Remittances Review
May, 2023
Volume: 8, No: 3, pp. 212-226
ISSN: 2059-6588 (Print) | ISSN 2059-6596 (Online)

Received: 28 November 2022 Accepted: 28 March, 2023
DOI: https://doi.org/10.33182/rr.v8i3.15

Research on GEM Stock Price Prediction Based on Grey System and Neural Network

Jianjun Shen[1]

Abstract

Financial markets have become the most important part of every country's economic system, and the performance of financial markets will reflect the state of a country's economic development. In this paper, we use the gray system model and long and short-term memory network (LSTM) model to predict the stock price of GEM. Firstly, the paper introduces the algorithms related to gray system model and long and short-term memory network model, and proposes a new model hybrid scheme based on the LSTM model and GM(1,1) model with the core objective of stock price data prediction. This hybrid model can combine the advantages of the neural network model as well as the gray model and can get a better expected result around the stock price prediction problem. In terms of the fitting effect of the model, the fit of the gray system model is about 85%, the fit of the neural network model is about 86.3%, and the fit of the GM(1,1)-LSTM hybrid model is much improved and is above 90% overall. The GEM stock price prediction carried out in this paper is a very significant and valuable research problem in the current financial research field, and the effective analysis and prediction of stock prices can provide reference advice for investors and investment institutions in their investment decisions.

Keywords: *grey system, long and short-term memory network, hybrid model, GEM stock, price prediction*

Introduction

The boom in the financial industry in recent years has led investors to pay increasing attention to the allocation of financial assets (Nguyen, 2021). And stocks are becoming the focus of investors' attention in addition to relatively traditional investment and financial management methods such as savings and bonds (Göçken, Özçalıcı, Boru, & Dosdoğru, 2016). Predicting the prices of stocks and other financial assets is of great importance for investors(Chen et al., 2021; Dash & Dash, 2016; Yang, Guo, & Li, 2022). Predicting and allocating asset budgets is very challenging because there are many factors that can affect stock prices (Cui, Li, & Yu, 2019). Therefore, many investors use technical and quantitative methods to try to forecast asset price volatility (Ko & Chang, 2021). There is still no conclusive answer as to whether stock market movements can be predicted (Bao, Wei, Zhou, Jiang, & Watanabe, 2021).

The literature (Abdel-Nasser & Mahmoud, 2019) investigates deep learning and how it analyzes the complex and irregular behavior of financial data and shows that RNNs perform differently in

[1] Seoul School of Integrated Sciences & Technologies, Seoul, 03767, Korea.
Corresponding author: Jianjun Shen (13910075665@139.com)

Volume: 8, No: 3, pp. 212 - 226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

identifying trends and behaviors of time series. These models can store the information of time series for a period of time and perform better in identifying their behavior. The gradient vanishing problem is a difficult problem for RNNs with long sequences, and it is difficult to sort out the long-term dependencies between sequences. In the literature (Nikou, Mansourfar, & Bagherzadeh, 2019), in order to solve the nonlinear and hard-to-predict problems that exist in financial time series, studies have used neural networks, SVR, random forests, and LSTM, and LSTM was found to be the model with higher prediction accuracy.

In the literature (Zhang, Wang, Li, & Chen, 2016), an improved gas concentration prediction model was developed by improving the GM(1,1) model using BP neural network, and the experimental results showed that the model effectively improved the prediction accuracy. The literature (Hu, 2017) used different GM(1,1) models to validate the energy demand scenario in China. The experimental results show that the proposed FLNGM(1,1) model has better forecasting results compared to other gray residual correction models based on sign estimation. The literature (Hamayel & Owda, 2021) mentions that the long short-term memory network (LSTM) and the threshold recurrent unit (GRU) address the serious impact of this problem on the model performance to some extent.

In this paper, we study the prediction ability of the modified LSTM model and GM(1,1) model based on traditional RNN networks on stock closing price, and design and implement a set of GM(1,1)-LSTM hybrid model to improve the shortcomings of the two single models, LSTM model and GM(1,1,) model, on GEM stock price prediction. The article introduces the algorithmic processes related to the gray system model and the long and short-term memory network model, and carries out the research on the establishment and prediction of the GM(1,1) model and the LSTM model on the basis of the constructed dataset. Both of these models perform well on the stock closing price prediction task, but to further improve the prediction accuracy, a hybrid GM(1,1)-LSTM model is constructed in this paper and the performance of this hybrid model is experimented on the same dataset. The construction of different GEM stock price prediction models is of particular importance as a guide to the reform of China's financial market.

