Research on Price Forecasting of Bitcoin and Gold Quantitative Transactions Based on LSTM

Jingong Chen*, Jingwen Liu

College of Mathematics and Statistics, Huajin Campus, Anhui Normal University, Anhui, 241000

*Corresponding author: chenjingong0301@163.com

Keywords: Artificial Neural Network; Reading difficulty factor; Hierarchical Clustering.

Abstract: Common market investment assets include gold and bitcoin. Bitcoin is an encrypted currency with a regional chain as a basement technology, which market traders increasingly recognize, but high-profile means high risk. Gold as a general equivalent is a frequent visitor in asset configuration. We have established a long short-term memory neural network model, quantifying the price of Bitcoin and gold. Through 2016, Bitcoin and Gold's transaction prices were trained on the Internet, and then the price forecast for the next three days from September 11, 2021, and the prediction results were tested, and the errors were 0.92 and 0.99. It can be seen that the model is well predicted.

1. Introduction

In the actual transaction process, investors can only give the best daily trading strategy according to the price data of gold and bitcoin, [1] then they need to predict and assess investment projects in the next period and plan a transaction plan. Market traders buy and sell volatile assets in order to maximize profits. [2] The level of traders has a considerable impact on each transaction. With the arrival of the big data information era, a large amount of digital information makes people dazzled and how to find the most effective digital information from these complex macro figures.[3] Traders rely on a wealth of daily trading data and policy guidance to predict the price of investment assets and profit from it.[4]

2. Price Prediction Model Based on LSTM

The price prediction of financial assets has always been a hot research field at home and abroad. Among the prediction methods of financial time series such as stock price, there are generally two kinds: one is the financial time series prediction model based on statistical methods, and the other is the financial time series prediction model based on machine learning algorithm [5]. In the category of statistical method, the most commonly used is the difference autoregressive integrated moving average model (ARIMA)[6] and generalized autoregressive conditional heteroscedasticity (GARCH)[7] model. However, the use of these methods is the premise of the variables of financial time series by linear assumption, despite all kinds of single forecasting methods to the specific time, which can show good prediction performance. However, real-life financial time series are relatively complex. Many distractions and economic data samples tend to have very big random fluctuations, making the financial time series not meet the linear. Stationary assumptions, facing such data used to create a single model method to consider all factors of fluctuations is very unrealistic. The effect of traditional statistical methods in the financial market is not ideal. At the same time, the machine learning algorithm has been widely used in financial time series and proved to have a good prediction effect in practice. Among them, RNN has the problem of long dependence in practical application. As one of the variants of recurrent neural network RNN, LSTM effectively solves the prolonged dependence of RNN and has good performance in such problems as image recognition, speech recognition, and time series. Therefore, based on the above considerations, we used LSTM neural network for prediction.[8]

There are some missing data in the experimental data. The linear interpolation method is used to complete the training set of the experimental data. The Daily gold settlement price data in the United

States from January 02, 2008, to May 16, 2018, are used in a total of 2290 days. Moreover, the daily bitcoin settlement price from April 28, 2013, to August 14, 2018, was then verified using the data given in the question from September 11, 2017, to September 10, 2021. In order to eliminate the influence of different dimensional and dimensional units between prices on the data results, we standardize the data to solve the comparability between data. For the price of gold and bitcoin, we use today's price/yesterday's price instead of today's price for the subsequent forecast, which makes the data more impressive.

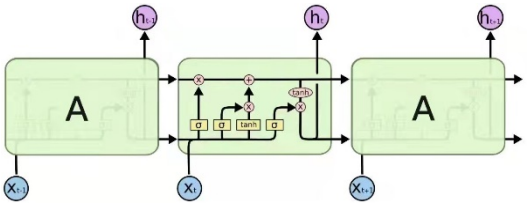

Figure 1. LSTM flowchart

Long and short-term memory network (LSTM) is an improved recurrent neural network.[9] Its unique feature is that it can realize long-term memory of information by introducing Gates structure and effectively filter LSTM through three gate structures: o_t、x_t、f_t, the purpose is to achieve selective screening of information at every moment.

