Price Prediction of Bitcoin Based on LSTM Model

Qihuan Zhou(✉)

College of Liberal Arts and Science, University of Connecticut, Storrs, CT 06269, USA
qihuan.zhou@uconn.edu

Abstract. Contemporarily, Long-Short Term Memory Model is becoming a widely applied machine learning models in time series data prediction. This article introduces background of LSTM Models and its concept with an application example to propose some perspectives and expectations on cryptocurrency price prediction (specifically, bitcoin in this paper). According to the results, the prediction results tally well with the real price, indicating the feasibility for the implementation of such a state-of-art machine learning models in cryptocurrency pricing. These results shed light on future development of price prediction based on machine learning for investors.

Keywords: Deep Learning · LSTM · Bitcoin · Big Data Component

1 Introduction

To trace the origin of digital currency, a significant time period is 1982. In this year, David Chaum, who is considered as the 'Godfather' of digital currency published his revolutionary work toward the world, a dissertation named Computer Systems Established, Maintained, and Trusted by Mutually Suspicious Groups. In this article, Chaum mentioned the idea of decentralization and proposed several potential solutions toward how to set up trusted computing among mutually suspicious groups [4]. In 1995, Chaum created eCash that is known as the first digital currency in human history. Chaum's contribution toward digital world is undeniable. The person who really presented digital currency to public's vision, however, is Satoshi Nakamoto. In 2008, this well-known cyberpunk published his famous work Bitcoin: A Peer-to-Peer Electronic Cash System. Nakamoto presents public a new way of trading in 'purely peer to peer version' [8]. He explains how to establish electronic trading system with decentralization in detail. Only one year of the article was published, Bitcoin was created and came to public's view. After Bitcoin was created, its value is one of the most important things that people care about. The price for Bitcoin would not be defined as 'stable' ever since it has been created. Bitcoin official website defines it as 'a high risk asset' for now, and the price of Bitcoin is unpredictable due to its young economic [1]. This proclamation is a great proof that fluctuation over Bitcoin's price is dramatic. Due to its unstable price level, the prediction over Bitcoin becoming a hot topic recently.

© The Author(s) 2023
D. Qiu et al. (Eds.): ICBBEM 2022, AHIS 5, pp. 255–261, 2023.
https://doi.org/10.2991/978-94-6463-030-5_27

There are uncountable models in predicting Bitcoin's price level. The most fundamental model is named Linear Regression Model. In simple words, linear regression model could be used to predict new value of variables as output with the provided variables. When there's only one input variable, the method is called simple linear regression. When there are multiple input variables, the methods would become multiple linear regression. Although Linear Regression Model is useful, it isn't always efficient enough to predict precise data. Basic Linear Regression Model could be simply represented with

$$h(x) = w1x1 + w2x2 + w3x3 + \ldots + wnxn + b \tag{1}$$

Sometimes, the output for Linear Regression Model may not be a linear line, which leads to the drawbacks of Linear Regression Model. For non-linear data with correlated polynomial regression feature, it's hard for Linear Regression Model to testing out the correct output. Besides, Linear Regression Model doesn't own strong ability to handle with highly complex data. To improve the prediction result, plenty of new models are developed from the idea based on Linear Regression Model such as Auto Regressive Integrated Moving Average, Attention Model, DNN, etc. Among all the models, LSTM Model is definite the one which attracts the attention of companies and researchers.