Gray Systems and Neural Network Models

Gray system theory

The GM(1,1) model, as a special case of the previously discussed GM(1, N) model, is an approximate differential differential equation model with differential, differential and other properties (Liming). The model has made a major breakthrough in modeling methods and ideas.

GM(1,1) modeling approach:

with the original non-negative series $X^{(0)}$:

$$X^{(0)} = \left\{ X^{(0)}(1), X^{(0)}(2), ..., X^{(0)}(n) \right\}, \quad X^{(0)}(k) \geq 0, k = 1, 2, \cdots, n \qquad (1)$$

$X^{(0)}$ of one accumulation to generate sequence $X^{(1)}$:

$$X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)\} \quad (2)$$

$Z^{(1)}$ is the sequence of immediately adjacent mean values generated by $X^{(1)}$:

$$Z^{(1)} = \{Z^{(1)}(1), Z^{(1)}(2), \dots, Z^{(1)}(n)\} \quad (3)$$

Among them:

$$X^{(1)}(k) = \sum_{i}^{k} X^{(0)}(i) \quad (4)$$

$$Z^{(1)}(k) = 0.5 * X^{(1)}(k) + 0.5 * X^{(1)}(k-1), k = 1, 2, \cdots, n \quad (5)$$

The differential equation of this change law:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (6)$$

In order to have only one variable in the model, u in the above differential equation must be an endogenous variable, which is classified as a parameter to be defended in GM(1,1), and the parameters of the parameter to be defended are listed as :

$$\hat{a} = \begin{pmatrix} a \\ u \end{pmatrix} \quad (7)$$

In this case, the differential equation is a linear combination containing only $\frac{dx}{dt}$ and the background value x:

$$a^{(1)}[x^{(1)}(k+1)] + ax^{(1)}(k+1) = u \quad (8)$$

Among them

$$a^{(1)}[x^{(1)}(k+1)] = x^{(0)}(k+1) \quad (9)$$

$$x^{(1)}(k+1) = \frac{1}{2}[X^{(1)}(k) + X^{(1)}(k+1)] \quad (10)$$

Substituting the above two equations, we have the generated sequence:

$$\begin{aligned} x^{(0)}(2) &= a[-0.5(x^{(1)}(1) + x^{(1)}(2))] + u, (k=1) \\ x^{(0)}(3) &= a[-0.5(x^{(1)}(2) + x^{(1)}(3))] + u, (k=2) \\ &\vdots \\ x^{(0)}(n) &= a[-0.5(x^{(1)}(n-1) + x^{(1)}(n))] + u, (k=n) \end{aligned} \quad (11)$$

May 2023
Volume: 8, No: 3, pp. 212 - 226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

For simpler writing, the following matrix notation is introduced

$$y=\begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \cdots \\ X^{(0)}(n) \end{bmatrix},\ \chi=\begin{bmatrix} -(X^{(1)}(1)+X^{(1)}(2))/2 \\ -(X^{(1)}(2)+X^{(1)}(3))/2 \\ \vdots \\ -(X^{(1)}(n-1)+X^{(1)}(n))/2 \end{bmatrix},\ E=\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (12)$$

then there are

$$y=a\chi+uE=[\chi\vdots E]\cdot\begin{bmatrix} a \\ u \end{bmatrix}=[\chi\vdots E]\cdot\hat{a} \quad (13)$$

Let $b=[\chi\vdots E]$, then we have

$$b=\begin{pmatrix} -(X^{(1)}(1)+X^{(1)}(2))/2 & 1 \\ -(X^{(1)}(2)+X^{(1)}(3))/2 & 1 \\ \vdots & \vdots \\ -(X^{(1)}(n-2)+X^{(1)}(n-1))/2 & 1 \\ -(X^{(1)}(n-1)+X^{(1)}(n))/2 & 1 \end{pmatrix} \quad (14)$$

Using the least squares method. Find the least squares solutions of the model parameters:

$$\hat{a}=\begin{pmatrix} a \\ u \end{pmatrix}=(b^t b)^{-1}b^t y \quad (15)$$

Let $x^{(0)}$ be a non-negative sequence, $x^{(1)}$ be a once-accumulated generating sequence of $x^{(0)}$, $z^{(1)}$ be an immediately adjacent mean generating sequence of $x^{(1)}$, and $\begin{pmatrix} a \\ u \end{pmatrix}=(b^t b)^{-1}b^t y$, call $\frac{dx^{(1)}}{dt}+ax^{(1)}=u$ a shadow equation of the gray differential equation $x^{(0)}(k)+az^{(1)}(k)=u$, or white equation (Dong & Zhao, 2021).