In each time step of LSTM, there is an LSTM selective memory function, which enables LSTM to freely select the contents to be remembered in each time step. f_t gate determines whether or not we need to remember information about the previous state, that is, how much of the current state of memory comes from the last memory (output a number between 0 and 1, where 1 means we keep the information we remember and 0 means we do not remember the previous information).

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

x_t and tanh function are controlling the new data and adding the tanh function produces a new candidate, x_t for each term in $\tilde{C}_t$ produces a value within [0,1] that controls how much new information is added.

$$C_t = f_t * C_t + i_t * \tilde{C}_t \tag{2}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1,} x_t\right] + b_i\right) \tag{3}$$

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1,} x_t] + b_c) \tag{4}$$

O_t determines how much of the current input value x_t is saved to the cell state. LSTM uses the output gate to control how much of the cell state $C_t$is output to the current output value h_t of LSTM.

$$o_t = \sigma\left(W_o\left[h_{t-1,} x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t * tanh(C_t) \tag{6}$$

Thus, we get the LSTM neural network model.

$$\begin{cases} i_t = \sigma(\overline{i_t}) = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t = \sigma(\overline{f_t}) = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ g_t = tanh(\widetilde{g_t}) = tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \\ o_t = \sigma(\widetilde{o_t}) = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ c_t = c_{t-1} \odot f_t + g_t \odot i_t \\ m_t = tanh(c_t) \\ h_t = o_t \odot m_t \end{cases} \tag{7}$$

The basic computing unit of the neural network is the neuron, usually called node or unit. It receives input from other nodes or external sources and computes the output. Each input has an associated

weight (weight, w), which is assigned according to its importance relative to other inputs. The output at one point is the superposition of many previous inputs multiplied by their attenuation coefficients to form the output at one point, and then the output points at different times are put together to form a function, which is the convolution. Sometimes, it is necessary to avoid information loss by filling, and sometimes, it is necessary to compress part of the information by setting step size during convolution. Therefore, step in convolution is another essential operation for constructing convolutional neural networks.

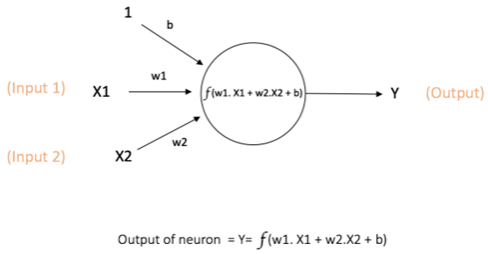

Figure 2. Neural network

The data used in LSTM neural network should be to extract time sequence features of time series data. We set the window length as 30 days, the rolling window length as 1 day, and the prediction period 3 days. It can be seen from the chart that we will forecast the stock price 3 days later based on the current time of each day by looking at the data for the previous 30 days. These settings also apply to CNN time series data's window length and prediction length.[10]

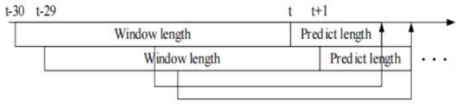

Figure 3. LSTM time series

The gradient descent algorithm is mainly used to optimize the value of a single parameter. In contrast, the backpropagation algorithm provides an efficient way to use the gradient descent algorithm on all parameters to minimize the loss function of the neural network model on the training data. Compared with the RNN model, LSTM can alleviate the gradient disappearance. In RNN, for any time at t, there is the following relationship:

$$\frac{\partial L_t}{\partial W_t} = \frac{\partial L_t}{\partial W_x{}^{(t)}} + \cdots + \frac{\partial L_t}{\partial W_t{}^{(1)}} = \frac{\partial L_t}{\partial Q_t}\frac{\partial Q_t}{\partial S_t}\frac{\partial S_t}{\partial W_x{}^{(t)}} + \sum_{k=1}^{t-1}\frac{\partial L_t}{\partial Q_t}\frac{\partial Q_t}{\partial S_t}\left(\prod_{j=k+1}^{t}\frac{\partial S_j}{\partial S_{j-1}}\right)\frac{\partial S_t}{\partial W_x{}^{(k)}} \tag{8}$$

