ISSN: 2688-9323

Research on Price Prediction in Quantitative Trading based on LSTM Neural Network

Yihang Guo, Chenxin Yang

College of Hydraulic and Architectural Engineering, Northwest A&F University, Xianyang, Shaanxi, 712100, China

Abstract

Quantitative trading refers to the use of modern statistical and financial theories and methods. And using computer technology to select multiple "obvious probability events" from a large amount of historical data to develop strategies that can generate excess returns, and using mathematical models to verify, solidify and execute these strategies to achieve consistent, stable excess returns. We have developed a model that uses large amounts of past data to train neural networks and make predictions accordingly. In this paper, we built the LSTM neural network prediction model based on the preprocessed daily price of gold and daily price of bitcoin by setting the window length, rolling window length, and prediction period. We cited the historical data that we found from 2010.1 to 2016.9 to train the model. Based on this model, we predicted the daily price of gold and bitcoin from September 12, 2016, to September 2021 10 for the daily prices of gold and bitcoin.

Keywords

Quantitative Trading; LSTM Neural Networks; Price Forecast.

1. Introduction

Quantitative trading is a scientific and systematic investment solution, which can make sustainable and robust profits through the hedging of financial derivatives and strategies based on big data [1], compared with traditional investment methods, quantitative investment mainly relies on models and large amounts of data to find investment targets and investment strategies.

2. Price Prediction Model based on LSTM Neural Network

Regarding asset (gold, bitcoin) price forecasting, some scholars choose artificial neural networks, some use generalized linear models and random forest models to set the asset price forecasting problem as a binomial classification problem [2], while we use an improved LSTM recurrent neural network for price forecasting. The LSTM is compared with typical time series forecasting models (gray forecasting models, ARIMA models, BP neural networks), the LSTM model has an overall better fitting and forecasting performance, and the LSTM model has an overall higher fitting and forecasting accuracy compared to other types of recurrent neural networks (RNN and GRU). Then the improved LSTM neural network is based on the recurrent neural network RNN, which is based on the principle of passing input data to the network across time steps so that the information is retained, thus solving the problem of "forgetting" previous inputs. The basic RNN performs well in modeling short sequences but suffers from the problem of gradient disappearance in long sequences. The long short-term memory model (LSTM) developed by Hochreiter and Schmidhuber [3] can retain some of the information, making the neural network long-term dependent and thus improving the accuracy of the prediction model. An LSTM has good performance in problems such as image recognition,

ISSN: 2688-9323

speech recognition, and time series problems [4]. Therefore, based on the above considerations, we choose to use LSTM neural network for price prediction.

Because the LSTM neural network training process requires a large training set, we searched for historical price data for gold and bitcoin and found data from January 1, 2010, to September 11, 2016, a total of 2740 days, and daily bitcoin settlement prices from January 1, 2014, to September 11, 2016, as the training set. Then the period from September 11, 2019, to September 10, 2021, given in the question was selected as the test set for validation.

Long-Short Term Memory Network (LSTM), a recurrent neural network after improvement, can regulate the duration of time, has memory nodes storing historical information, and can be trained on sample data by a back-related algorithm. It fundamentally improves the problem of gradient disappearance and the emergence of long-standing sex problems in RNN neural networks.

LSTM neural network has the feature of directed recurrence, which can mine the time series relationship existing in the prediction problem. Its advantage is that it can fit well the nonlinear data present in the time series and the model can store the information in the time series for a long time, so it can extract information with long time intervals and relatively long extensions. Therefore, LSTM has a high accuracy for trend prediction and long-term prediction of time series [5].

The structure of a typical RNN network is as follows:

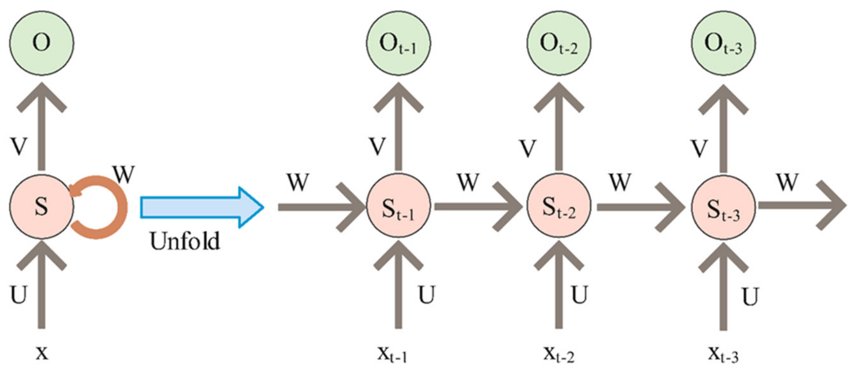

Figure 1. Typical network structure of RNN

The right side is the structure that is produced for easy understanding and memory when calculating. Briefly, x is the input layer, O is the output layer, S is the implied layer, and t refers to the calculation of the tenth time; V, W, and U are the weights, where the state of the implied layer when calculating the t time is:

$$S_k = f(U * X_t + W * S_{t-1}) \tag{1}$$

Thus, the purpose of linking the current input result to the previous calculation is achieved.