The derivation of the predicted values proceeds as follows:

(1) The solution of the shadow equation $\frac{dx^{(1)}}{dt}+ax^{(1)}=u$, also known as the solution of the time response function:

$$x^{(1)}(t)=[x^{(1)}(0)-\frac{u}{a}]e^{-at}+\frac{u}{a} \quad (16)$$

(2) Time response series of the gray differential equation $x^{(0)}(k)+az^{(1)}(k)=u$:

$$\hat{x}^{(1)}(k+1)=[x^{(1)}(0)-\frac{u}{a}]e^{-ak}+\frac{u}{a},(k=1,2,\cdots,n) \tag{17}$$

(3) $x^{(1)}(0)=x^{(0)}(1)$, then we have

$$\hat{x}^{(1)}(k+1)=[x^{(0)}(1)-\frac{u}{a}]e^{-ak}+\frac{u}{a},(k=1,2,\cdots,n) \tag{18}$$

(4) Reduction gives

$$\hat{X}^{(0)}(k+1)=\hat{X}^{(1)}(k+1)-\hat{X}^{(1)}(k) \tag{19}$$

Recurrent neural network theory

Human neurons are one of the principles of recurrent neural networks. Figure 1 shows its structure. When human beings think about a certain problem, they usually make a decision based on their recent "memory" instead of thinking from the beginning of the problem, because human memory is limited and they cannot remember all the relevant information completely. For example, a person likes to travel, and the place he or she wants to visit most is Guizhou, and he or she must visit a certain place sometime in the future. Normally, to complete the sentence, a certain place should be Guizhou. This is the judgment that a person makes based on the recent memory of the context of this sentence, so the recurrent neural network is designed so that a simple artificial neural network can have the ability of short-term memory. Simple artificial neural networks are not very good at handling problems with correlation because they use the error back propagation principle, while recurrent neural networks are better suited to handle that type of problem.

However, the recurrent neural network was not well used when it was first proposed due to the limitation of the immature development of computer technology, and it was only gradually applied to various fields after the computer technology was deeply learned and developed.

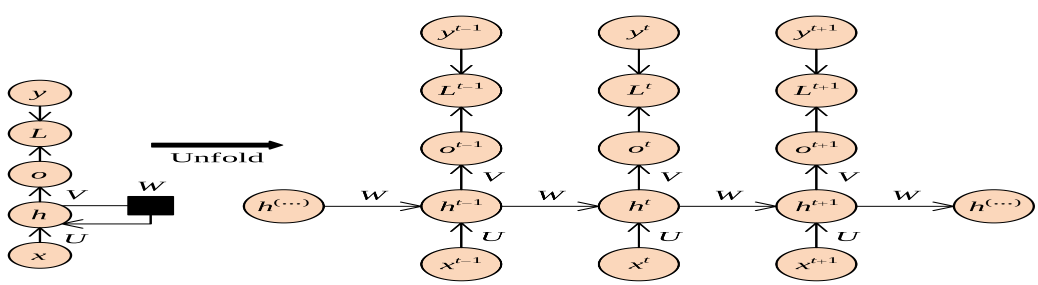

Figure 1: Recurrent neural network unfolding structure diagram

Long Short Term Memory Network LSTM

The long and short-term memory network model can solve the problem that traditional recurrent neural networks are prone to gradient disappearance and gradient explosion, and Figure 2 shows the schematic diagram of its structural unfolding. This model is improved on the basis of RNN model, which can effectively extract the required features from the data when the data volume is relatively large. Therefore, the LSTM model can be more effective in modeling and predicting the analysis of financial time series data.