The full name for LSTM Model is Long-short Term Memory Model. Schimdhube and Hochreiter brings out the idea of LSTM Model based on network element structure's improvement in Recurrent Neural Network (RNN) [6]. The original Recurrent Neutral Network Model isn't able to efficiently manage a large number of strings, which caused so called 'Vanishing Gradient Problem'. Long-short Term Memory Model was created for solving this tough work. Compared with RNN Model, LSTM Model has a better adaptation in dealing with data which contains higher time series correlation. Many firms used or are using LSTM Model into the daily life application. For example, Google applied LSTM into the area of language recognizance and translation. Apple used LSTM Model to strengthen the identification ability of Siri. On the micro level, uncountable scholars adopted LSTM as method of financial analysis. Song Gang applied LSTM Model into predicting the stock's price which randomly selected from Shanghai Stock Market, Hong Kong Exchanges and Clearing Market and Shenzhen Stock Exchange [9]. In 2018, Fisher used LSTM Model to predict the stock market's tendency and analysis what is the most profitable combination of stock by applying LSTM [5]. From these examples, LSTM exhibits its powerful function as a deep learning model. Besides, LSTM Model also embraces an excellent flexibility. There are many developed model or improved model appeared to public depends on LSTM Model. In 2019, Tima Taewook and his team mentions a combined model with LSTM Model and CNN to predict the fluctuation of stock market [10]. Kim and Change united GARCH Model and LSTM Model to a whole new model [7]. Tang and his team created the concept of TD-LSTM Model and TC-LSTM Model [11]. Chen Lu proposed the idea of LSTM-ATTE Model to predict consequence in different time period [3]. To better understand the way the LSTM Model could be used to predict Bitcoin's price fluctuation, this article would present an experiment by analyzing data and results. The rest part of the paper is organized as follows. The Sect. 2 will introduce the data origination as well as the LSTM model. Subsequently, the Sect. 3 will present the results of the prediction models. Afterwards, the Sect. 4 will demonstrate the explanation of the results as well as the limitations of the paper.

Fig. 1. Bitcoin's candlestick chart from 2016–2021 [2] (Photo credit: Original)

Eventually, a brief summary is given in Sect. 5 with drawbacks and expectation toward the future of LSTM Models application.

2 Data and Method

In this experiment, data is collected from the database of Investing.com, which is one of the top financial websites in the world which provides reliable market information. Data includes Bitcoins' useful index which could used for prediction from starting day October 30 2016 to October 30 2021. The time frame for the set of data is counted daily. Data includes Open Price at the beginning of the day, Close Price at the end of the day, Highest and Lowest price for Bitcoin of the day and trading volume. Figure 1 presents Candlestick chart for Bitcoin since 2016.

The typical structure that LSTM Model has but RNN Model doesn't is it contains several gates in different layers. There are generally three kinds of gate which are forget gates, input gates and output gates. These gates could be used to select the information which has certain value, or the invalid data would be forget and never being used for calculation on the next layer. These gates fit LSTM system perfectly. Each LSTM unit has two states. One is cell state and another is hidden state. Cell states are usually update faster in speed that hidden state. When LSTM Model dealing with input, there are four steps for input to go through. First stage is the forget gate, the general equation for forget stage is

$$f_t = \sigma[W_f\left[h_{t-1}, x_t\right] + b_f \tag{2}$$

where σ is sigmoid function which decides whether the input should be forgotten or remembered for the next layer. The second gate is called input gate, with the equations

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

and

$$C^t = tanh\left(W_C \times \left[h_{t-1}, x_t\right] + b_c\right) \tag{4}$$

It is also determined by sigmoid function which decides the updating rate of a function. C^t decides how many data would add into cell. By combining forget gate and output gate together, the equation for the state of cell should be upgraded to C^t which is represented as $f_t \times C^{t-1} + i_t \times C^t$. The last gate the input should go through is output gate and the equation is $h_t = o_t \times \tanh(C^t)$. After input go through these three gates in this unit, the output generated is next unit's hidden input. This is the basic idea and procedure of how LSTM works. For this experiment, the Python3.0 is chosen as the tool to do the prediction. The input variables include Close Price, Open Price, Highest Price of the day, Lowest Price of the day and Trading Volume. These input values are used to generate an output, which is the fluctuation percentage in Bitcoin price. The output created by LSTM is utilized to compare with the historical data and do the conclusion. First, data is imported into to LSTM Model. There are totally 1828 sets of data. These sets of data are divided into 92% and 8%. 92% which is around 1686 sets of data are used as the data for training in LSTM Model. 8%, which is around 140 pieces of data is used to test the accuracy of prediction and would also be used to do the comparison. For the variable setting, the variables are set up that every other 100 epochs, the studying efficiency would become 1/10 of the origin. For batch size, it is set as 256, this means that 256 sets of data are used to do the training each time. This number is used to speed up the process of training and fits the memory of normal computers. For epochs, the number that set as 40. These two settings represent that we are training using 256 sets of

