Bitcoin Price Prediction Based on

CNN-LSTM Hybrid Neural Network Model

Team member:	Li Yan	A0215664A
	Han Yuyao	A0215635H
	Liu Xiaojuan	A0215644H
	Li Xin	A0197955E
	Zhu Jiatong	A0215594Y

Content

1. Introduction

In recent years, electronic payments are mainly based on traditional electronic payment tools, while digital currency relies on virtual currency which is on the basis of the blockchain technology. Compared with traditional currency and electronic transaction, it uses decentralized point-to-point network, does not need a third-party payment platform, and has faster transaction speed and lower cost. In addition, due to the use of encryption algorithm and automatic authentication mechanism, the transaction is more secure and transparent, and is not easy to be cracked or forged. Therefore, since the birth of digital currency, it has been widely concerned.

Figure 1: the price changes of Bitcoin

In particular, Bitcoin is the world's first distributed super sovereign digital currency. It was proposed and built by Japanese programmer Satoshi Nakamoto in 2009. It relies on the electronic payment system based on encryption technology and P2P (peer-to-peer) technology. In April 2010, Bitcoin was publicly traded for the first time at the price of US $0.03. In January 2013, the price did not exceed US $15, and the peak price in December of the same year was close to US $1000. In addition, the price reached a new high of $19000 in December 2017, but then dropped sharply to $6700 in February 2018, even to $3200 in December of the same year. At the end of 2018, after the price of Bitcoin went through a low ebb, the price rose all the way and was in a state of rapid growth, exceeding $61000 in March 2021.[1] Recently, Bitcoin has its best first quarter since 2013, spending more than 30 days above $50,000 for the first time in its history. It can be seen that the price of Bitcoin has been rising all the way, accompanied by shocks and large fluctuations. Bitcoin also trades 24 hours a day at a price dominated by the non-government. There is no price limit. Therefore, the price trend is obviously characterized by sharp rise and fall.
Because the price of Bitcoin fluctuates greatly, many investors use it to speculate, so Bitcoin has also been widely concerned by the public. However, just because of the

large fluctuation of Bitcoin price, it is quite difficult to predict the price of Bitcoin. Our project is to solve this problem. We study whether we can predict the price of Bitcoin based on internal information (such as the historical price of Bitcoin) and external information (such as market factors). The data used for forecasting is not only limited to Bitcoin's historical trading information, but also includes technical indicators, macroeconomic variables and investors' attention to Bitcoin. At the same time, artificial intelligence technology is introduced into Bitcoin price forecasting. In our project, CNN is used to extract features that have a significant impact on the Bitcoin price in the data set, and then LSTM is used to predict the price. A price forecasting model based on CNN-LSTM hybrid neural network is proposed. The current research focuses on the accuracy and direction of Bitcoin price forecast.

Based on this, the structure of this report is as follows: Section 2 reviews the relevant literature. The third section introduces the research methods used. Section 4 introduces and analyzes the results. Finally, section 5 is the conclusion of this project.

2. Related Literature

At present, many domestic and international scholars have conducted a great deal of in-depth discussions on the price of digital cryptocurrencies especially Bitcoin.

Due to the dramatic changes in the price of Bitcoin, many scholars have studied the factors affecting the price of Bitcoin. Jermain Kaminski (2014) studied the correlation between investor's Twitter sentiment and Bitcoin price and trading volume. Cheoljun Eom and Taisei Kaizoji (2019)[2] claim investor sentiment can help explain changes in Bitcoin volatility for future periods significantly. Ladislav Kristoufek (2013) studied Bitcoin price , finding that there is a relationship between Bitcoin price and search engine. Additionally, the price of Bitcoin would be affected by investors' speculative behavior, and there is a bubble in it. S.Vassiliadis (2017)[3] points out that there is a strong correlation between the price of Bitcoin and trading volume and transaction cost, and there is a certain relationship with gold, crude oil and stock market index. Pavel Ciaia (2016)[4] also draws the conclusion that there is a short-term guidance lag relationship between Bitcoin price and macroeconomic variables in the empirical test.Huang JZ and Huang W (2018)[5] uses 124 technical indicators based on the historical price of Bitcoin to build a return prediction model, showing that the combination of big data and technical analysis can help predict the return of Bitcoin.