$$\frac{\partial S_j}{\partial S_{j-1}} = tan(\theta_j)W_s \tag{9}$$

$$S_1 = tanh(W_X X_1 + W_s S_0 + b_1) \tag{10}$$

$$S_2 = tanh(W_X X_2 + W_s S_1 + b_1) \tag{11}$$

$$S_3 = tanh(W_X X_3 + W_s S_2 + b_1) \tag{12}$$

$$O_1 = W_o S_1 + b_2 \tag{13}$$

$$O_2 = W_o S_2 + b_2 \tag{14}$$

$$O_3 = W_o S_3 + b_2 \tag{15}$$

In LSTM, the value can be flexibly controlled by adjustingW_{hf}、W_{hi}、W_{hg}, we can control what $\frac{\partial C_t}{\partial C_{t\text{-}1}}$ is. When long-term memory continues from time n to time m, the value of $\prod_{t=n}^{m}\frac{\partial C_t}{\partial C_{t-1}} \approx 1 * 1 * 1 * \ldots * 1$. Thus the gradient disappearance is greatly alleviated.

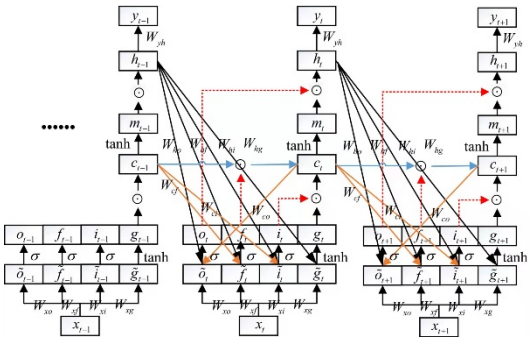

Figure 4. Causes of gradient disappearance

LSTM neural network model is used to train the training set, and the parameters are set as follows:

1. Time step: The parameter distinguishes the LSTM neural network model from other models. LSTM requires the input data to be priced during training data input. The initial input sequence step size is set to 3, that is, the length of the loop unit. Each time, only the next three days of data are predicted, and then all the data are used to predict the next three days.

2. LSTM units: Similar to other neural network models, LSTM neural network also needs to define the number of neurons at each neural layer. The input data shape should be taken into account during the definition, and the number of neurons in the LSTM layer should be slightly larger than the input data shape to ensure the normal operation of the model. However, it should not be set too large in comparison. Otherwise, the performance of the model will be affected. Considering the above problems, since the number of factors we input is 1, the number of neurons at the LSTM layer is initially defined as 20.

3. Optimizer: The purpose of the optimizer is to minimize the loss function. Adam is the most commonly used optimizer among all neural network optimizers, so the initial optimizer is Adam.

4. Input shape: We only use gold price/bitcoin prices to invest in prediction. Therefore, the input data dimension is directly 1-dimensional.

3. Model Solving

Based on our model, we can predict the next three days based on the trading data we have so that we can adjust the portfolio according to our results in time to achieve the purpose of maximum return. Depending on the trained model, we can predict the next three days from today's data to make changes to the portfolio strategy.

From Figure 5, we get bitcoin price forecasts for the next three days.

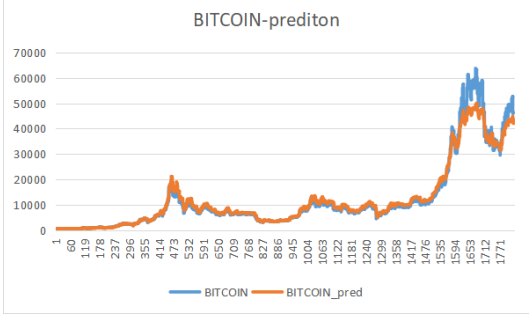

Figure 5. BITCOIN-prediction

From Figure 6, we get the gold price forecast for the next three days.

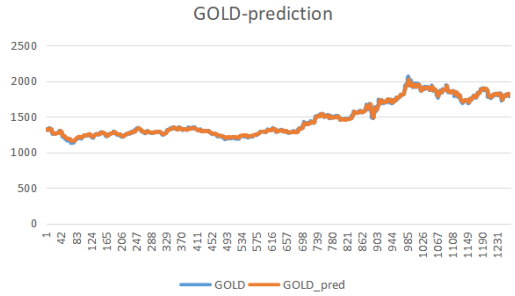

Figure 6. GOLD-prediction

In order to verify the accuracy of the prediction model, we used the prediction results of day N+1 to calculate the error.

