ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0034

The Nested Two-Layer LSTM Model and Stock Prediction Application

Qinghao Kong, Shuo Guo

Department of Electrical Engineering and Information Technology, Shandong University of Science and Technology, Jinan 250031, China

Abstract

LSTM models are now widely used in various real-life situations, and one of the main uses is for making forecasts. In this paper, the principle and construction process of Nested Two-Layer LSTM model are explained in detail and optimized to predict the vagaries of stock market by constructing LSTM model. Two typical stock data, gold and bitcoin, are selected, and according to the results, the prediction is accurate with goods.

Keywords

Nested Two-Layer LSTM; Stock Forecast; Gold and Bitcoin.

1. Introduction

1.1 Background

Long Short-Term Memory (LSTM) is a kind of optimized recurrent neural network, which overcomes the problem of gradient disappearance and explosion of RNN easily because it introduces the concept of gating; at the same time, the multiple gating inside LSTM neurons have their own functions to form independent storage of memory data, and thus can solve the problem of long-order dependence of data. Compared with the traditional time series model, LSTM has certain advantages in the analysis of nonlinear correlation data due to the construction of internal nonlinear activation function. In this paper, we try to verify the applicability of Nested Two-Layer LSTM for stock data analysis by conducting an empirical analysis of Nested Two-Layer LSTM for stock price prediction. In this paper, the daily trading data of gold and bitcoin for the past 5 years are selected as the sample data. In the experiment of stock price prediction, a Nested Two-Layer LSTM model is used to predict the closing index of the future day and analyze the applicability of the Nested Two-Layer LSTM in stock price prediction.

1.2 Literature Review

The stock market is a very complex system. At present, there are two main prediction methods, one is traditional Machine Learning methods, such as Decision Tree Algorithm, Boost Algorithm, SVM Algorithm, etc.; the other is based on Deep Learning methods, such as Convolutional Neural Network, Long Short-Term Memory, etc.

Mei[1] proposed the combination of SVM and ARIMA to predict the direction of the stock price. First, the ARIMA model was used to predict the target, and then the error was input into the SVM model as data, and the final output was used as the result of predicting the stock price. Kim[2] proposed a Long Short-Term Memory based on feature fusion. The model will extract stock time series data and picture image data, and realize that the fusion prediction model is better than a single model in stock price prediction. Jayanth Balaji et al. [3] explored the effect of various deep learning models on stock prediction, and the experiments proved that the Nested Two-Layer LSTM has the best effect on a specfic stock.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0034

2. Nested Two-Layer LSTM Predicts Future Returns

2.1 The Structure of the Nested Two-Layer LSTM Model

The Long Short-Term Memory Model ,the time series refers to the sequence of the same statistical indicators arranged in the order of their occurrence time. The main purpose of time series analysis is to predict future data based on data that has existed in the past. The LSTM Model is mainly realized through a structure called a gate. The function of the gate structure is to selectively allow information to pass through. The information selection and passing property of the gate structure is realized by using the product of the sigmoid neural layer and the point-by-point structure. The structure of the Nested Two-Layer LSTM Model is shown in the following Figure 1:

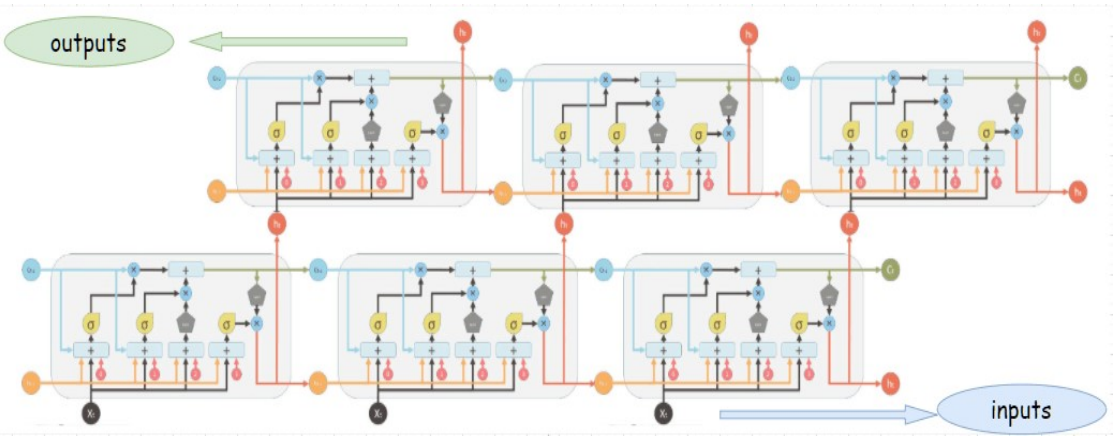

Figure 1. The Structure of Nested Two-Layer LSTM Model

Inputs: We first perform data cleaning and perform information enhancement process on the data. After the information enhancement process, we normalize the data into a uniform data format and use the normalized data as the input vector for this model.