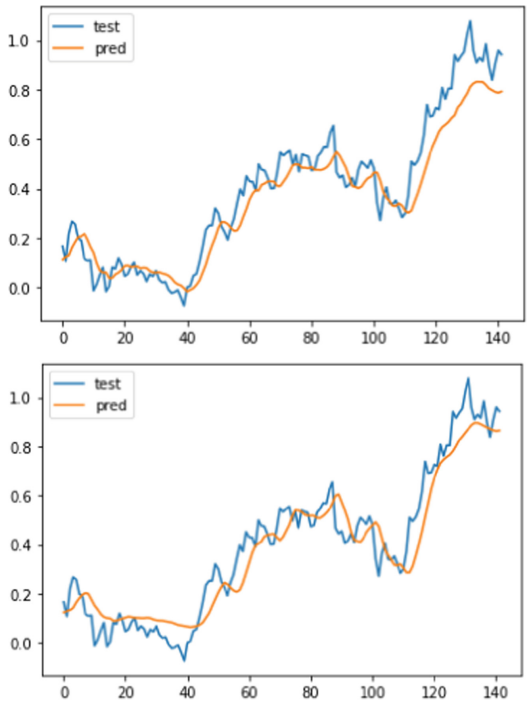

Fig. 2. The test and prediction of the first experiment and second experiment (from the upper to the lower) (Photo credit: Original)

data out of 1868 sets of data to do 40 times training. In order to minimize the influence of parameters setting inside the LSTM model. Experiment is done again with different batch size and epochs. For the second experiment, new epochs are set up as 50 and batch size as 128. In general, the experiment is done by twice to eliminate the error and two experiments results are obtained.

3 Results

Figure 2 shows 140 prediction data and compare it with real data. Blue lines represent 140 real data that is used for prediction and orange lines represent the predicted data. X axis represents the number of data that put in testing in unit of sets and Y axis represents the fluctuation ratio in unit of %. During the process, The Mean Square Error for predicted data in the first experiment is approximately 0.0089 and predicated data' MSE in second experiment is around 0.0079. During the process while LSTM is generating output. The model also presents the level of fit between predicted output and real output. This is related with loss value. The relationship between Loss value and number of Epochs is

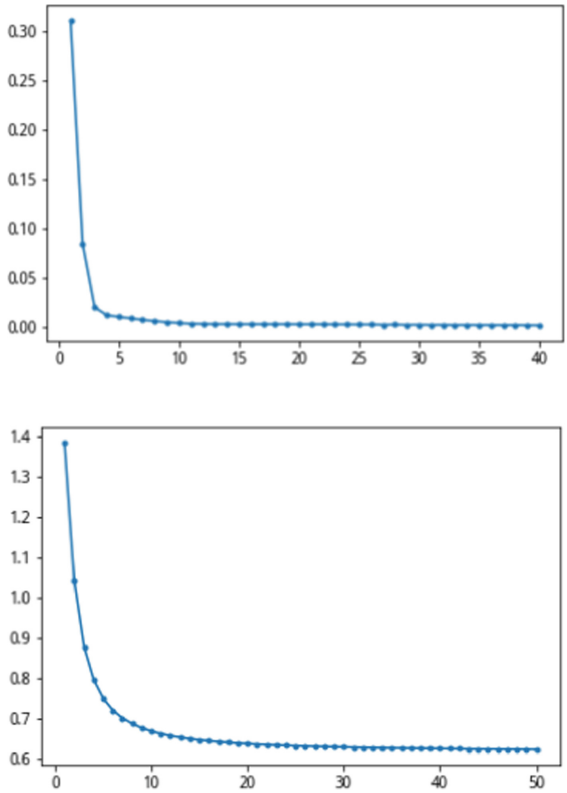

Fig. 3. Loss Value as a function of Number of Epochs in the first (upper) and second experiment (lower). (Photo credit: Original)

