Advances in Computer, Signals and Systems (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2023.070112
ISSN 2371-8838 Vol. 7 Num. 1

A Time Series Data Prediction Model Based on Adaptive Weighted LSTM

Wanlu Shen*, Ruotong Wu

School of Science, Northeast Electric Power University, Jilin, 132011, China
**Corresponding author: wurtong03@163.com*

Keywords: Stock Price Prediction, LSTM, FPCA, PCA, Model Averaging Method, Distance correlation coefficient

Abstract: Financial time series prediction has always been a hot topic in the field of statistics learning. Aiming at the step selection problem of LSTM time series prediction model, this paper proposes an adaptive weighted LSTM model based on model average method. The model average is mainly reflected in two aspects: On the one hand, the proposed method takes intraday price information into account. Firstly, functional and nonlinear information of intraday price series are extracted through functional principal component analysis and kernel principal component analysis, and then Bagging is used to fit the residual sequence generated by the original LSTM model. On the other hand, the proposed method integrates the information of the model under different time Windows by using the weight based on distance correlation coefficient, and adaptively solves the step size selection problem, so as to improve the effectiveness of the overall model. The actual data analysis results show that the proposed method can effectively improve the prediction accuracy of the original LSTM model and has a certain robustness. Due to the flexibility of the proposed method, it can be used in time series prediction such as energy consumption prediction, environment detection and road traffic flow monitoring.

1. Introduction

Time series prediction is a method to analyze the data of the past period of time based on the development track of things in the past, so as to seek a continuous rule based on time information, and finally to further predict the development trend of things in the future period of time based on statistical methods. Because it can predict the future development trend to some extent, time series prediction can be used in many fields of research, among which stock price prediction is an important application direction. Specifically, for stock investors, it is undoubtedly the best choice to invest in a stock with low risk and high return, so it is particularly important to choose an appropriate and relatively accurate stock price prediction method. This paper takes the price of financial stocks as an example and selects three stocks for analysis and prediction. The price fluctuation chart is shown in Figure 1.

