Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 4, No. 2, 2023

LSTM-Based Stock Price Prediction

Hongwei Lu *

College of Computer Science, Chengdu University, Chengdu, 610100, China
*** Corresponding author Email:** hongweiluhappy@gmail.com

Abstract: Predicting stock prices is a job that researchers and analysts have been working on for many years. Investors have shown great interest in this area so that they can better manage their assets. Accurately predicting changes in stock prices in the market can generate huge economic benefits. In view of the high noise, nonlinearity and non-stationarity of stock price data, which makes it very difficult to accurately predict the stock price, this paper intends to use the long-short-term memory network (Long Short-Term Memory, LSTM) Recurrent Neural Network (RNN) architecture to establish A model that predicts the future value of a stock.

Keywords: Stock Market; Recurrent Neural Network; Long Short-term Memory.

1. Introduction

1.1. Research Background

With the continuous development of computer technology and the advent of the era of big data, data acquisition is easier, the data management system is more perfect, and the data quality is constantly improving [1]. Financial data research is also becoming more mature. In the long-term research and exploration, people found that financial data does not conform to the normal distribution commonly used in theoretical research, but has nonlinear characteristics. The rise of machine learning has opened up a path for solving nonlinear problems. Machine learning models are different from traditional quantitative models. By imitating the structure of neurons in the human brain, machine learning models try to reproduce the way humans perceive the world. A neural network is essentially a collection of thousands of linear and non-linear functions, through the joint action of these functions to analyse the distribution of data.

1.2. Research Significance

In the network of deep learning, RNN (Recurrent Neural Network) has the concept of time series with its unique self-connection structure, which can be used to process time series data and automatically learn the time relationship between data. However, there will be two major problems in the training process of RNN hidden layer, namely gradient explosion and gradient disappearance, so RNN is not suitable for training longer time series data. LSTM (Long Short-Term Memory Neural Network) is improved from RNN. Its hidden layer adds a gate structure to filter information, which can alleviate the phenomenon of gradient disappearance and gradient explosion [2].

1.3. Literature Review

Theoretically, there are three classic theories recognized by investors in predicting the trend of stock prices, namely efficient market hypothesis, behavioral finance theory, and adaptive market. Related studies on stock price trend forecasting are described. In terms of means, the earliest methods based on econometric models include ARIMA model, GARCH model and so on. Introduce machine learning and deep learning methods later, such as GRU, LSTM models, etc [3]. for stock price prediction. In the end, the optimization of the model based on the method of deep learning, and the attempt to improve the prediction accuracy of the model by fusing stock data from different sources are elaborated in detail.

In recent years, with the rise of machine learning and deep learning, many researchers have begun to focus on using new technologies to study problems. LSTM is an improvement to RNN (cyclic neural network), adding a " gate " structure in the hidden layer to reduce the probability of gradient disappearance and gradient explosion. With the deepening of research, especially the maturity of natural language processing technology, researchers began to try to analyze stocks from other aspects [4].

1.4. Research Content and Main Problems to be Solved in this Paper

(1) Theoretical knowledge. Familiar with the principle of LSTM.

(2) Feature selection. Based on the thermal feature map, the feature evaluation of historical stock transaction data is carried out to lay a theoretical foundation for the subsequent fusion with stock data from other sources.

(3) Multi-source fusion. Start research in the intersecting fields of finance and computer, and build an LSTM stock price prediction model based on multi-source data fusion.

(4) This article uses the stock price of Jiangsu Subote New Material Co., Ltd, and uses the LSTM model to make predictions.

2. Introduction to LSTM Model

2.1. LSTM Neural Network

In 1997, Sepp Hochreiter, Jürgen Schmidhuber and others proposed LSTM, the long short-term memory neural network, to solve the long-term dependence phenomenon in RNN due to the chain rule. In the previous section, we proposed that RNN (Recurrent Neural Network) is prone to two major phenomena, namely, gradient disappearance and gradient explosion. And LSTM reduces the possibility of these two phenomena very well. This section first introduces the network structure and principle of LSTM, and then introduces the parameter learning in LSTM.

The basic structure diagram of LSTM is shown in the figure below. The loop structure of LSTM is more complicated, and a gating structure is added inside it, that is, three gates (forgetting gate, input gate, output gate) are added to decide which information to forget, and What information to continue to pass on. The so-called gate means a value between (0,1). If the value is closer to 1, it means that the door is more open, and the information is more likely to be passed on; Closer to 0, it means that the more the door tends to be closed, the more likely the information will be forgotten. Therefore, in LSTM, the sigmoid function is usually used to construct the gate. The main function of LSTM is to process long-term sequence data, so that the network can remember relatively distant information hidden in the time series data. The following will introduce the various gates of LSTM in detail structure [5].