Outputs: After a nested two-layer LSTM model, the output data goes through a linear layer and then a formatted output vector is obtained. We denormalize the output vector, and after processing, we can get the predicted feature vector.

The Nested Two-Layer LSTM Model is composed of two layers of LSTM Model. Each layer of LSTM is composed of three gate structures. These three gate structures are input gate, forget gate and output gate. Consists of a single door structure. The following Figure 2 shows the meaning of a single door structure, that is, the meaning of each module:

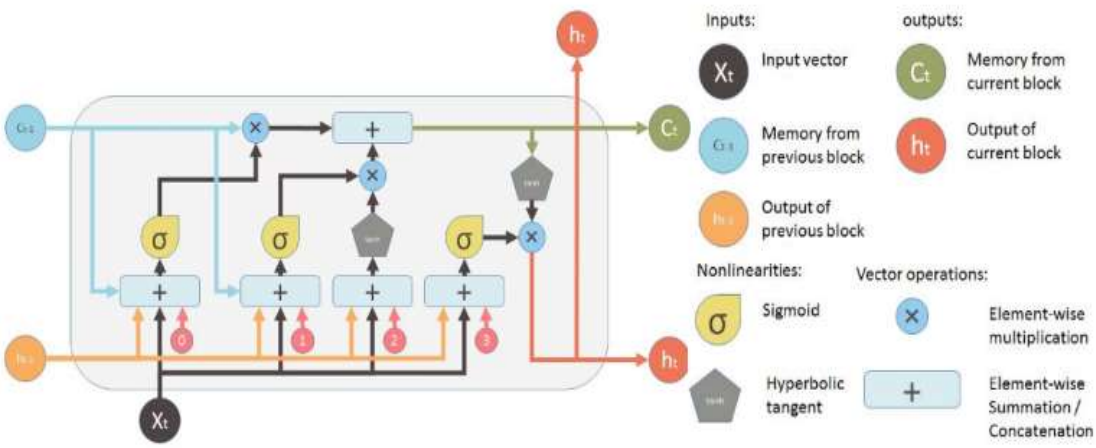

Figure 2. A Single Door Structure of Nested Two-Layer LSTM Model

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0034

The three gate structures of the LSTM Module are Input Gate, Forget Gate and Output Gate. Below we describe the functions of these three gates and their mathematical principles:

1) Forget Gate: Decide what information to discard

The first step in LSTM is to decide what information to discard from the cell state. This decision is made through a Forget Gate. The gate reads and outputs a value between 0 and 1 for each number in the cell state . 1 means "completely reserved", 0 means "completely discarded".

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

2) Input Gate: decides the information to be newly added

The input gate determines how much new information is added to the cell state. Achieving this involves two steps: first, a sigmoid layer called the "input gate layer" determines which information needs to be updated; a tanh layer generates a vector, which is the alternative content to update. In the next step, we combine the two parts to make an update to the cell's state.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

3) Output gate: decide what information to output

Ultimately, we need to determine the value of the output. First, we run a sigmoid layer to determine which part of the cell state will be output, then we tanh the cell state (getting a value between -1 and 1) and compare it to the output of the sigmoid gate Multiplying, it will eventually output the part of the value that we determined to output.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

As for our Nested Two-Layer LSTM, a LSTM unit is overlaid with another LSTM unit. We feed the hidden state of the bottom LSTM into the upper LSTM again and use the hidden state from the upper state as the input for the output module.

2.2 Optimizing the Model

Before we show the analysis results, there are several training details to be stated:

1) Loss function

During the training, we use mean square error (MSE) as the loss function:

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

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0034

Where y is the true target vector and $\hat{y}$ is the predicted feature vector and N is the number of dimension.

2) Normalization

Due to different domains of each dimension in the input vector v_{input} ,a normalization technique is needed. Before we feed the input vector into the LSTM module, we apply the following per-feature normalization operation:

$$x_i = \frac{x_i - \mu_i}{\sigma_i}$$

Where μ_i and σ_i are the mean and standard deviation of the ith feature.

3. Result Analysis

According to the literature review, after a comprehensive comparison of various models and methods for prediction, the Nested Two-Layer LSTM works best on a specific stock. So we use Nested Two-Layer LSTM for stock prediction.

In the process of training and testing the model, we used MSE as the loss function of the model, and the change of the loss function for predicting the price of gold and bitcoin is shown in the figure below. As you can see from the image, whether it is predicting the price of gold or bitcoin , the final loss function is very low, and the loss function when predicting Bitcoin is lower than the loss function when predicting gold price. It can be concluded that the prediction effect of the prediction model is good and the prediction effect of Bitcoin price is slightly better than that of gold price.