Since the RNN model needs to link the computation of the current implicit state to the previous n computations if it needs to achieve long-term memory:

$$S_k = f(U * X_t + W * S_{t-1} + W_2 * S_{t-2} + \cdots + W_n * S_{t-1}) \tag{2}$$

In that case, the computational effort would grow exponentially, resulting in a significant increase in model training time, so RNN models are generally used directly for long-term memory computation.

ISSN: 2688-9323

LSTM is a variant of RNN with the main feature of adding the valve nodes of each layer outside the RNN structure. There are three types of valves: forget gate, input gate, an output gate, which can be opened or closed and are used to determine whether the output of the memory state of the model network (previous state of the network) in that layer reaches a threshold value and is thus added to the computation of the current layer. The valve nodes use the sigmoid function to compute the memory state of the network as input; if the output reaches the threshold, the valve output is multiplied by the current layer's computation as input to the next layer (multiplication means element-by-element multiplication in the matrix); if the threshold is not reached, the output is forgotten. The weights of each layer, including the valve nodes, are updated during each model backpropagation training.

The specific LSTM's judgment calculation process is shown in Figure 2 below:

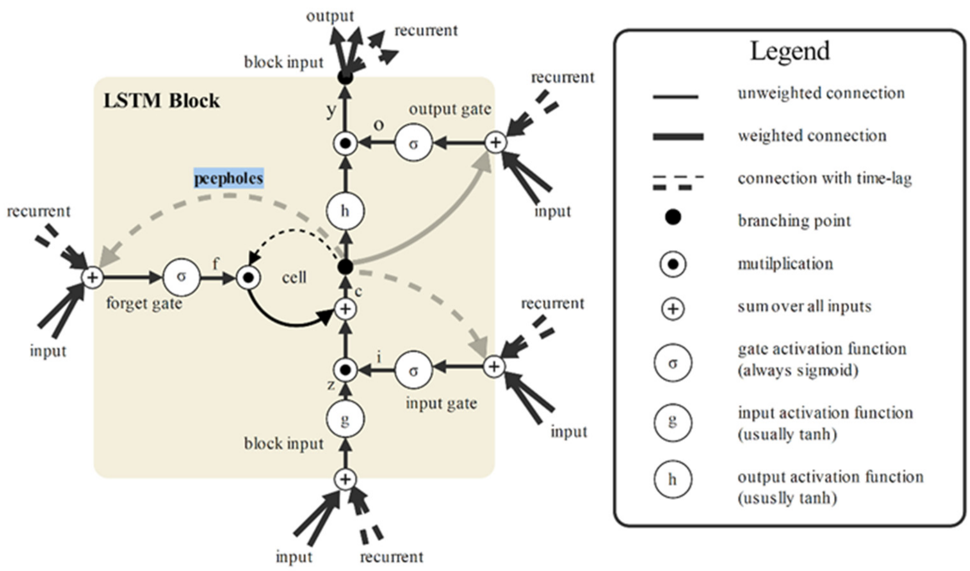

Figure 2. LSTM calculation judgment process

The kernel of the LSTM idea is based on the hidden layer in the RNN model, with an additional hidden layer containing only a single state of h that preserves the long-term state, called the unit state c, and expands the two states according to the time dimension.

At time t, there are three inputs to the LSTM: the input value x_t of the network at the current moment, the output value h_{t-1} of the LSTM at the previous moment, and the cell state C_{t-1} at the previous moment. there are two outputs of the LSTM: the output value h_t of the LSTM at the current moment, and the cell state C_t at the current moment.

The control of the long-term state is influenced by three control switches, which are the three important gates in the LSTM neural network:

a) Forget gate: It determines how much of the cell state C_{t-1} from the previous moment is preserved to the current moment C_t.

b) Input gate: It determines how much of the input x_t of the network at the current moment is saved to the cell state C_t.

c) Output gate: Control unit state C_t How many outputs to the LSTM with the current output value h_t.