In terms of forecasting, machine learning has made breakthrough progress in recent years. More and more scholars use deep learning technology to predict Bitcoin price.Liang qiu and Wang Fanbin(2015)[6] introduces the wavelet analysis to the prediction of the trend of Bitcoin price over a quarter, using the time series of Bitcoin price. Li Jing (2016)[7] collects Bitcoin transaction data from January 2009 to March 2016, establishing a Bitcoin market forecasting model which uses the data of the previous day, week and month of Bitcoin market on the BP neural network. The author suggests that the longer the prediction period, the larger the prediction error, and the transaction volume is more difficult to predict compared with the price. Dennys C.A.Mallquia (2018)[8] employs Artificial Neural Networks (ANN), SVM, and

Ensemble algorithms (based on Recurrent Neural Networks and K-Means clustering methods) to predict the direction of Bitcoin price, and analyzes the behavior of ANN and SVM for the maximum, minimum and closing prices predictions. The study concludes that the combination of RNN and a Tree classifier can better predict the direction of Bitcoin price, meanwhile SVM algorithm obtains a more precise prediction than ANN in forecasting the Bitcoin price.
Therefore, most of the current studies show that the price of Bitcoin is affected by a variety of factors, showing the characteristics of high volatility, which is difficult to predict in traditional methods. However, machine learning can train and learn complex nonlinear data, which is suitable for Bitcoin price forecasting. At the same time, considering that the price of Bitcoin has changed greatly, our project mainly forecasts the short-term price of Bitcoin.

3. Proposed Methodology

3.1 Data Collection and Pre-processing

Sources of information can be divided into internal information (different parameters of Bitcoin) and external information (macroeconomic factors and investor attention).
Among them, the internal information includes the opening price, the highest price, the lowest price, the closing price, the trading volume, and the transaction amount of Bitcoin. They are derived from Kraken Bitcoin exchange trading data provided by Quandl (https://www.quandl.com). Huang JZ(2018)[5] declares that technical indicators can be used to predict the price of Bitcoin, so our project chooses three technical indicators: RSI,MFI and OBV.
As a new investment tool, the price of Bitcoin is considered to be related to macroeconomic variables. Instead of gold as a hedging tool, when the macro-economy is well, people have extra money in their hands to invest, and the increase in Bitcoin buyers increases the price. Therefore, we selected the following indicators: crude oil futures price, gold price, S & P500 index, New York Stock Exchange Index, NASDAQ index, federal funds rate and RMB dollar exchange rate to reflect the running state of macro-economy. All the data come from the wind database.
Da Z and Engelberg J (2011)[9] points out that search activities reflect investors' concerns. When investors search for virtual currency in search engines, we think they have paid attention to virtual currency. The search index is based on the number of keyword searches in the search engine. The Internet Finance Laboratory of Tsinghua University's Wudaokou School of Finance, together with Huobi and Sina Technology, jointly released the Global Bitcoin Research Report 2014—2016, which was China's first independently released global Bitcoin development report. According to the report, China accounts for 80 percent of the world's Bitcoin transactions, followed by Europe and the United States. And the main search engine in China is Baidu. Therefore, we use Baidu Index to accurately quantify investors' attention to Bitcoin [10].The data comes from the Baidu website (https://index.baidu.com).

According to[7], the long prediction period may lead to enormous prediction error. So the forecast period used in this paper is 3 days, namely, the characteristic parameter data of the previous 3 days is used to predict the closing price of Bitcoin on the 4th day. Each set of samples in the constructed data set has a 51-dimensional feature that contains all the features of the previous 3 trading days. Moreover, we adopted a rolling prediction method with a time interval of six months. In each time interval, the data from the first five months is used as a training set to predict the closing price of the sixth month of Bitcoin. Table 1 shows the attributes which compose the data set.
And the data we used ranges from October 30th, 2017 to April 1st, 2019. We didn't use the data of 2020, because we believe that the market in 2020 will be affected by the COVID-19 pandemic, and the market fluctuated greatly. At this time, some variables we use to measure the macro-economy, such as the stock index, cannot reflect the market well, and the market is invalid. So we discarded this part of the data.
In order to prevent the influence of different dimensions of the original variables on the prediction accuracy, the original data is standardized. The calculation formula is:

$$y_{i,j} = \frac{x_{i,j} - \overline{x}_j}{s_j}$$

where $y_{i,j}$ is the data after standardization, $\overline{x}_j$ is the mean and s_j is the standard deviation of each dimension component.