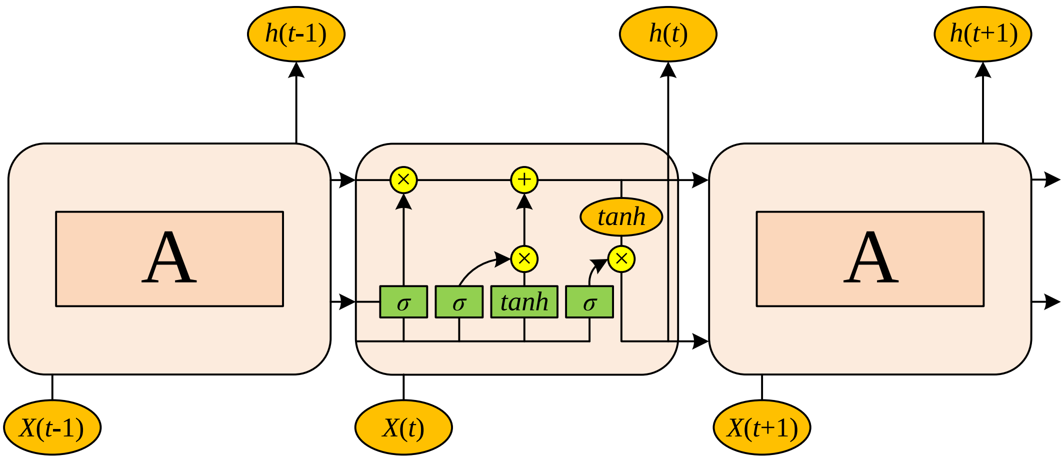

Figure 2: LSTM memory cell structure diagram

Three gating structures are added to the LSTM network model, namely forgetting gates, input gates and output gates.

The three gates in the network structure of the long- and short-term memory model have different roles, where the forgetting gate mainly determines how much of the cell state at moment $t-1$ needs to be saved to moment t; the input gate mainly determines how much of the input $X(t)$ at moment t needs to be saved to the cell state; and the output gate mainly controls how much of the cell state in the input gate needs to be output as $h(t)$.

The expression of the forgetting gate is:

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \qquad (20)$$

The value of f_t in equation (20) ranges from 0 to 1. The smaller the value, the more forgetting there is, and vice versa, the less. The calculation equation is as follows:

$$i_t = \sigma\left(W_u h_{t-1} + U_i x_t + b_i\right) \qquad (21)$$

$$C_t = \tanh\left(W_c h_{t-1} + U_C x_t + b_c\right) \qquad (22)$$

The expression of the input gate is:

$$C_t = C_{t-1} * f_{t+} \tilde{C}_t * i_t \qquad (23)$$

May 2023
Volume: 8, No: 3, pp. 212-226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

The cell state at moment $t-1$ plus the new input information can update the cell state at the current t moment, and the output gate can then extract information from the updated cell state to generate the hidden state h_t.

The expression of the output gate is:

$$o_t = \sigma\left(W_o h_{t-1} + U_o x_t + b_o\right) \qquad (24)$$

$$h_t = o_t * \tanh\left(C_t\right) \qquad (25)$$

Equation (25) in the value range of o_t is also between 0 and 1, the above equations can be calculated by associating the final information left behind, the formula is as follows:

$$h_t = o_t * \tanh\left(C_{t-1} * f_t + i_t * \tanh\left(W_c h_{t-1} + U_C x_t + b_C\right)\right) \qquad (26)$$

$$C_t = C_{t-1} * f_t + i_t * \tilde{C}_t \qquad (27)$$

From the derivation of the above equation, it can be seen that when the forgetting gate is opened, the gradient of cell state C_t can be passed to C_{t-1} at the moment of $t-1$.

In recent years, scholars have begun to devote research to the analysis and prediction of stock price data.

Due to the extremely complex and influential factors in the financial market environment, there are extremely many market factors that influence the trend of each single stock.

It is difficult to comprehensively extract the hidden patterns and change characteristics in the trend using traditional time series analysis models, but deep learning-based models have certain advantages in this regard.

Construction and Study of Stock Price Prediction Model Based on GM(1,1) and LSTM

Research on stock price data prediction model based on LSTM model

GM(1,1) model development and prediction

In this chapter, GM(1,1) modeling will be performed on the basis of the closing price data of Hanbang Hi-Tech (300449).

The raw series are shown in Table 1 below.

The data are obtained from the stock price dataset constructed in this paper, and the data also ensure the objectivity and correctness of the experimental results.

Table 1: Zhaoxun Media Closing Price Data (Unit: RMB)

DataSample	Date	ClosingPrice
Hanbang Hi-Tech(300449)	8.06	24.11
	8.07	23.74
	8.08	23.56
	8.09	25.15
	8.10	23.47
	8.13	23.83
	8.14	23.84
	8.15	25.97
	8.16	23.24
	8.17	25.55
	......	......