The forge gate, input gate, and output gate are the three important internal structures that exist in the memory unit of LSTM neural networks. the LSTM model can train cyclically and optimize its ground weights adaptively on a certain basis, thus being able to solve the problem of gradient disappearance and explosion to some extent.

ISSN: 2688-9323

LSTM recurrent neural networks can model problems with multiple input variables almost perfectly. This is a great benefit in time series forecasting, where classical linear methods are difficult to adapt to multivariate or multi-input forecasting problems.

Compared with the static method, the dynamic rolling method adds new data to the training data with each model training, and each time step of prediction is trained by the information of the period before that step to ensure that the latest information is used for prediction.

Regarding the selection of the training step length, usually the less the training set time step prediction can prevent the phenomenon of overfitting, but there may also be a state that cannot fully learn the data structure, so we tried to observe the prediction effect of the model by constantly changing the time step length, and finally selected the method of 50 steps to predict 3 steps, setting the window length to 50 days, the rolling window length to 1 day, and the prediction period length is 3 days. This indicates that we will predict the price for the next 3 days by looking at the previous 50 days of data based on the point in time of the day. The neural network training process is shown in Figure 3.

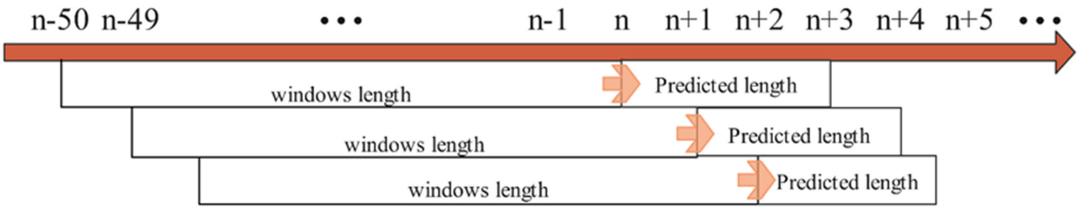

Figure 3. Dynamic scrolling training chart

3. Model Solving

The training set is trained using an LSTM neural network model with the following training parameter settings.

a) Time step: It means that the recurrent neural network considers each input data to be associated with the previous number of successive inputs, it means that the recurrent neural network considers each input data to be associated with the previous number of successive inputs, the training input data is the price, the initial input time step is set to 3, that is, the length of the recurrent unit, each time to predict only the next three days of data, in turn, all the data are predicted in a loop.

b) Input shape: Forecasting by asset prices only, the input data dimension is 1 dimension.

c) Optimizer: The purpose of the optimizer is to minimize the loss function. Among all neural network optimizers, the initial setting is the most commonly used Adam optimizer, which has the advantage that after bias correction, the learning rate of each iteration has a definite range, making the parameters relatively smooth.

d) LSTM units: The LSTM neural network also needs to define the number of neurons for each neural layer, and the input data shape should be taken into account when defining it, and make the number of neurons in the LSTM layer slightly larger than the input data shape to ensure that the model can operate properly, but it should not be set too large in comparison, otherwise it will have an impact on the performance of the model. Considering the above problem, the number of neurons in the LSTM layer is initially defined as 20 because the number of factors we input is 1.

Based on the model we have trained, we can predict the next three days based on the data so far, so that we can adjust the portfolio based on the results of the prediction. Figure 4 below shows the prediction of a certain stage in the prediction model, i.e., the price of the next three days using the data of the first 50 days.

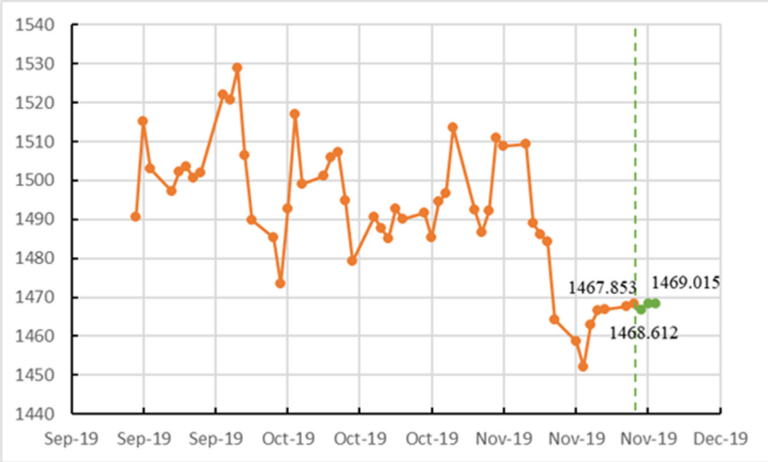

Figure 4. Sub-stage prediction diagram

Constantly predicting price movements for the next three days based on our data to date. To verify the accuracy of the prediction model, we can use the prediction results of day N+1 to compare with the original price trend as shown in Figures 5 and 6, and then perform error analysis.