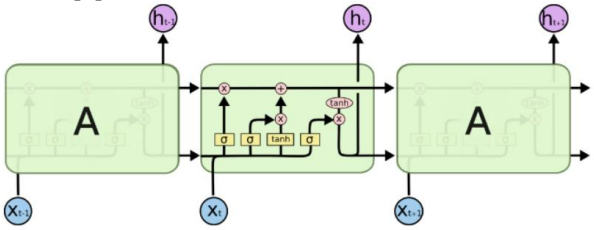

Figure 1. LSTM basic structure diagram

1. Symbol representation

x_t: the t-th input vector in the time series;

$\hbar_t$: The output vector of the hidden state at time t;

tanh: tanh activation function;

σ: sigmoid activation function;

C_t: The actual value of the cell state at time t;

$\widehat{C}_t$: Candidate value of the cell state at time t;

f_t: Forget the control value of the gate layer at time t;

i_t: Input the control value of the gate layer at time t;

o_t: Output the control value of the gate layer at time t;

W_f: Each time step forgets the weight matrix of the gate layer;

W_i: Each time step inputs the weight matrix of the gate layer;

W_o: Each time step outputs the weight matrix of the gate layer;

W_C: the weight matrix of the cell state at each time step;

W_y: weight matrix from hidden layer to output layer at each time step;

b_f: The bias vector of the forget gate layer;

b_i : The bias vector of the input gate layer;

b_o: the bias vector of the output gate layer;

b_C: the bias vector of the cell state;

b_y: the bias vector of the output layer;

⊙: Indicates the multiplication between corresponding elements in different matrices

2.2. Principle of LSTM

2.2.1. Forgetting Gate Layer

The local structure diagram of the LSTM forgetting gate layer is shown above. Its function is to control the proportion of useless information that can be forgotten between the previously transmitted information and the current input information. Therefore, in the forgetting gate structure, the LSTM layer will calculate the forgetting gate layer control parameters f_t[6].

$$f_t = \sigma\left(W_f \cdot [\hbar_{t-1}, x_t] + b_f\right) \quad (1)$$

2.2.2. Input Gate Layer

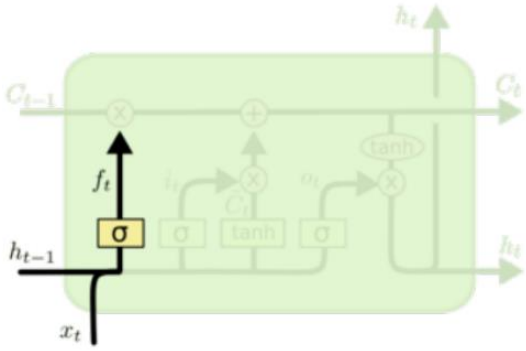

Figure 2. Local structure diagram of LSTM forgetting gate layer

The main function of the LSTM input gate layer is to update information, including two steps, the first is to determine new information, and the second is to update information.

(1) Determination of new information

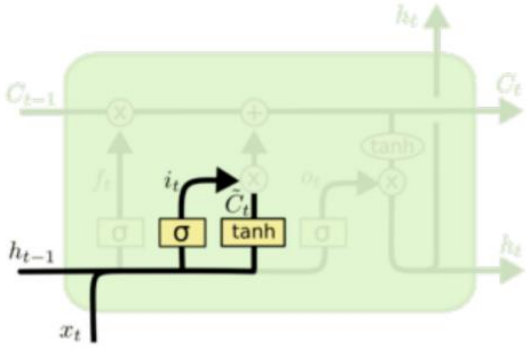

Figure 3. LSTM input gate layer determines new information map

In this step, it will be determined what new information enters. First, the LSTM layer will calculate the control parameters of the input gate layer number, so as to determine the proportion of new information that can be transmitted, and then calculate the candidate value of the hidden state $\widehat{C}_t$.

$$i_t = \sigma(W_i \cdot [\hbar_{t-1}, x_t] + \hbar_{t-1}, b_i) \quad (2)$$

$$\widehat{C}_t = tanh(W_C \cdot [\hbar_{t-1}, x_t] + \hbar_{t-1}, b_C) \quad (3)$$

