	0.897	0.889
LSTM	0.350	0.389	2.506	0.895	0.887
MLP	1.080	0.673	4.201	0.677	0.652
RF	0.195	0.288	1.855	0.942	0.937
SVR	0.213	0.310	2.018	0.936	0.931
BLS	0.185	**0.272**	**1.748**	0.945	0.943
PCC-BLS	**0.166**	0.285	1.864	**0.950**	**0.947**

TABLE VII
SH.600837: THE EVALUATION VALUE OF DIFFERENT MODELS

model	MSE	MAE	MAPE	R^2	R^2_{adj}
Adaboost	0.099	0.227	1.754	0.939	0.934
Bagging	0.082	0.202	1.549	0.949	0.945
CNN	0.227	0.350	2.746	0.857	0.846
GBDT	0.079	0.195	1.499	0.951	0.947
GRU	0.203	0.320	2.470	0.872	0.862
LSTM	0.154	0.275	2.114	0.902	0.895
MLP	0.788	0.773	6.222	0.513	0.476
RF	0.090	0.215	1.650	0.944	0.940
SVR	0.286	0.421	3.335	0.823	0.810
BLS	0.106	0.238	1.848	0.935	0.930
PCC-BLS	**0.078**	**0.194**	**1.484**	**0.952**	**0.950**

TABLE VIII
SH.600019: THE EVALUATION VALUE OF DIFFERENT MODELS

model	MSE	MAE	MAPE	R^2	R^2_{adj}
Adaboost	0.141	0.199	2.871	0.936	0.931
Bagging	0.162	0.177	2.448	0.927	0.921
CNN	0.079	0.193	3.054	0.965	0.963
GBDT	0.114	0.192	2.854	0.948	0.944
GRU	0.068	0.165	2.508	0.970	0.968
LSTM	0.071	0.165	2.493	0.969	0.967
MLP	0.397	0.406	6.264	0.820	0.806
RF	0.188	0.195	2.686	0.915	0.908
SVR	0.196	0.292	4.153	0.911	0.904
BLS	0.031	0.116	1.796	0.986	0.985
PCC-BLS	**0.028**	**0.109**	**1.724**	**0.987**	**0.986**

TABLE IX
SH.600050: THE EVALUATION VALUE OF DIFFERENT MODELS

model	MSE	MAE	MAPE	R^2	R^2_{adj}
Adaboost	0.011	0.078	1.587	0.968	0.966
Bagging	0.007	0.060	1.192	0.979	0.978
CNN	0.020	0.106	2.184	0.937	0.932
GBDT	0.008	0.061	1.214	0.979	0.977
GRU	0.013	0.084	1.727	0.959	0.956
LSTM	0.012	0.080	1.627	0.962	0.959
MLP	0.122	0.287	5.709	0.659	0.633
RF	0.009	0.065	1.308	0.976	0.974
SVR	0.058	0.198	4.165	0.837	0.825
BLS	0.007	0.062	1.248	0.980	0.978
PCC-BLS	**0.006**	**0.054**	**1.092**	**0.982**	**0.981**

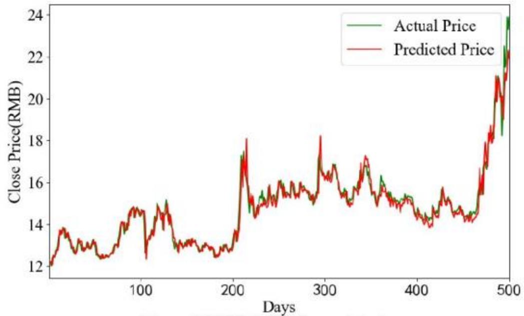
features for the PCC-BLS. It can be seen that the stock price data and the derived technical indicators of the experimental stocks are correlated with their closing prices to some extent. In addition, there was also a strong correlation between the close price of the experimental stock and the SSEC or SZI. Actually, we also adopt the threshold as 0.4∼0.6. Of the above range of thresholds, setting the threshold at 0.5 gives the best predictions.