Table 1: List of input attributes

transaction information	technical indicators	macroeconomic variables	investor attention
lowest price	RSI	gold price	Baidu index
closing price	MFI	exchange rate	
highest price	OBV	NYSE index	
opening price		NASDAQ index	
trading volume		S&P500 index	
transaction amount		Federal Funds Rate	
		crude oil futures price	

3.2 CNN-LSTM Hybrid Neural Network

CNN Neural Network

CNN model is one of the most typical and widely used artificial neural networks nowadays. It mimics the perception of local information by biological visual cells. Local connection and layer by layer calculation are used to extract data features. Finally, the global information is synthesized through full connection. Its basic structure

includes convolution layer, pooling layer and complete connection layer.
The convolution layer uses convolution instead of matrix multiplication, and each convolution kernel can extract the features of the input data. Convolution operation adopts weight distribution method, which effectively reduces the number of parameters, reduces the complexity of neural network training, and improves the training speed. Pooling layer can reduce the size of input data and data volume. Common merging methods include mean merging, maximum merging and so on. The formula is as follows:

$$y_j^l = f(\sum_{i \in M_j} x_i^{l-1} k_j + b_j^l)$$

$$x_j^l = g(\beta_j^l pooling(x_j^{l-1}) + b_j^l)$$

Where, f and g is the activation function, M is the convolution kernel, b_j^l is the bias, k_j is the weight matrix of the convolution kernel, β_j^l is the coefficient of the channel corresponding to the pooling layer and pooling() is the pooling function.

LSTM Neural Network
Long Short-Term Memory (LSTM)[11] is an improved RNN model, which can effectively solve the problem of gradient disappearance and gradient explosion in the RNN model. It is suitable for processing long-term sequence data and solving long-term dependence. Its basic unit is the memory module, containing the memory unit and three gates controlling the memory unit, namely Input Gate, Output Gate and Forget Gate. The gate is the structure that determines the selective passage of information. If the output value of the sigmoid function is 0, it is discarded completely, while if it is 1, it passes completely. Figure 2 shows the basic unit of LSTM Neural Network.

(1) Forget Gate: The purpose of Forget Gate is to determine what information will be discarded. Reading in the output of the previous layer h_{t-1} with the current input x_t, the gate outputs f_t and assigns the current cell C_{t-1}, the calculation formula of f_t is as follows:

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

Where σ presents sigmond function

(2) Input Gate: The role is to update based on existing information. Firstly, run the sigmoid function to get i_t and decide which values to enter. Then, according to

the tanh function, a candidate value vector $\tilde{C}_t$ is obtained ,which is multiplied with i_t and added to the state C_t. The formula for this part is as follows:

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

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

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

(3) Output Gate: Output the information of the current point. After running a sigmoid function to get o_t and determining which parts will be output, C_t is processed by the tanh function to obtain a value between -1 and 1. Finally, the value is multiplied with o_t to decide the ultimate output:

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

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

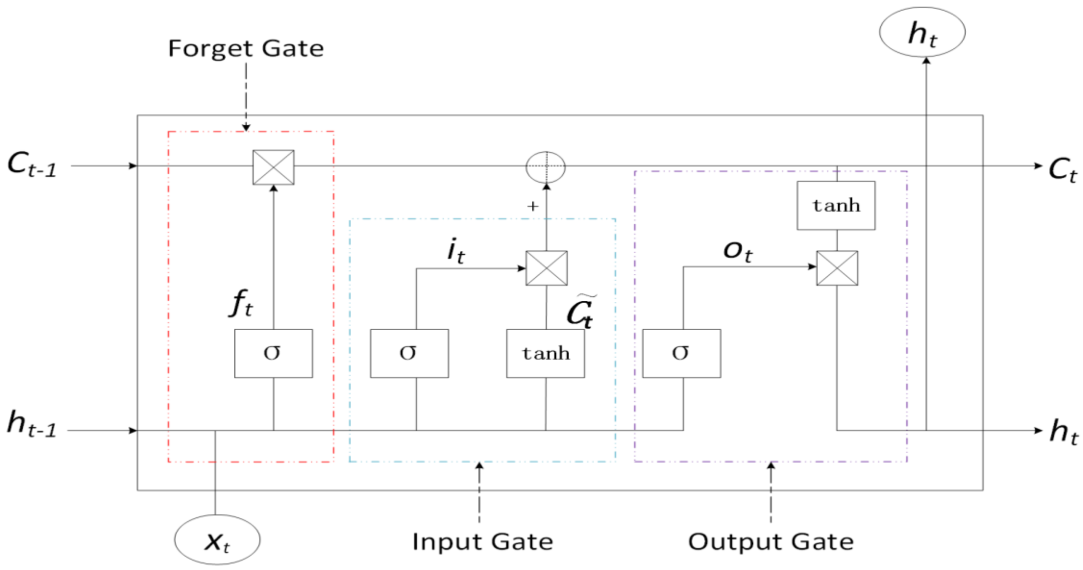


Figure 2: Basic unit of LSTM Neural Network

CNN-LSTM Hybrid Neural Network

The structure of the CNN-LSTM Hybrid Neural Network model proposed in this paper is shown in Figure 3. The model consists of two parts. The first part is the CNN part, which is mainly responsible for data input and feature extraction. And the input is a feature graph with the size of 3*17, arranged in time series. There are three convolution layers (Conc2D) in the CNN part. The number of the convolution kernels are all 15, and sizes of them are also the same of 2*2. The pooling layers (MaxPooling2D) all adopt the maximum pooling method. First of all, there is a convolution layer followed by a pooling operation. Next are two successive convolution operations. Ultimately, there is the full connection layer (Dense), extracting the characteristic data as a one-dimensional vector array whose length is 25.

In the second part, that's to say, the LSTM part, the output of the CNN part is used as the input of the LSTM Neural Network. This section consists of one LSTM layer and a full connection layer. Among them, the number of hidden layer nodes is 50 and the learning rate is 0.01. MSE is used as the loss function and Adam is used as the optimization method. Meanwhile, to avoid overfitting, the Dropout layers are added to randomly inactivate some neurons. Moreover, with the increase of training times, the model will produce the phenomenon of over-fitting. On the contrary, the fitting effect is not ideal if the training times are not sufficient. Hence, for this model, the number of iterations is set to 100, and the size of training batch is 50, which can obtain a better performance. This algorithm is based on the Keras deep learning framework.[12]

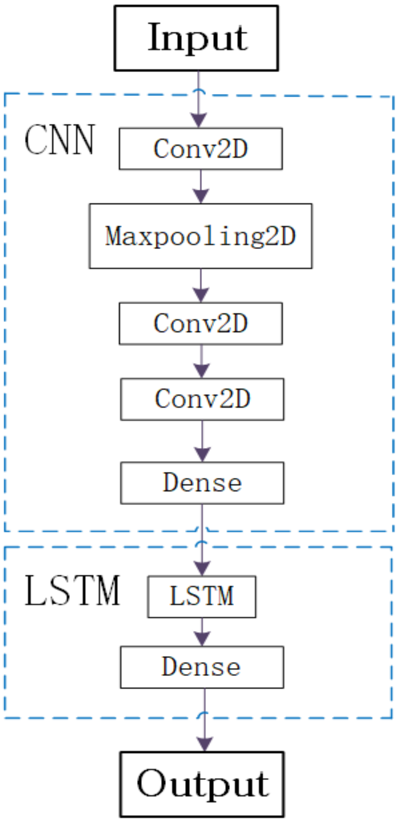


Figure 3: Structure of the CNN-LSTM Hybrid Neural Network model

3.3 Performance Metrics

The prediction ability of different neural networks to compare the price of special currency is reflected in two aspects: the accuracy of value prediction and direction

prediction. Average absolute error (MAE), root mean square error (RMSE) and Mean Absolute Percentage Error (MAPE) are used as performance indicators to quantify the value prediction ability of neural network. At the same time, recall and F1 measure are introduced to measure the prediction ability of price direction.
calculation formula is:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - f_i\right|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - f_i\right)^2}$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i - f_i}{y_i}\right|$$

where y_i is the ith true value, f_i is the ith predicted value, and N is the number of data to be evaluated. MAE, RMSE and MAPE reflect the degree of deviation of the predicted value from the true value. The smaller MAE, RMSE and MAPE, the higher the accuracy of the prediction.