First of all, the first order cumulative series of closing price data of Hanbang Hi-Tech (300449) between August 6, 2020 and July 31, 2020 is generated in a group of 10 days. The initial series is x_0, the cumulative series is X_1, and the development coefficient is α. The development coefficient $\alpha = -0.0024$ is calculated, and from the limitation of the GM(1,1) model, the development coefficient $-\alpha = 0.0024 < 0.3$ is calculated, which indicates that the model can be used to predict the closing price of Hanbang Hi-Tech (300449), and also ensures the feasibility of the modeling. After modeling calculations and predictions obtained Hanbang Hi-Tech (300449) closing price fitted value, because GM (1,1) for the stock closing price of such values up and down unstable time series data in the longer the length of time the worse the fitting effect, so the subject will use the sliding window to take the value of the method to rearrange the fitted value, to August 19 fitted value as an example, the day of the stock closing price fitted value for August 6 to The fitted value of the stock closing price on that day is the fitted value of the eleventh day of modeling prediction from August 6 to August 17 for ten days as the original series, and similarly, the fitted value of August 20 is the fitted value of the eleventh day of modeling prediction from August 7 to August 19 as the original series, and the fitted situation is shown in Table 2.

Table 2: GM(1,1) Model Hanbang Hi-Tech (300449) Fitted Values

Data Sample	Date	Closing Price	Fitted value
Hanbang Hi-Tech(300449)	8.19	26.64	26.21
	8.20	26.33	26.01
	8.21	26.68	26.34
	8.22	26.59	26.28
	8.23	25.41	25.13
	8.26	26.15	25.96
	8.27	25.46	25.31
	8.28	26.52	26.14
	8.29	25.93	25.43
	8.30	25.58	25.12

Modeling effect analysis

Analysis of statistical indicators:

In this paper, the posterior difference test is used to test the accuracy of the grey prediction model about the test. Now two indicators will be used to test the prediction results of the above experiments. They are the ratio of posterior differences C and the probability of small errors P. The two indicators can be combined to judge the prediction effectiveness of the gray model. For the degree of fit of the closing prices of the six different stock samples, this paper will use these two evaluation metrics to measure the GM(1,1) model. Table 3 shows the calculation results. The GM(1,1) model fits the closing prices of six different stocks such as Hanbang Hi-Tech (300449) better and the model accuracy reaches the ideal level.

Table 3: GM (1, 1) model fitting results

Data Sample	Ratio of post-test difference C	Small error probability P	Model accuracy
Ouya Design(300949)	0.12	0.9057	Excellent
Jiji Corporation(300553)	0.27	0.9472	Excellent
Kangtai Bio(300601)	0.29	0.9341	Excellent
Tianmai Technology(300807)	0.15	0.9751	Excellent
Sunlight Power(300274)	0.49	0.8451	General
Hanbang Hi-Tech(300449)	0.57	0.9347	Qualified

In this section, a GM(1,1) model is built and forecasted based on the historical closing price data of six stocks, and the new time series is obtained by accumulating the initial series, and the development factor α can be calculated. In other words, the model is theoretically valid and reasonable.

The modeling analysis and prediction of the closing prices of six different stocks can be judged by the visual comparison of the fitting results and the analysis of statistical indicators. The GM(1,1) model performs relatively well in the problem of stock closing price prediction and can predict the trend of the data as a whole, but there are still cases where the prediction effect is not satisfactory and the prediction effect is relatively unstable, which may be due to the fact that the stock prices are influenced by more factors that can affect the prediction effect of the model.

The experimental results are examined using the posterior difference test. The ratio C of the posterior difference of the model and the probability of small error P can be obtained, and the accuracy of the model can be judged whether it meets the expected standard by comparing it with the model accuracy reference standard introduced in the previous section.

Basically, it can be concluded that the GM(1,1) model is reasonable and effective in modeling on the basis of stock closing prices, and it is guaranteed and credible for subsequent predictions.

Remittances Review
May 2023
Volume: 8, No: 3, pp. 212 - 226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

Research on stock price data prediction model based on LSTM model

In this project, on the same data sample as the GM(1,1) model modeling, for the stock samples screened above, the individual stock samples involved in the experimental measurement are taken as the training set with a total of 1811 historical closing price data from the first trading day of 2020 to September 6, 2021, and the latter 250 historical closing price data as the test set for stock closing price prediction experiments.

Here the modeling and prediction of LSTM model was carried out using the closing prices of four individual stocks as the data base, and the prediction effect is shown in Figure 3~Figure 6, where the red line is the actual value and the blue line is the predicted value.