D. Experimental Results

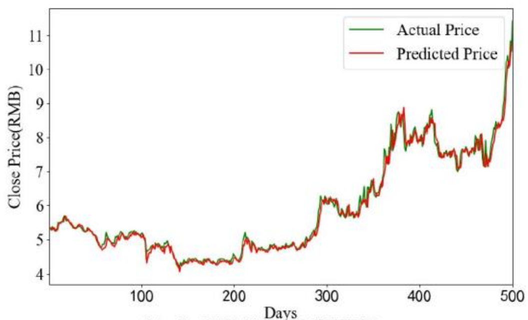
Tables VI, VII, VIII, and IX show the experimental results of the performance of the ten other models with the PCC-BLS for four experimental stocks. It can be seen that the PCC-BLS outperformed the other models on the four experimental stock prediction results. In the case of sz.000776, compared to the second-ranked model BLS, PCC-BLS has 10% less MSE. The R^2 is improved from 0.945 to 0.950 through the Pearson Correlation Coefficient screening feature. In the case of sh.600019, the PCC-BLS has 9% less MSE compared to the second-ranked model BLS. The R^2 of the proposed model is 0.987, which is an outstanding fit. In the case of sh.600050, compared to the second-ranked model Bagging, PCC-BLS achieves 14% decreases of MSE, and 0.3% increase of R^2, respectively. In the case of sh.600837, compared to the second-ranked model Bagging, PCC-BLS achieves 4.8% decreases of MSE, and 0.8% increase of R^2, respectively. The results show that the PCC-BLS network achieves higher performance in the stock price forecasting problem. The forecasting results of the four stocks by the PCC-BLS model are shown in Fig. 2. In addition, PCC-BLS also can save lots of training time. The running time of training the PCC-BLS model for the four stocks was about 1.14s, 1.34s, 1.95s, and 2.07s, respectively, which is faster than most of the other models. For instance, the GRU model consumes 82.35s, 104.56s, 110.78s, and 102.41s, respectively, for training.

IV. CONCLUSION

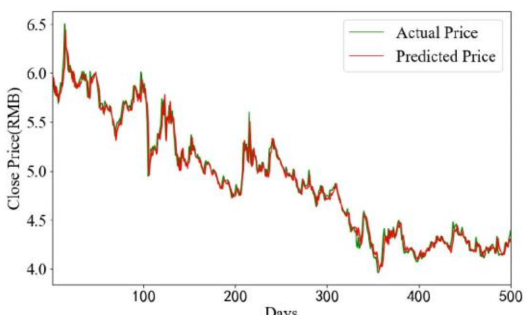
The main purpose of this brief is to help investors make reasonable trading decisions by accurately predicting stock price movements using machine learning methods. In this brief, we proposed a new framework composed of PCC and BLS, and applied this framework for a short-term prediction of stock price recorded on the Shenzhen Stock Exchange or Shanghai Stock Exchange. As feature selection aims at selecting more representative features for enhancing forecasting performance, we adopted suitable input variables by PCC from 35 variables. The combinations of input variables were then input to the BLS for training. Finally, we compared the proposed model with ten machine learning methods without feature selection on five evaluation metrics. The experimental results showed



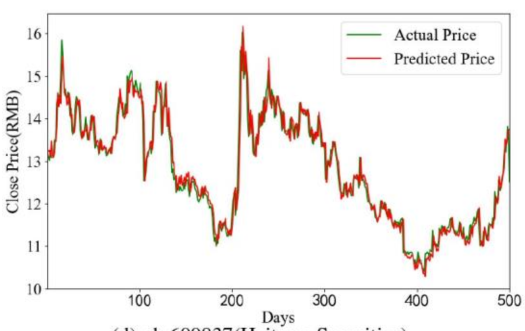
(a) sz.000776(GF Securities)



(b) sh.600019(BAOSTEEL)



(c) sh.600050(China Unicom)



(d) sh.600837(Haitong Securities)

Fig. 2. Actual and predicted price of the four experimental stocks using PCC-BLS.

that PCC-BLS provided more accurate forecasting of future values than other prediction models mentioned above.

However, in this brief, we are short on the practical application of the proposed model. In the future, we will adopt an actual stock investment strategy and evaluate the practical application of the model through returns. Future research will also focus on the comparison of different feature selection methods.

REFERENCES

[1] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Comput. Appl.*, vol. 33, no. 10, pp. 4741–4753, 2021.