(2) Information update

In this step, all new information entered in the previous step will be determined, and what new information can be updated. At this time, the LSTM layer will calculate the actual value of the hidden state.

$$C_t = f_t \odot C_{t-1} + i_t \odot \widehat{C}_t \quad (4)$$

2.2.3. Output Gate Layer

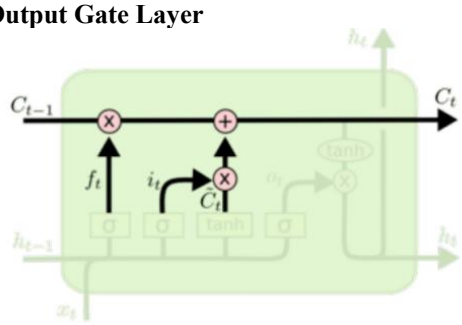

Figure 4. LSTM input gate layer information update diagram

The output gate layer is the last gate in the LSTM gate structure. Its main function is to determine what information can be output and control the proportion of LSTM layer information output. Therefore, in this step, the output gate control parameters will be calculated first, and finally Then

calculate the output value h of the LSTM layer. The partial structure diagram of the LSTM output gate layer is shown below.

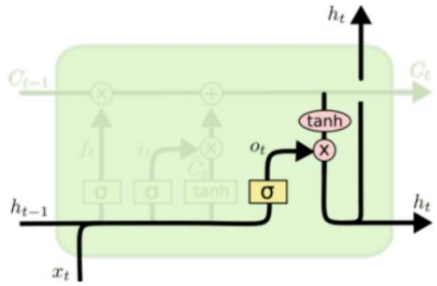

Figure 5. LSTM input gate layer partial structure diagram

$$o_t = \sigma W_i \cdot [h_{t-1}, x_t] + h_{t-1} + b_o \quad (5)$$

$$h_t = o_t \odot tanh(C_t) \quad (6)$$

LSTM controls the rejection and addition of information through the "gate structure". In the above formula, h is the output state, which can be to store short-term memory. It is the cell state, used to store long-term memory. Therefore, LSTM because of its unique structure make it capable of handling long dependent data.

3. Experimental Results and Analysis

3.1. Selection of Eigenvalues

Determine four characteristic values according to the characteristic heat map, namely: closing price, highest price, lowest price, and opening price. The feature heat map is shown below [6].

Figure 6. Feature heat map

3.2. Result Display

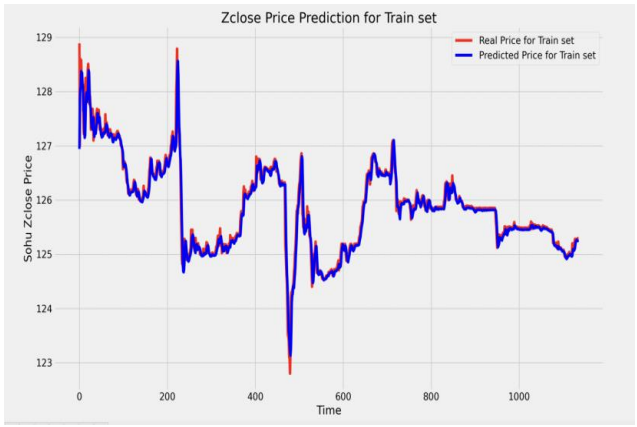

Figure 7. Result 1

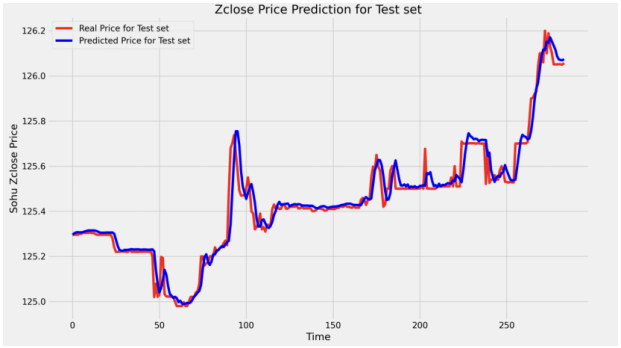

Figure 8. Feature heat map

3.3. Stock Portfolio

By using LSTM to predict 20 stocks, and based on the results, two groups were screened out, each with 4 stocks for combination. Two cases were calculated using Markowitz's portfolio theory and Sharpe ratio, as shown in the figure below.