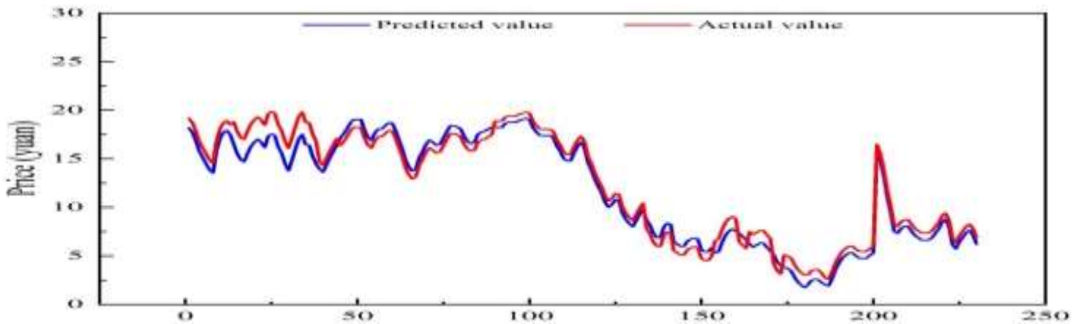

Figure 3: Ouya Design

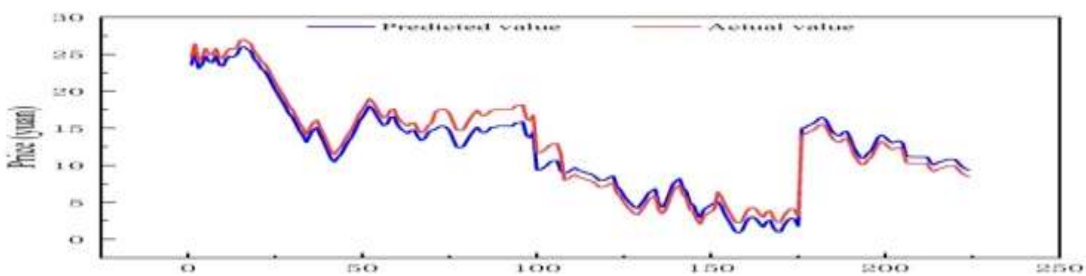

Figure 4: Jizhi Corporation

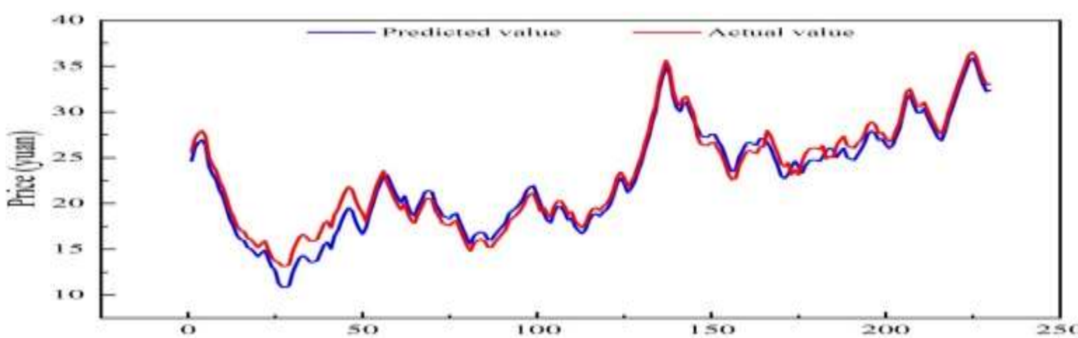

Figure 5: Kangtai Bio

Remittances Review
May 2023
Volume: 8, No: 3, pp. 212-226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

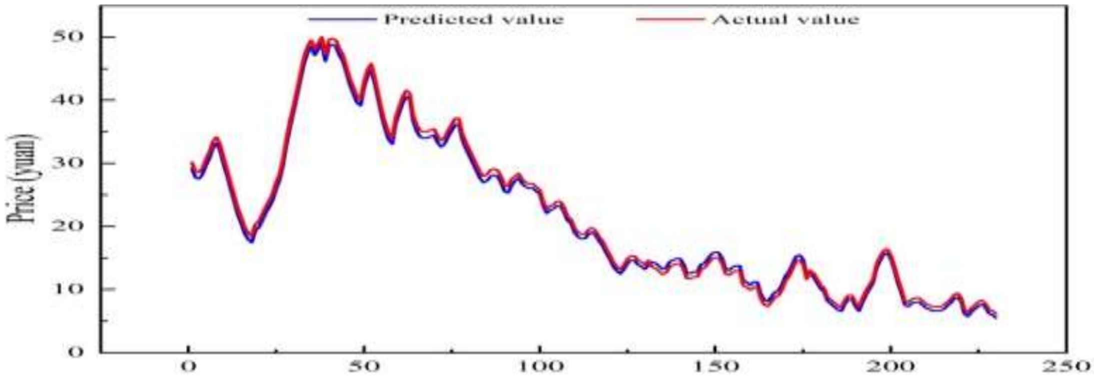

Figure 6: Hanbang Hi-Tech

The prediction error of the model is small on most of the stock samples. However, there are still some samples with large errors between the predicted and actual values, for example, Aoya Design (300949) and Jiji (300553). By analyzing this data sample, we can get that the data base of the above stocks is larger compared with other stocks, and their closing price data are generally in the thousands, while the closing prices of most other stock samples range from a few dollars to a few hundred dollars, and the above two stocks are different from other stocks The difference between the above two stocks and other stocks is large, so the fluctuation range of the forecast value is also much higher than most other stocks sample.