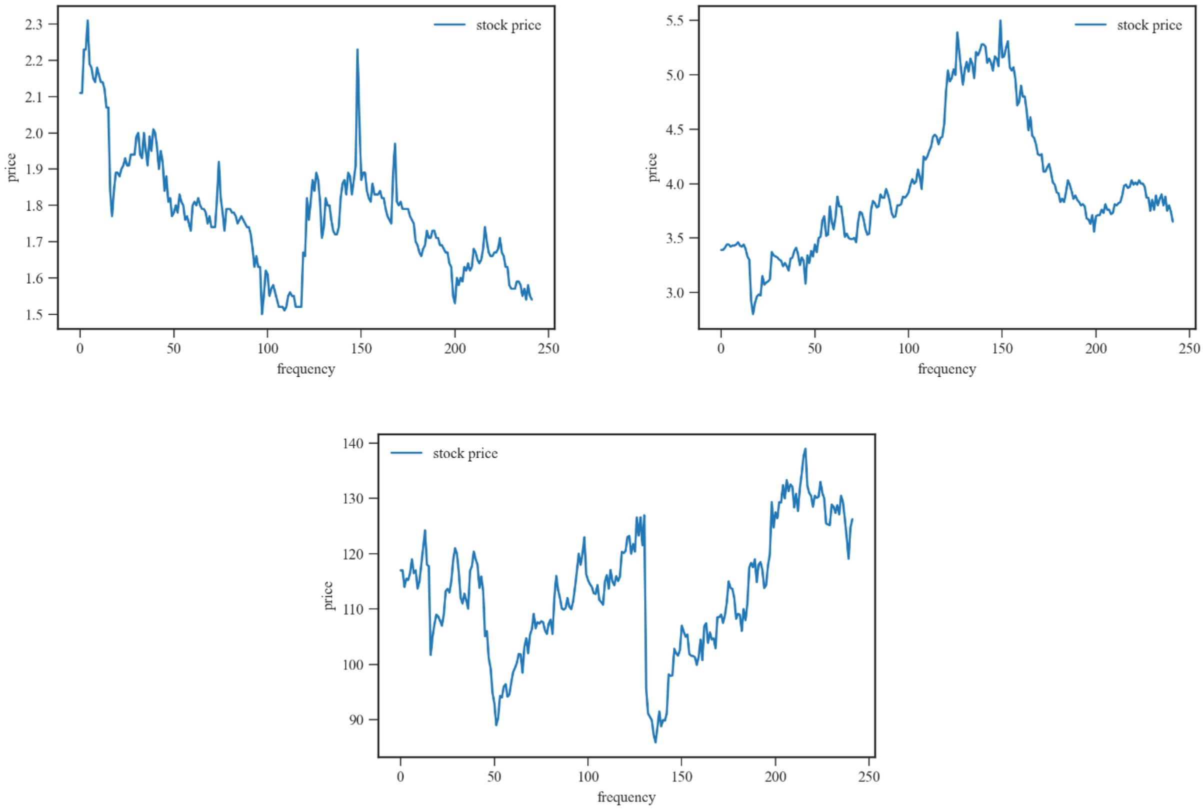

Figure 1: Three stock price fluctuation chart

In the prediction research of time series, the classical prediction models mainly include ARIMA model [1] and GARCH model [2]. Li and Engle et al. proved through experiments that the two models have certain accuracy in stock price prediction. Yang Qi and Cao Xianbing et al. combined the two and got the Armia-Garch fitting model to predict stock prices, which showed the better effectiveness and accuracy of the model [3]. However, the traditional time series prediction model has strict requirements on the data stationarity. If the data stationarity does not meet the requirements, the difference method should be used to process the data, which will cause the loss of original information. With the continuous development of deep learning, models such as LSTM neural network and BP neural network emerge in the field of stock price prediction. By using the selective memory function of LSTM neural network model, Dou Wei dug deeply into the time series of stock prices and captured its inherent law, thus realizing the prediction of stock price trend [4]. Xu Xingjun et al. demonstrated that the stock price prediction model based on BP neural network can make a correct judgment on the stock price trend to a certain extent [5]. Xu Youen et al. compared the advantages and disadvantages of traditional statistical forecasting model, machine learning, neural network and other single forecasting models with combination forecasting models, and forecast the effectiveness and stability of various combination model stock forecasting methods [6]. Chen Yu combined machine learning technology and optimization algorithm to build a combination strategy, and achieved excellent performance in the back test period [7]. Fang Yi et al. found that deep neural network had the best prediction effect by comparing the results of various models [8]. However, in the process of constructing LSTM neural networks, step size selection is an important problem to be solved. In addition, raw time series information may not be sufficient to predict future stock prices.

Based on this, this paper proposes an adaptive weighted LSTM model based on model average method, which organically combines functional principal component analysis [9], kernel principal component analysis [10] and model average method [11] with long and short term memory network, so as to improve the prediction accuracy of the original LSTM model while ensuring its robustness. The innovation points of this paper are mainly reflected in two aspects: firstly, the functional information and nonlinear information of intra-day price are taken into account in the prediction model, and Bagging algorithm is used to fit the residual sequence generated by the original LSTM model to solve the problem of insufficient information of the original time series. Secondly, the model average method is used to replace the original model selection method, so as to solve the step selection problem of LSTM neural network adaptively. The specific method is to use the normalized distance correlation coefficient to weight the predicted value of LSTM under each step length. The actual data analysis shows that the proposed method has obvious improvement on the original LSTM model. Finally, the proposed method can also be applied to energy consumption prediction and environmental monitoring.

2. Theory and method

2.1 LSTM neural network

Long Short-Term Memory neural network (LSTM) is a kind of time cyclic neural network, which is a variant of cyclic neural network (RNN). It solves the problem of information redundancy by adding "gate" structure at appropriate location. When it flows through neurons, information is allowed to be selectively retained or abandoned, thus enhancing the weight of original information and weakening the weight of irrelevant information, and solving problems such as gradient disappearance, gradient explosion and inability to deal with long-term dependence in traditional recursive neural networks [12]. The main formula of LSTM neural network is as follows:

(1) Determine the information to be discarded. This operation is completed through the forgetting gate. The calculation formula is as follows:

$$f_t = \sigma(W_f \times (h_{t-1}, X_t) + b_f) \tag{1}$$

(2) Take the new information stored and the information retained at the last moment together as the updated state of the input door layer. The calculation formula is as follows:

$$i_t = \sigma(W_i \times (h_{t-1}, X_t) + b_i) \tag{2}$$

$$C_t = i_t \times a_t + f_t \times C_{t-1} \tag{3}$$

(3) Use the output gate layer to determine the output part. The calculation formula is as follows:

$$a_t = \tanh(W_c \times (h_{t-1}, X_t) + b_c) \tag{4}$$

$$h_t = \sigma(W_0 \times (h_{t-1}, X_t) + b_0) \times \tanh(C_t) \tag{5}$$

2.2 Model average algorithm

2.2.1 Bagging Algorithm

Bagging algorithm is one of the most widely used model averaging algorithms. It is mainly a technique to reduce generalization errors by combining several models, which can improve the generalization ability of the whole prediction model by reducing the instability of the weak learning algorithm [13]. Specifically, the Bagging integration process is given a weak learner and a training set. Firstly, the training set is extracted from the original sample set, and n training samples are extracted from the original sample set in each round. A total of k rounds are extracted to obtain k training sets.

Use one training set at a time to get a model, k training sets to get k models. After k training, a sequence of prediction models can be obtained: $f_1, f_2, \cdots, f_k$. Finally, the result of strong prediction model is obtained by using arithmetic average method.

2.2.2 Heuristic model averaging method based on distance correlation coefficient

Distance correlation coefficient can be used to measure the correlation between multiple variables, especially when the relationship between variables is nonlinear. It mainly studies the independence between two variables. If the distance correlation coefficient is 0, the two variables are independent. We can use Distance Correlatin to study the independence between two variables and, denoted as, and its calculation formula is as follows:

$$dcorr(\mu,v)=\frac{dcov(\mu,v)}{\sqrt{dcov(\mu,\mu)dcov(v,v)}} \tag{6}$$

$$dcov^2(\mu,v)=\widehat{S_1}+\widehat{S_2}-2S_3 \tag{7}$$

$$\widehat{S_1}=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}||\mu_i-\mu_j|\,|d\mu|\,|v_i-v_j||dv \tag{8}$$

$$\widehat{S_2}=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}||\mu_i-\mu_j||d\mu\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}||v_i-v_j||dv \tag{9}$$

$$\widehat{S_3}=\frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}||\mu_i-\mu_l|\,|d\mu|\,|v_j-v_l||dv \tag{10}$$

Through the above formula, we can calculate the distance correlation coefficient between the predicted value and the real value under each step size on the training set, and use the result as the weight of different prediction models, so as to predict the predicted value of the test set.

2.3 Time series data prediction model based on adaptive weighted LSTM

This experiment mainly adopts the adaptive weighted LSTM time series data prediction model, calculates the residual difference between the predicted value and the real value, obtains the corresponding prediction result by predicting the residual series, and calculates the final predicted value by calculating the weight. The specific steps of this experiment are as follows:

(1) Under a fixed step size, use LSTM to predict the time series to get $\hat{y}$ and calculate the residual $e=y-\hat{y}$.

(2) By selecting different step size T, calculate the residual sequence set $\{e_1,e_2,e_3,\ldots,e_n\}$, where n is the number of elements in the step size set.

(3) kPCA and FPCA are used to reduce the dimension of intra-day price data X, and the reduced dimension data is used to predict each residual sequence in the above residual sequence set. The prediction model is Bagging algorithm, and the predicted value set $\{\hat{e}_1,\hat{e}_2,\hat{e}_3,\ldots,\hat{e}_n\}$ of residual sequence is obtained.

(4) Calculate the sequence set $\tilde{y}=\hat{y}+\hat{e}$ of predicted values, and obtain the predicted result $\tilde{y}_1$, $\tilde{y}_2$, ..., $\tilde{y}_n$ with asynchronous length.

(5) The normalized distance correlation coefficient between $\tilde{y}$ under asynchronous length and y predicted by the training set is used as the weight of the predicted value under the current step size.

(6) Finally, based on the model trained on the training set, the time series predicted value $\widetilde{\widetilde{y_{T+1}}}$ on the test set was obtained.

3. Data analysis

3.1 Data source and experimental environment

As the stock price data belongs to a random system, which reflects the market economic data and trading dynamics, various factors and their effects on the stock price are complex internal laws. We want to investigate whether our proposed method has advantages for stock price prediction in different situations where stochastic systems are affected by different levels of complexity. Therefore, we choose stocks randomly to avoid other factors. This study used the stock data are from the Shenzhen stock exchange (http://www.szse.cn/index/index.html), three stocks are new energy (600777), Oriental Group (600811A) shares and Opai Home (603833) daily opening price series and intraday price series within 242 days.