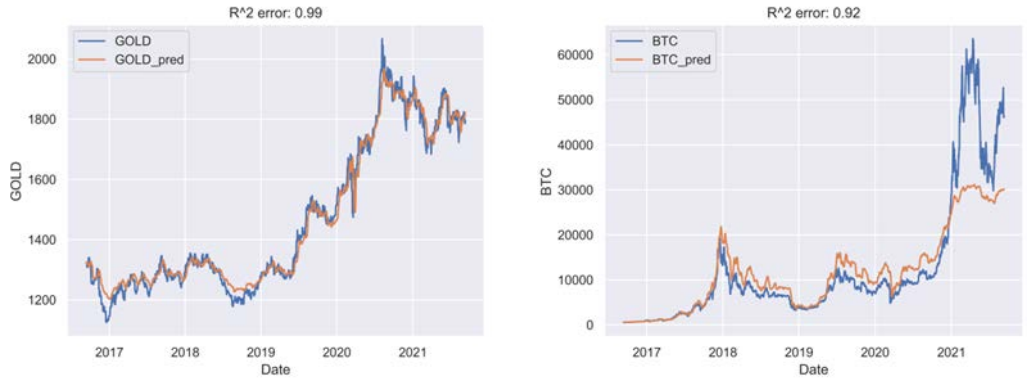

Figure 7. Prediction error

According to the above Figure 7, the prediction error of gold is roughly 0.99, and that of bitcoin is 0.92, so the prediction effect is good (prediction error = error quantity/price of the day).

4. Conclusion

The model established in this paper is used to study gold and Bitcoin to consider the most significant return for rational investment. LSTM is like RNN to better handle the task of time series than CNN. At the same time, LSTM solves the long-term dependence of RNN and alleviates the "gradient disappearance" problem brought by RNN in reverse communication during training. The sequence data of the appropriate length is greatly advantageous. This model has a particular reference value for the actual stock, futures, and financial investments.

References

[1] Zhang Siyu. Taking Bitcoin as an example to explore virtual currency investment [J]. Modern Business, 2022 (03): 101-103.

[2] Ming Lei, Wu Yifan, Xiong, Message. Bitcoin price bubble inspection, evolution mechanism and risk prevention [J]. Economic Review, 2022 (01): 96-113.

[3] Du Linfeng. Is the Bitcoin market still impact my country's exchange rate and stock market? Empirical Study Based on TVP-VAR Model [J]. North China Finance, 2021 (11): 20-31.

[4] Xu Wei, Gorn Yang.Agferability [J]. Today's Wealth (China Intellectual Property), 2021 (11): 103-107.

[5] Zhang Yuan, Leisure, Qi Charhui, Zhang Zhihong.The related transaction behavior detection method based on Bi-LSTM [J / OL]. Computer Science: 1-10 [2022-02-26].

[6] Song Yi Meng, Chen Xingchi. Construction of Prediction Model of Square Index of Photovoltaic Construction Integrated Industry - An Empirical Analysis Based on ARIMA Model [J]. Technology and Industry, 2021,21 (12): 238-243.

[7] Wu Lingrui, Jiang Zi FAN, Wang Fen.Xu to the New Crown "Black Swan" Wave Trend Discussion - An Empirical Study Based on GARCH Model [J] .Journal of Hubei University of Economics (Humanities Social Science Edition), 2022,19 (02): 51-55.

[8] Chen Jia. Research on the application of RNN neural network in stock index prediction [D]. Tianjin University of Science and Technology, 2019.

[9] Du Kangji. Improved circular neural network method and its application [D]. Northeast Electric Power University, 2021.

[10] Fang Yiqiu, Lu Zhuang, Ge Junwei. Shares prediction of the LSTM-CNN model of the joint RMSE loss [J / OL]. Computer Engineering and Application: 1-10 [2022-02-26].