Through the analysis of the above experimental results, we can conclude that the long and short term memory network model applied to the stock closing price prediction problem and achieved the desired results, but in some sample data, we need to make the corresponding normalization process and the adjustment of the model parameters to make the model get a more ideal effect.

In this chapter, a variant of the traditional RNN-based network, the Long Short Term Memory Network (LSTM), is selected for modeling prediction in the stock closing price prediction problem to compensate for the shortcomings of the traditional RNN.

Construction and Study of GM(1,1)-LSTM Hybrid Model

Combines the two models based on the above study, explain the implementation scheme and principles of the hybrid model, and conduct prediction analysis comparison experiments on the same data set for the hybrid model, so that we can make a reasonable analysis and draw conclusions on the effect of the hybrid model on the stock price prediction problem in a more objective way.

Architecture principle of hybrid model and model implementation

The hybrid model proposed in this study is based on the LSTM model and the GM(1,1) model, both of which perform relatively well in predicting stock price data, but there is a significant difference between the two. magnitude is higher. There are many factors affecting stock trading data, and a particular model may not be able to predict all the changes well, so in this paper, a new

May 2023
Volume: 8, No: 3, pp. 212 - 226
ISSN: 2059-6588 (Print) | ISSN: 2059-6596 (Online)

hybrid model will be constructed based on the long and short-term memory network model as well as the gray model, and the prediction results will be fused by the average weighting method, and the formula is shown as follows.

$$Value = \frac{Weight_{LSTM} * Value_{LSTM} + Weight_{GM(1,1)} * Value_{GM(1,1)}}{2} \quad (28)$$

As shown in Equation (28), the hybrid model is able to weight the prediction results of the two sub-models with different weights to obtain the final hybrid model prediction values.

Since stock prices are affected by many factors, the same model will have different prediction accuracy on different data samples, therefore, the weights of the sub-models of the hybrid model will be different in different data samples, and this topic will subsequently use the data standard deviation to calculate the specific proportion of the weights of the sub-models, which is calculated as shown below:

$$Weight_{LSTM} = \left(1 - \frac{LSTM_{diff}}{LSTM_{diff} + GM(1,1)_{diff}}\right) \times 2 \quad (29)$$

Equation (29) in:

$$Weight_{GM(1,1)} = 2 - Wgight_{LSTM} \quad (30)$$

$$LSTM_{diff} = Sample_{std} - LSTM_{std} \quad (31)$$

$$GM(1,1)_{diff} = Sample_{std} - GM(1,1)_{std} \quad (32)$$

From the results of the study, it can be obtained that the prediction results of the LSTM model are less volatile, so the LSTM model should be considered to occupy a larger weight when the fluctuation of the predicted data samples is small.In addition, in the statistics of the fluctuation of the time series, the standard deviation can not only be very intuitive to determine the fluctuation of a certain time series, but also can clearly reflect the rise and fall of the series.

Comparison of prediction results of LSTM model, GM(1,1) model and hybrid model

Considering the specificity of stock price data, in order to make a more objective comparison of model prediction performance, this topic will build corresponding model predictions for different stocks with the same time interval data, and a total of 1811 historical closing price data from the first trading day of 2020 to September 6, 2021 in each sample data set will be used as the training set, and the data of the last 10 working days will be used as the test set. The LSTM model and GM(1,1) model were used for the prediction experiments respectively, and then the hybrid model prediction values were calculated based on determining the weight of each model in the hybrid model. Taking Hanbang Hi-Tech (300449) as an example, the prediction results of LSTM model, GM(1,1) model and hybrid model are shown in Table 4.

It can be seen that the hybrid model achieves better results in all indicators compared with the single model, which proves that the hybrid model performs better than the single model in most of the individual stock sample forecasting tasks. By comparing the indicators, it is found that the experimental results of the GM(1,1) model group and the LSTM model group are also more volatile compared to the hybrid model.