This experiment is mainly run in R 4.2.2 software and Python 3.9 software. In order to test the robustness of the proposed method, periods 135, 140, 145, 150, 155, 160, 165, 170, 175 and 180 were selected as training sets to test the robustness of the model. Finally, this experiment mainly compares the original LSTM and the improved LSTM through three error estimation methods, namely mean square error, absolute error and posterior error, and obtains the final result. Mean square error (MSE) is usually used as a loss function for regression problems, and the formula is:

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2 \tag{11}$$

Relative error is the ratio between the absolute error of measurement and the measured true value, and the formula is:

$$MAE=\left(\Sigma_{t=1}^{T}\left|\widetilde{Y}-Y_t\right|\right)/T \tag{12}$$

A posteriori error refers to the error obtained after the exact approximate solution is obtained. The formula is:

$$MAPE=\left(\Sigma_{t=1}^{T}\frac{\left|\widetilde{Y}-Y_t\right|}{Y_t}\right)/T\times 100\% \tag{13}$$

3.2 Empirical results

Table 1: Error result of stock 600777

Number of training sets	FDMA-MSE	FDMA-MRE	FDMA-BE	LSTM-MSE	LSTM-MRE	LSTM-BE
145	0.001493	0.013883	0.295824	2.068407	0.827941	0.473127
150	0.001393	0.013742	0.285318	2.053823	0.825125	0.470292
155	0.001168	0.013135	0.261418	2.066336	0.827335	0.454603
160	0.001039	0.012751	0.245106	2.044190	0.823353	0.460311
165	0.001013	0.012507	0.243458	2.047939	0.823813	0.44393
170	0.001034	0.012605	0.245453	2.055161	0.825384	0.424164
175	0.001000	0.012420	0.240342	2.060150	0.826460	0.432099
180	0.000950	0.012131	0.234660	2.051318	0.824373	0.444838
185	0.000957	0.012164	0.235280	2.057807	0.825645	0.442771
190	0.000912	0.012097	0.229687	2.066120	0.827329	0.442263

Through the above three evaluation methods, our model was named FDMA and the original model was LSTM. Set the step size set as 5, 10, 15, 20, 25, 30. As for the result of LSTM, the optimal result under all step sizes in the set is selected. The final experimental results are shown in Table 1-3 and Figure 2-4:

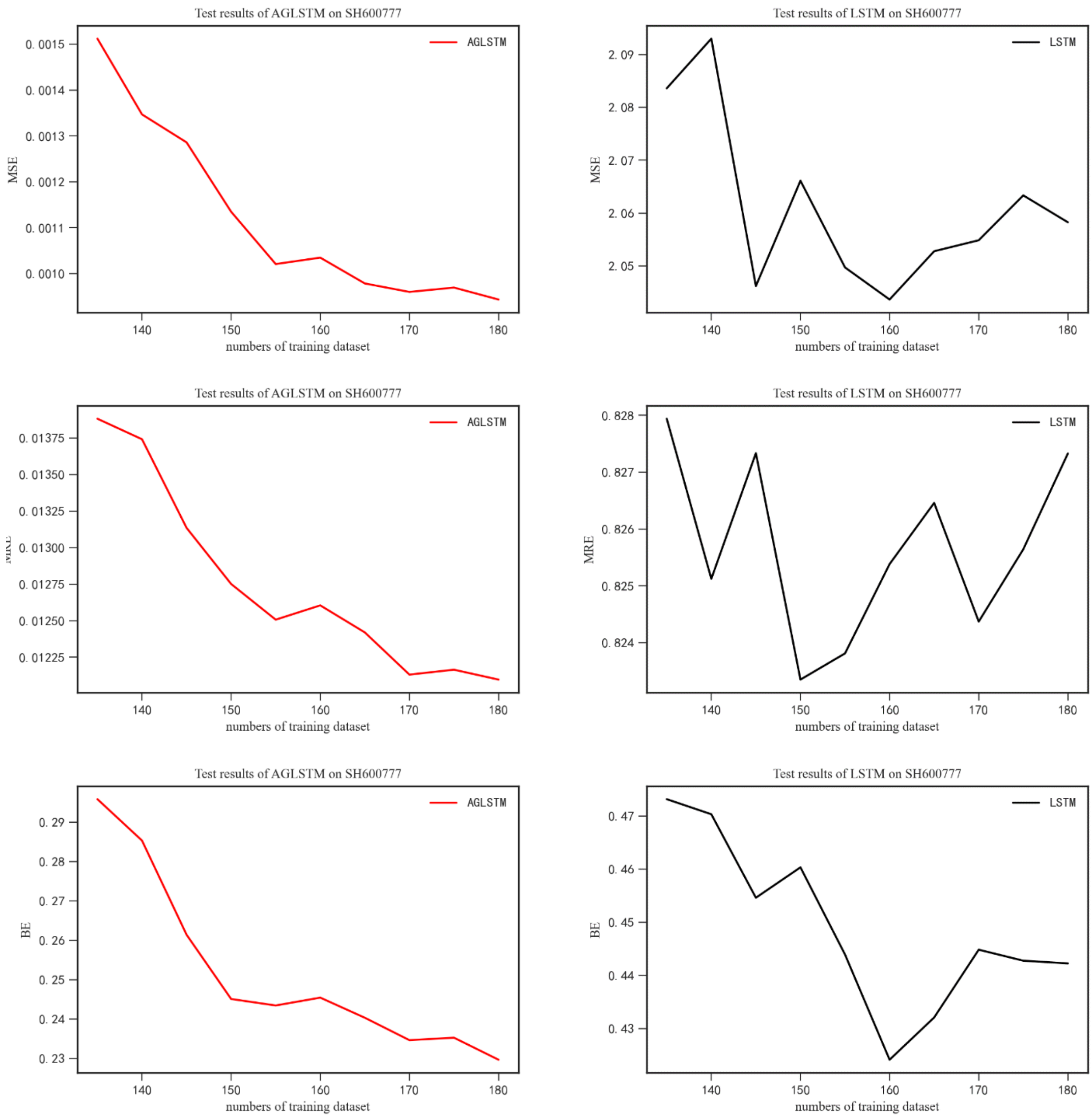

Figure 2: Stock 600777 error fluctuation in different training sets

Table 2: Error result of stock 600811A

Number of training sets	FDMA-MSE	FDMA-MRE	FDMA-BE	LSTM-MSE	LSTM-MRE	LSTM-BE
145	0.005868	0.01195	0.119895	12.58745	0.880849	0.613633
150	0.004880	0.011268	0.110226	12.54461	0.878961	0.622168
155	0.004005	0.010673	0.100308	12.66929	0.882536	0.642868
160	0.003282	0.010029	0.091720	12.63498	0.881082	0.649761
165	0.002658	0.009314	0.083276	12.64151	0.880983	0.659056
170	0.002181	0.008681	0.076254	12.79063	0.885999	0.663571
175	0.001871	0.008160	0.071324	12.76190	0.884502	0.676593
180	0.001642	0.007508	0.067130	12.86278	0.888469	0.662899
185	0.001605	0.007542	0.066369	12.86794	0.888567	0.665365
190	0.001481	0.007258	0.063776	12.87149	0.888428	0.672989