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

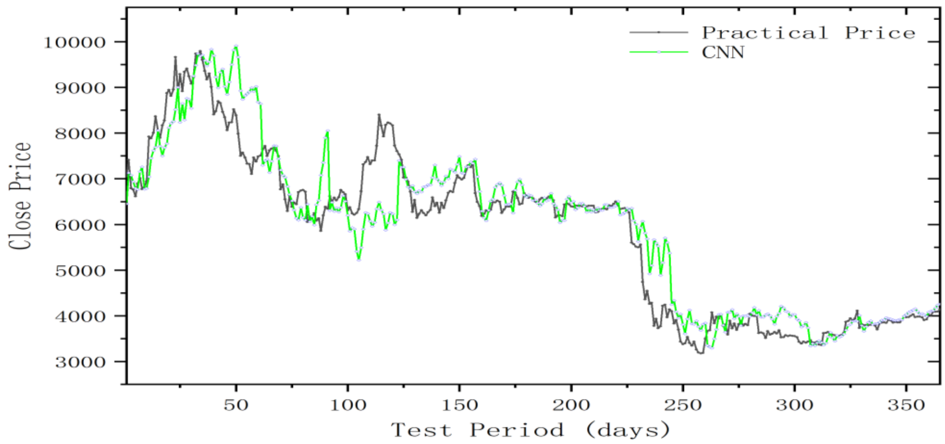
Where TP = True Positive; TN = True Negative; FP = False Positive; and FN = False Negative. F1 is the harmonic mean of Precision and Recall, which is used to comprehensively reflect the classification effect of the prediction model.

$$F_1 = \frac{2 * precision * \mathrm{recall}}{precision + recall} = \frac{2TP}{2TP + FP + FN}$$

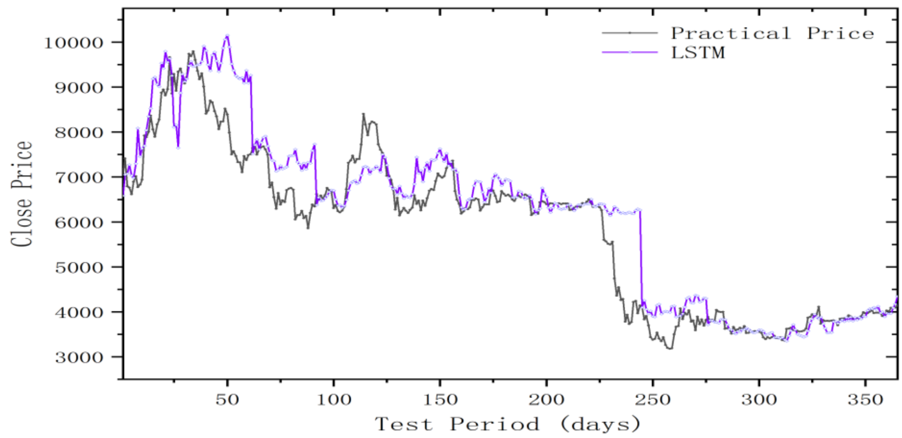
In order to obtain statistically significant values, each model was evaluated (trained and tested) 20 times and averaged as the performance metrics scores.

4. Result and Discussion

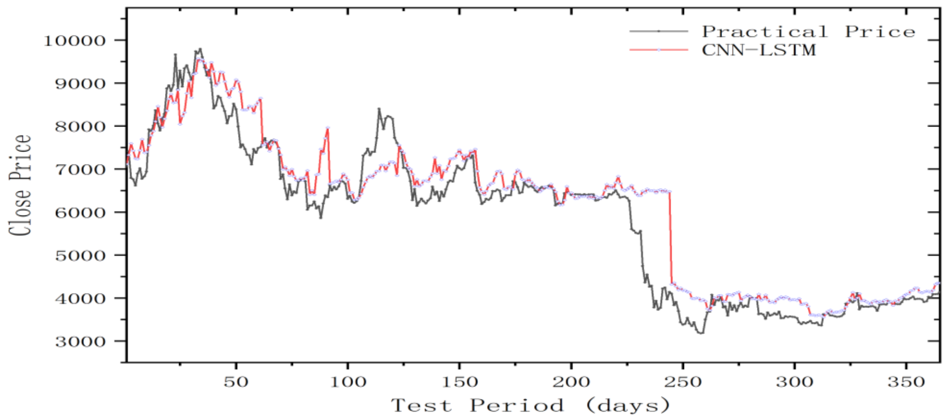
After adjusting the parameters, this paper also uses CNN and LSTM to train the same data and predict the closing price of Bitcoin. All models were trained 20 times. In addition, the prediction of single structure neural network model is compared with that of CNN-LSTM hybrid neural network.
Figure 4 shows the comparison between the Bitcoin price predicted by different models and the actual price. It can be seen that the forecast results of CNN-LSTM are very consistent with the actual price trend of Bitcoin.