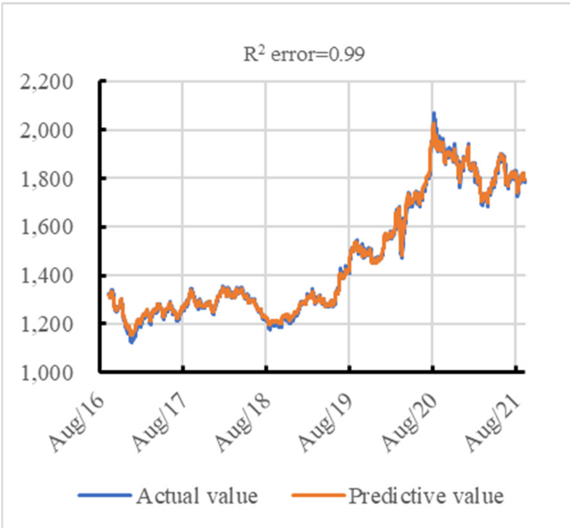

Figure 5. Gold price forecast trend

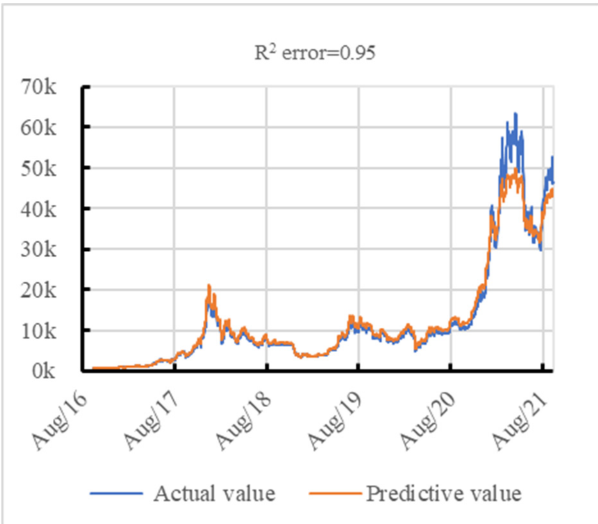

Figure 6. Bitcoin price forecast trend

ISSN: 2688-9323

Here we use R^2 (average error) to evaluate the error between the prediction results and the true value, and the results of the calculated error indicator are 0.99 and 0.95 are greater than 0.8, which can be considered as good prediction results, so the prediction results can be used in the next model, and then deduce the maximum income of daily trading and the trading decision of assets.

4. Conclusion

This paper develops a price prediction model to assist quantitative trading based on LSTM neural network. Our LSTM neural network is the best method for analyzing time series data today. In the specific prediction model, the neural network is trained with gold and bitcoin price data until 2019, so that you can get the bitcoin and gold prices on that day, and then use our model to predict the prices for the next three days for your reference, and then adjust your portfolio. We also measure the deviation of the predicted value from the true value by solving for the R^2 index, and calculate error indicators of 0.99 and 0.95 for gold and bitcoin, respectively, which are both greater than 0.8, proving that the prediction models are accurate. LSTM recurrent neural network solves the problem of gradient disappearance and gradient explosion that exists during the training of long sequences. In LSTM recurrent neural network, the nodes in the implicit layer are interconnected, and the features and patterns learned in the previous moment are transferred to the next moment, which makes the network have the function of memory; meanwhile, LSTM recurrent neural network has the forgetting mechanism, which better simulates the forgetting mechanism of the human brain.

References

[1] Liu Lijun, Liang Guopeng, Wang Haitao. Research on Quantitative Trading Strategy System Model [J]. Shanghai Business, 2021(09):158-159.

[2] Ta, Liu, Tadesse. Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading[J]. Applied Sciences, 2020, 10(2):437.

[3] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[4] Duan Y, Lv Y, Wang F Y. Travel time prediction with LSTM neural network[C]// IEEE International Conference on Intelligent Transportation Systems. IEEE, 2016.

[5] Venter G, Jaroslaw S S. Particle Swarm Optimization[J]. AIAA Journal, 2003, 41(8):129-132.