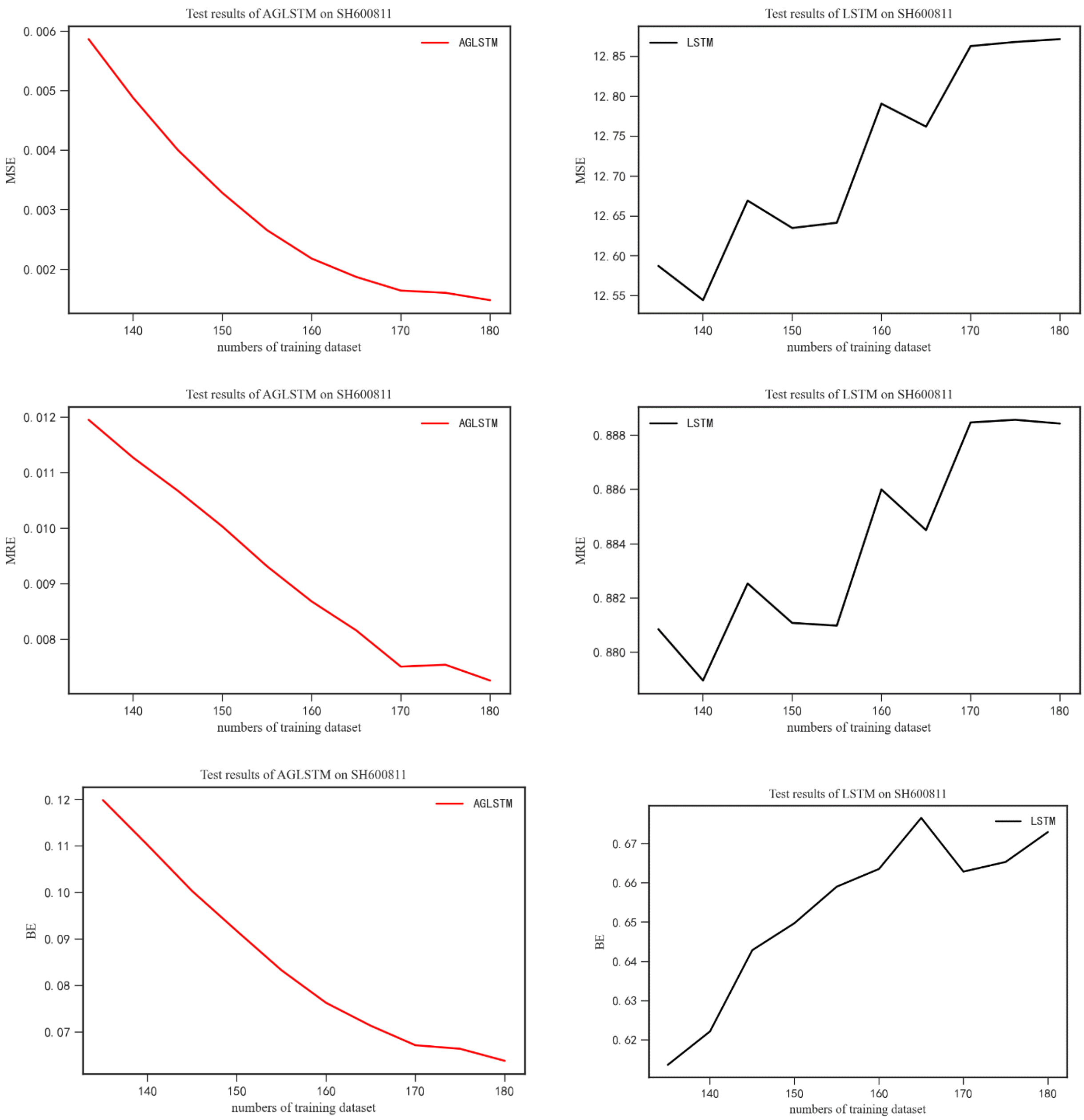

Figure 3: Stock 600811 error fluctuation in different training sets

Table 3: Error result of stock 603833

Number of training sets	FDMA-MSE	FDMA-MRE	FDMA-BE	LSTM-MSE	LSTM-MRE	LSTM-BE
145	29.24002	0.02302	0.430422	12738.65	0.995885	0.986544
150	26.61215	0.022014	0.410102	12735.67	0.995793	0.985404
155	24.76308	0.021179	0.39725	12734.33	0.995738	0.985446
160	28.17282	0.022299	0.421388	12737.59	0.995859	0.985905
165	26.03291	0.02137	0.405932	12737.42	0.995851	0.98595
170	25.40049	0.021057	0.400884	12737.56	0.995859	0.985826
175	32.87636	0.0232	0.452466	12736.98	0.995834	0.985907
180	31.31447	0.022603	0.443442	12737.13	0.995843	0.985788
185	26.44281	0.021119	0.40871	12736.96	0.995838	0.98572
190	24.53549	0.020267	0.396464	12736.92	0.995832	0.985855