(a) CNN and practical curve



(b) LSTM and practical curve



(d) CNN-LSTM and practical curve

Figure 4: Comparison diagram of prediction and practical curve

Table 2 shows the average results of performance indicators for value prediction and direction prediction, respectively. This shows that CNN neural network and LSTM model both have their drawbacks. That's to say, CNN is better than LSTM in value prediction, and LSTM is better in direction prediction. As we can see, CNN-LSTM performs better in value prediction and direction prediction. The hybrid neural network has better prediction effect than single structure neural network. Despite the high volatility of Bitcoin, the scores of these performance indicators are still not very well, so it is still a challenge for our future work.

Table 2: Comparison of model prediction result

Neural	Value Prediction			Direction Prediction		
Network	MAE	RMSE	MAPE(%)	Precision	Recall	F1
CNN	451.04	669.10	7.78	0.60	0.51	0.55
LSTM	487.12	746.53	8.49	0.69	0.54	0.60
CNN-LSTM	438.29	669.59	8.26	0.81	0.53	0.64

5. Conclusion

As mentioned above, our project uses deep learning technology to predict the price of Bitcoin based on CNN-LSTM hybrid neural network. Different from the traditional research, our project not only considers the trading information of Bitcoin itself, but also comprehensively considers various factors that may affect the price of Bitcoin, including macroeconomic variables and investors' attention and other external factors. In terms of value prediction and direction prediction, the hybrid neural network is compared with the single structure neural network. The results show that CNN-LSTM hybrid neural network performs well in Bitcoin prediction and is more suitable for Bitcoin prediction compared with CNN model and LSTM model.

In quantifying investors' attention, we use Baidu Index with the keyword "Bitcoin". So one of the possible improvements is to incorporate Bitcoin's popularity on other and global social software, such as twitter and instagram, into the data.

References:

[1]https://finance.sina.com.cn/money/forex/hq/BTCBITSTAMP.shtml

[2]Cheoljun Eom,Taisei Kaizoji,Sang Hoon Kang,Lukas Pichl. Bitcoin and investor sentiment: Statistical characteristics and predictability[J]. Physica A: Statistical Mechanics and its Applications,2019,514.

[3] S.Vassiliadis, P.Papadopoulos, M.Rangoussi, T.Konieczny, J.Gralewski, Bitcoin

value analysis based on cross-correlations, J.Internet Bank.Commer. 22(S7)(2017)1.

[4] Pavel Ciaian,Miroslava Rajcaniova d'Artis Kancs.The Economics of Bitcoin Price Formation [J]. Applied Economics,2016,48(19)

[5] Huang JZ, Huang W, Ni J, Predicting Bitcoin Returns Using HighDimensional Technical Indicators, The Journal of Finance and Data Science,

[6] LIANG Qiu, WANG Fanbin, YU Longxiu, et al.Prediction of Bitcoin price based on Wavelet Analysis [J]. Neijiang science and technology,2015,36(05):129-130.

[7] LI Jing. Using BP Neural Network to construct Bitcoin market forecasting model[J].Finance and Accounting Monthly,2016(21):33-36.

[8] Dennys C.A. Mallqui, Ricardo A.S. Fernandes,Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques,Applied Soft Computing,

[9] Da Z, Engelberg J, Gao P. In search of attention[J]. The Journal of Finance, 2011, 66(5): 1461-1499

[10] LIU Yizhe. The study of investor attention's impact on the tock performance based on baidu index[D].Hunan University,2015.

[11] Hochreiter, Sepp and Schmidhuber, Jürgen Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[12] Li Yan, Dai Wei. Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model [J]The Journal of Engineering Volume 2020, Issue 13. 2020. PP 344-347