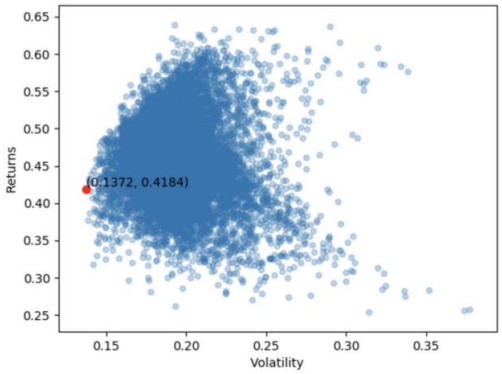

Figure 9. Risk Minimum Graph

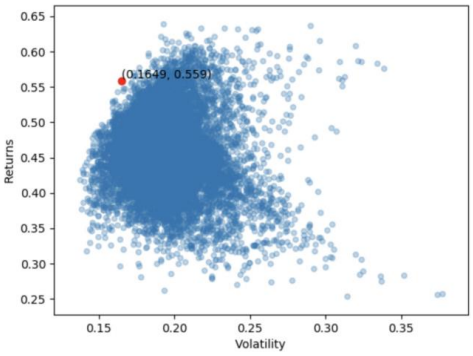

Figure 10. Combination optimal graph

Figure 9 shows the investment portfolio with the least risk. When the volatility is 0.1372, the rate of return can reach 0.4184. Figure 10 shows the optimal investment portfolio. When the volatility is 0.1649, the rate of return can reach 0.559.

4. Conclusion

This paper mainly does two things. One is to select the features of the stock indicators. Too many features are prone to overfitting, so this paper performs feature selection on the indicators of stock historical transaction data. The second is to screen multiple stocks and Portfolio, and calculate the rate of return under different volatility.

It was found that, for some stocks, the prediction effect of the LSTM model in some stocks can meet expectations, but it cannot be extended to all stocks. At the same time, when predicting the stock price, you can use the LSTM model to

predict the stock price, select a few stocks with good effects, and make a reasonable combination.

For investment institutions, the essence of quantitative strategies is to earn stable returns, and the primary goal is to maintain.

A certain income can have a lower drawdown at the same time. This article hopes to give some reference suggestions to quantitative investment institutions and provide a feasible high-frequency strategy research direction through the exploration of high-frequency trading data.

For individual investors, although there are no quantitative channels for individual investors in the market, the pursuit of stable returns is the philosophy of most ordinary investors. At present, the fundamental data required for value investment has more channels for institutions to obtain, but in general, company financial report data and some economic macro data have obvious lags, which often cause certain obstacles to individual investors' investment. Judgment impairment. The research on high-frequency data in this paper also provides a certain reference for future individual investors and researchers, and can also play a certain auxiliary role for every investor.

One is that this paper only incorporates stock data, and there are data from other sources in the financial market. For example, investor sentiment, news events, etc. all have an impact on investors, so this information can be used to improve the forecasting effect of the model. Second, the stock data in this article comes from 2020 to 2021, and the data cycle may be a bit short. Therefore, in the follow-up Longer-period stock data can be used in the research to better train the model.

References

[1] Cao Yanyan. (2023). LSTM model optimization and its application research in stock index forecastin. Master's degree thesis of Dongbei University of Finance and Economics, 55.

[2] Hui Wenwen. (2022). Research on stock price prediction based on LSTM model and multi-source data fusion. Master's thesis of Heilongjiang University, 62.

[3] Li Xuanli. (2023). Prediction of stock price trends based on LSTM-CNN hybrid model based on investor sentiment (Master's thesis of Shanghai International Studies University, 121.

[4] Zhu Wenchao. (2023). Stock price forecasting - LSTM-based financial time-series data modeling and decision-making. Modern Marketing (Second Period) (03), 39-41.

[5] Zhang Ni. (2021). Research on the Application of Stock Price Prediction Based on LSTM Neural Network. Modern Business (16), 116-118.

[6] Zhang Xuntao & Fan Yongsheng. (2023). LSTM model based on XGBOOST feature selection to evaluate stock trend analysis. Computer Knowledge and Technology (09), 91-94+97.

[7] Ding Xin. (2022). The impact of individual investors' investment behavior on their stock market investment performance. National Circulation Economy (36), 121-124.

[8] Yue Songtao. (2023). Prediction of high-frequency stocks based on investor sentiment. Exhibition Economy (08), 108-111.