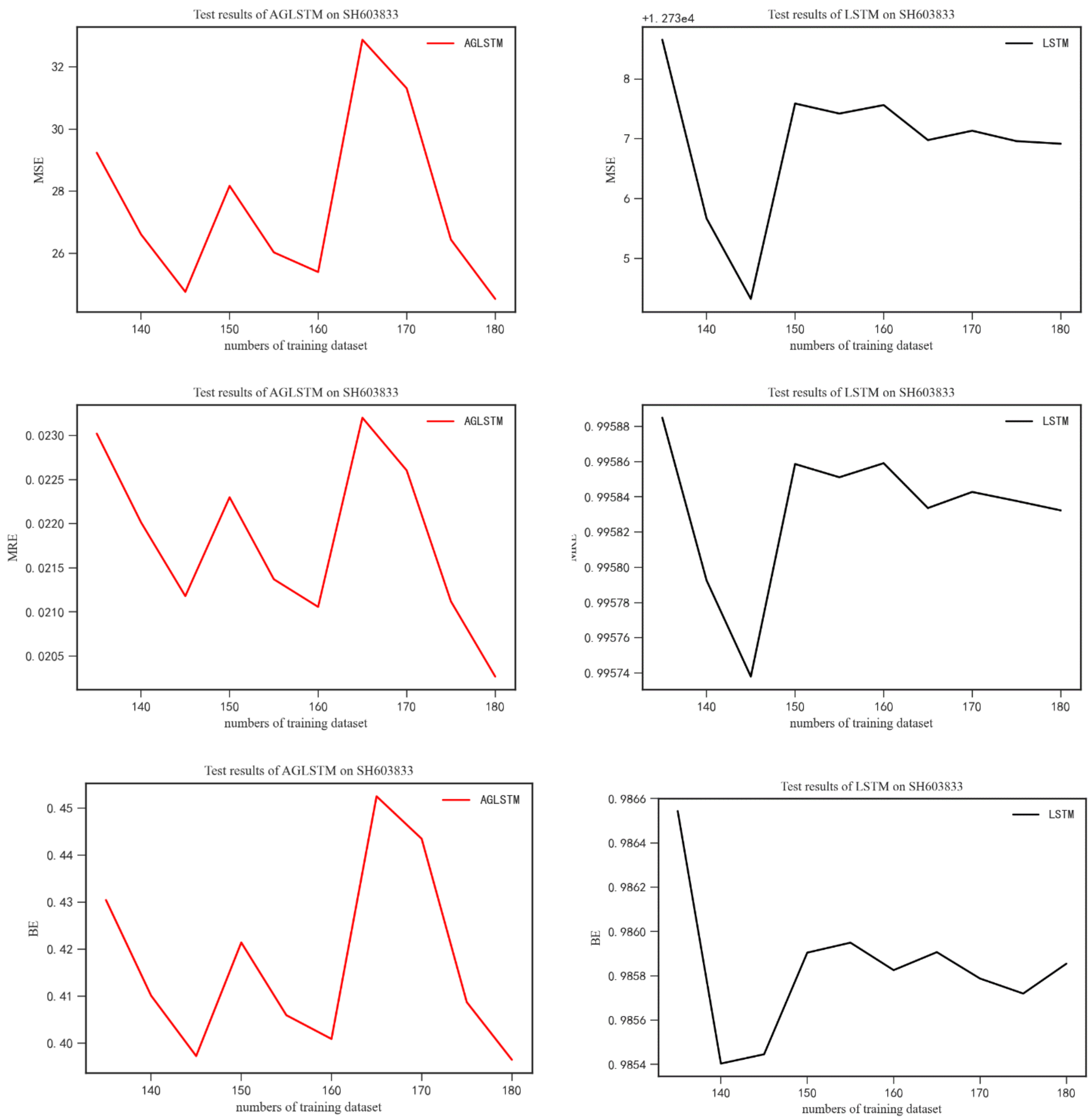

Figure 4: Stock 603833 error fluctuation in different training sets

Experimental results show that this method is superior to the original LSTM method in mean square error, relative absolute error and posterior error. In some experiments, it was three or four orders of magnitude lower than the original LSTM. In addition, it can be seen from Figure 2-4 that the performance of the original LSTM fluctuates to a certain extent when the number of training set samples is different. On the contrary, the error of the improved method decreases with the increase of the number of samples, which indicates that the improved method is more robust than the original LSTM. The proposed method is more accurate than the original LSTM in the middle and far distance experiment, because the additive model and average method are adopted to effectively balance the variance and deviation of the prediction model. Although most of the errors are calculated under the conditions of training set and test set, it can be proved that the proposed method can be well fitted to the timing curve. In conclusion, the proposed method can effectively improve the prediction accuracy of the original LSTM, and has certain robustness. It can be seen that the proposed method takes into account the functional information and nonlinear information of the day price, uses Bagging algorithm to fit the residual sequence generated by the original LSTM model, and uses the model average method to solve the step size selection problem of the LSTM neural network.