[2] C. K. Tse, J. Liu, and F. C. M. Lau, "A network perspective of the stock market," *J. Empirical Financ.*, vol. 17, no. 4, pp. 659–667, 2010.

[3] X. F. Liu and C. K. Tse, "A complex network perspective of world stock markets: Synchronization and volatility," *Int. J. Bifurcation Chaos*, vol. 22, no. 6, 2012, Art. no. 1250142.

[4] G. Chen, Y. Lou, and L. Wang, "A comparative study on controllability robustness of complex networks," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 5, pp. 828–832, May 2019.

[5] M. Small and C. K. Tse, "Determinism in financial time series," *Stud. Nonlinear Dyn. Econom.*, vol. 7, no. 3, pp. 1–29, 2003.

[6] J. Dong, W. Dai, Y. Liu, L. Yu, and J. Wang, "Forecasting chinese stock market prices using baidu search index with a learning-based data collection method," *Int. J. Inf. Technol. Decis. Making*, vol. 18, no. 5, pp. 1605–1629, 2019.

[7] D. Zhang and S. Lou, "The application research of neural network and BP algorithm in stock price pattern classification and prediction," *Future Gener. Comput. Syst.*, vol. 115, pp. 872–879, Feb. 2021.

[8] Q. Li *et al.*, "Integrating reinforcement learning and optimal power dispatch to enhance power grid resilience," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 1402–1406, Mar. 2021.

[9] X. Zhang *et al.*, "Braess paradox and double-loop optimization method to enhance power grid resilience," *Rel. Eng. Syst. Safety*, vol. 215, Nov. 2021, Art. no. 107913.

[10] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, 2018.

[11] B. N. G. Koneru and V. Vasudevan, "Sparse artificial neural networks using a novel smoothed LASSO penalization," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 5, pp. 848–852, May 2019.

[12] L. Bai, Y. Zhao, and X. Huang, "A CNN accelerator on FPGA using depthwise separable convolution," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 10, pp. 1415–1419, Oct. 2018.

[13] A. R. Mohamed, L. Qi, Y. Li, and G. Wang, "A generic nano-Watt power fully tunable 1-D Gaussian kernel circuit for artificial neural network," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 9, pp. 1529–1533, Sep. 2020.

[14] Q. Zhuge, L. Xu, and G. Zhang, "LSTM neural network with emotional analysis for prediction of stock price," *Eng. Lett.*, vol. 25, no. 2, pp. 167–175, 2017.

[15] A. Thakkar, D. Patel, and P. Shah, "Pearson correlation coefficient-based performance enhancement of vanilla neural network for stock trend prediction," *Neural Comput. Appl.*, vol. 33, pp. 1–16, Jul. 2021.

[16] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Philos. Trans. Roy. Soc. A*, vol. 379, no. 2194, 2021, Art. no. 20200209.

[17] C.-F. Tsai and Y.-C. Hsiao, "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches," *Decis. Support Syst.*, vol. 50, no. 1, pp. 258–269, 2010.

[18] D. Liu, C. K. Tse, and X. Zhang, "Robustness assessment and enhancement of power grids from a complex network's perspective using decision trees," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 5, pp. 833–837, May 2019.

[19] L.-P. Ni, Z.-W. Ni, and Y.-Z. Gao, "Stock trend prediction based on fractal feature selection and support vector machine," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5569–5576, 2011.

[20] J. Yang, R. Rao, P. Hong, and P. Ding, "Ensemble model for stock price movement trend prediction on different investing periods," in *Proc. 12th Int. Conf. Comput. Intell. Security (CIS)*, 2016, pp. 358–361.

[21] C.-L. Huang and C.-Y. Tsai, "A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1529–1539, 2009.

[22] C. L. P. Chen and Z. Liu, "Broad Learning System: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.

[23] K. Pearson, "Contributions to the mathematical theory of evolution," *Philos. Trans. Roy. Soc. London A*, vol. 185, pp. 71–110, Dec. 1894.

[24] A. Thakkar and K. Chaudhari, "Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions," *Inf. Fusion*, vol. 65, pp. 95–107, Jan. 2021.