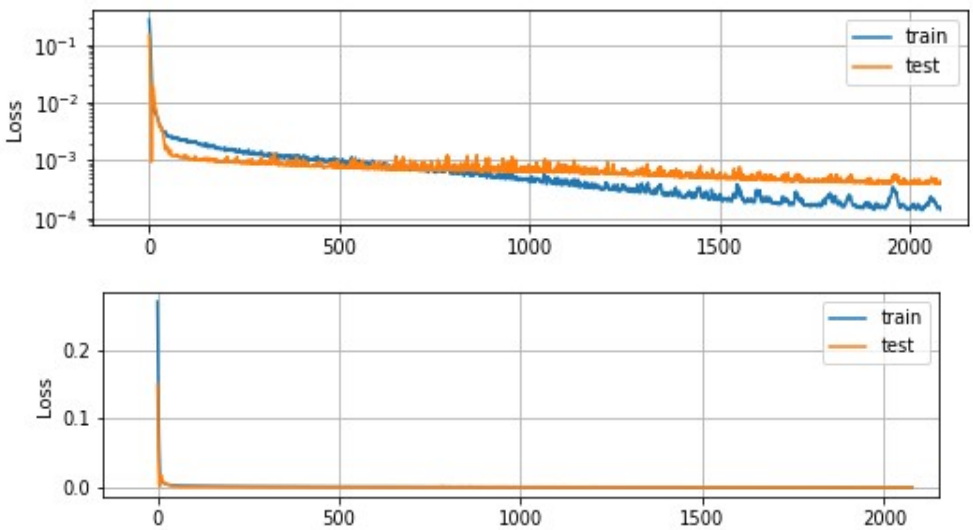

Figure 3. A Loss Function that Predicts the Price of Gold and Bitcoin

After a period of training and testing, we can get the final prediction result. Figure 4 below is the forecast data for gold, and Figure 5 is the forecast data for Bitcoin. The red line is the data we predicted, and the blue line is the real data. The data. It can be seen that the predicted value obtained by using the Nested Two-Layer LSTM is highly similar to the real value, which also shows that the constructed the Nested Two-Layer LSTM has a good prediction effect.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0034

Figure 4. Forecast Data and Real Data for Gold

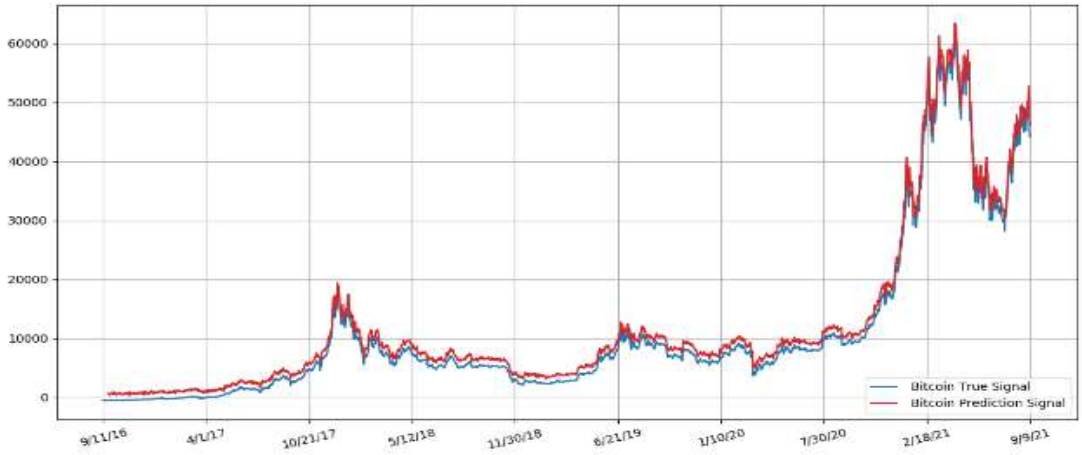

Figure 5. Forecast Data and Real Data for Bitcoin

4. Conclusion

Using the Nested Two-Layer LSTM model to predict the price of gold and Bitcoin, according to the prediction results, it can be seen that the Nested Two-Layer LSTM model has good applicability for the prediction of stocks. Has a wide range of application prospects.

References

[1] Wenjuan Mei.Stock price prediction based on ARIMA-SVM model. Institute of Management Science and Industrial Engineering.Proceedings of 2018 International Conference on Big Data and Artificial Intelligence. Computer Science and Electronic Technology International Society, 2018. 55~61.

[2] Kim Taewook,Kim Ha Young. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data PloS one,2019,14(2).

[3] A.Jayanth Balaji, D.S. Harish Ram, Binoy B.Nair. Applicability of Deep Learning Models for Stock Price Forecasting An Empirical Study on BANKEX Data. Procedia Computer Science,2018,143:947-953.