4. Conclusions

In this paper, combining PCA, FPCA and KPCA, an adaptive weighted LSTM model based on model average method is proposed, and the model is applied to the stock price prediction to test the model. Through the process of establishing the model in this paper and the analysis of the final prediction results, it is found that the advantages of this method lie in: Due to the rich information contained in the residual term [14], the functional information and nonlinear information based on intra-day prices adopted in this paper adopt the Bagging algorithm to fit the residual sequence, which can significantly improve the prediction effect of the model and improve the accuracy of the final stock price prediction. At the same time, the model information under different time Windows is integrated in the model, and the model average method is used to select the appropriate step size in the construction of LSTM adaptively. The application of the model average method can improve the effectiveness of the overall model.

For future work directions: Firstly, when Bagging algorithm is used for fitting in prediction, it is necessary to further determine its optimal base regression model. Secondly, this paper uses distance correlation coefficient to calculate the weight of the predicted value under multi-step size. This method belongs to the heuristic model averaging method, and the model averaging method is divided into three types: heuristic, autonomous and optimized. Therefore, we can explore other optimized model averaging methods, such as Mallos criterion, Bayes criterion and cross validation criterion.

References

[1] Li Z W, Han J, Song Y P. On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning [J]. Journal of Forecasting, 2020, 39(7): 1081-1097

[2] Engle R F, Rangel J G. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes [J]. The Review of Financial Studies, 2008, 21(3): 1187-1222

[3] Qi Yang, Xianbing Cao. Stock price analysis and prediction based on ARMA-GARCH model [J]. Practice and understanding of mathematics, 2016, 46(06): 80-86.

[4] Wei Dou. Stock Price Trend Forecast Based on LSTM [D]. Tianjin University, 2020. DOI: 10.27356/d. cnki. gtjdu. 2020.003177.

[5] Xingjun Xu, Gangfeng Yan. Stock Price Trend Analysis Based on BP Neural Network [J]. Zhejiang Finance, 2011(11): 57-59, 64

[6] Youen Chen, Yang Luo. A Review of Stock Forecasting Based on Portfolio Model [J]. Science and Technology and Industry, 2022, 22(11): 214-220.

[7] Yu Chen. Based on optimizing multi-dimensional characteristics of machine learning to pick stocks and timing

portfolio strategy to build [D]. Nanjing information engineering university, 2022. The DOI: 10.27248/ d.cnki.gnjqc. 2022.000489.
[8] Liu F, Cai M, Wang L, et al. An Ensemble Model Based on Adaptive Noise Reducer and Over-Fitting Prevention LSTM for Multivariate Time Series Forecasting [J]. IEEE Access, 2019, PP (99):1-1.
[9] Zhao Y, Ren X, Zhang X. Optimization of a Comprehensive Sequence Forecasting Framework Based on DAE-LSTM Algorithm [J]. Journal of Physics: Conference Series, 2021, 1746(1):012087 (12pp).
[10] Weijia Wu. Factor of compound factor quantitative stock selection scheme research [D]. Shanghai normal university, 2022. The DOI: 0.27312/d.cnki.gshsu.2022.000782.
[11] Mengmeng Yang, Xinghui Wang, Xing Zhao, Could enhance average method based on dynamic factor of stock price prediction [J]. Journal of Qilu University of Technology, 2022, 4 (5): 75-82. The DOI: 10.16442/ j.cnki. qlgydxxb. 2022.05.012.
[12] Chaobin Huang, Ximing Cheng. Research on stock price prediction based on LSTM neural network [J]. Journal of Information Science and Technology University of Beijing (Natural Science Edition), 2021, 36(1): 79-83.
[13] Yabo Zhao, Hua He. Stock Price Range Prediction Model Based on Information granulation and BP-Bagging [J]. Ploidy journal of Hebei university of technology, 2022 (05): 55 + 59-74. The DOI: 10.14081/j.cnki.hgdxb. 2022.05.005.
[14] Liu B, Tang X, Cheng J, et al. Traffic Flow Combination Forecasting Method Based on Improved LSTM and ARIMA[J]. 2019.