The experimental results are also in able to tentatively prove that the hybrid model constructed in this topic has some improvement in the stock closing price prediction problem. During the experiments, it was found that the GM(1,1) model was closer to the true value than the LSTM model for some samples, but the LSTM model was able to fit the trend of the true value better. After the hybrid calculation, the error fluctuation of the hybrid model can be reduced and the hybrid model prediction results are closer to the actual values

Table 4: Comparison of model prediction results

Date	Actual value	GM(1,1) predicted value	LSTM predicted values	Hybrid model predicted values
20200819	26.64	26.21	26.45	26.33
20200820	26.33	26.01	26.25	26.13
20200821	26.68	26.34	26.58	26.46
20200822	26.59	26.28	26.43	26.355
20200823	25.41	25.13	25.28	25.205
20200826	26.15	25.96	26.11	26.035
20200827	25.46	25.31	25.52	25.415
20200828	26.52	26.14	26.35	26.245
20200829	25.93	25.43	25.64	25.535
20200830	25.58	25.12	25.33	25.225

Conclusion

The GM(1,1) model was developed and predicted on the basis of the constructed dataset, and it was concluded through experiments that the model is more suitable for medium and long-term prediction, and it is theoretically possible to model and predict the stock price data effectively. At the same time, the LSTM model based on the traditional RNN network was chosen to forecast the stock closing price. The experimental results conclude that the LSTM model can better explore the hidden patterns of the stock closing price historical data and can make effective and objective forecasts. In the prediction experiments of the selected stock samples, good prediction results were also achieved, which also shows that the LSTM model can perform well in the task of stock price prediction. A hybrid model scheme based on the LSTM model and the GM(1,1) model was developed and experiments were conducted to show that the hybrid model can perform better and more consistently on the same data samples than the individual sub-models. From the fitting results of the three models, the fitting degree of the gray system model is about 85%, the fitting degree of the neural network model is about 86.3%, and the fitting degree of the GM(1,1)-LSTM hybrid model is much improved and is above 90% overall. The experimental results show that the hybrid model still has certain advantages in comparison with the traditional forecasting models, and also verify that the hybrid model proposed in this topic has certain advantages, effectiveness and objectivity in the forecasting ability of stock price data.

References

Abdel-Nasser, M., & Mahmoud, K. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Computing and Applications, 31*, 2727-2740.

Bao, Z., Wei, Q., Zhou, T., Jiang, X., & Watanabe, T. (2021). Predicting stock high price using forecast error with recurrent neural network. *Applied Mathematics and Nonlinear Sciences, 6*(1), 283-292.

Chen, Y., Fang, R., Liang, T., Sha, Z., Li, S., Yi, Y., . . . Song, H. (2021). Stock price forecast based on CNN-BiLSTM-ECA Model. *Scientific Programming, 2021*, 1-20.

Cui, W., Li, B., & Yu, D. (2019). Empirical analysis of stock price forecast based on GARCH Model and BP neural network model. *J Tianjin Normal Univ (Nat Sci Ed), 5*(39), 30-34.

Dash, R., & Dash, P. (2016). Efficient stock price prediction using a self evolving recurrent neuro-fuzzy inference system optimized through a modified differential harmony search technique. *Expert Systems with Applications, 52*, 75-90.

Dong, W., & Zhao, C. (2021). Stock price forecasting based on Hausdorff fractional grey model with convolution and neural network. *Mathematical Biosciences and Engineering, 18*(4), 3323-3347.

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Systems with Applications, 44*, 320-331.

Hamayel, M. J., & Owda, A. Y. (2021). A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. *AI, 2*(4), 477-496.

Hu, Y.-C. (2017). Grey prediction with residual modification using functional-link net and its application to energy demand forecasting. *Kybernetes*.

Ko, C.-R., & Chang, H.-T. (2021). LSTM-based sentiment analysis for stock price forecast. *PeerJ Computer Science, 7*, e408.

Liming, M. J. W. L. Z. Stock Price Prediction of GEm based on Grey System and Neural Network. *评论员文章*, 97.

Nguyen, D. T. H. (2021). Firm performance in the financial industry: the role of foreign entrants and their origin. *Multinational Business Review*.

Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management, 26*(4), 164-174.

Yang, S., Guo, H., & Li, J. (2022). CNN-GRUA-FC Stock Price Forecast Model Based on Multi-Factor Analysis. *Journal of Advanced Computational Intelligence and Intelligent Informatics, 26*(4), 600-608.

Zhang, S., Wang, B., Li, X., & Chen, H. (2016). Research and application of improved gas concentration prediction model based on grey theory and BP neural network in digital mine. *Procedia Cirp, 56*, 471-475.