concluded in Fig. 3. Here, y axis represents the loss value and x axis represents the number of Epochs. According to the results, the data could be concluded that Loss value is constantly decreasing with the increasing number of epochs. This relationship generally reflects the fact that predicted data had a relatively accurate toward the real data. Comparing two experiment, second experiment embraces a larger loss with the changing of batch size and epoch as shown in the Fig. 3. In general, both of the experiment does a relatively accurate prediction toward the real value.

4 Discussion

From the obtained results and the given parameters, Long-short Term Memory Model definitely has a strong effect in predicting the price fluctuation of Bitcoin. Nevertheless, the predicted data generated by using LSTM still has obvious different while both of them are presented in the same graph. It's time to consider why such difference appeared between real and predicted. The first reason can be attributed to the total number of sets of data. Even though in our experiment, our inputs contain 1686 sets of data, the frequency doesn't seem high enough to obtain more accurate value. Higher frequency in data might lead to a more precise prediction outcome. Second possible reason causes the difference is that there might not have enough input in each set of data. Even though five inputs are used try to figure out one output. The more input provided with LSTM, the more precisely the output model could generate. Third reason is ascribed to ISTM itself. As mentioned earlier, the LSTM is setting b_y layer to layer. During the its progress dealing with the data, i.e., the predicted data in LSTM model is mainly depends on previous predicted data. In LSTM, data has a strong relationship with previous predicted data. The later predicted data is the tendency of previous sets of data rather the all sets of data. This leads to the larger standard error of predicted data than real data.

5 Conclusion

In summary, this paper assesses the feasibility of Bitcoin price prediction based on LSTM model. Primarily, the theory towards LSTM Model and the way to apply LSTM Model works in predicting the price of Bitcoin are demonstrated. According to the results, LSTM Model exhibit an impressive value on predicting price fluctuation of Bitcoin. However, the high requirement toward data collection and LSTM Model's basic theory's shortage make LSTM Model become not that perfect. Moreover, it's crucial to not ignore that there are still unpredictable reasons in real life like psychological behavior made by human beings that couldn't be simply presented in a model, e.g., psychological behavior. Nevertheless, LSTM Model owns a high flexibility, hence there will definitely appear more new models which developed based LSTM model and present a better performance in prediction in the future. Overall, these results offer a guideline for price prediction for financial data based on LSTM model.

References

1. "Open Source P2P Money." Bitcoin, https://bitcoin.org/. 2021

2. "Stock Market Quotes & Financial News." Investing.com, https://www.investing.com/. 2021
3. Chen Lu, Research On Financial Time Series Prediction Algorithm Based On LSTM Model[D]. Harbin: Harbin Institute of Technology, 2019: 48–49
4. David Lee Chaum. "Computer Systems Established, Maintained and Trusted by Mutually Suspicious Groups." University of California, Berkeley, 1982.
5. Fischer T. KRAUSS C. Deep learning with long short-term memory networks for financial market predictions [J]. European Journal of Operational Research, 2018, 270(2): 654–669
6. Hochreiter, S., & Schmidhuber, J"urgen. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.
7. Kim H Y, Chang H W. Forecasting t Systems with the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models [J]. Expert System with Application, 2018
8. Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." Decentralized Business Review (2008): 21260.
9. SONG G, ZHANG Y F, BAO F X, et al. Stock prediction model based on particle swarm optimization LSTM[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(12): 2533-2542.
10. Taewook K, Young KH. Forecasting Stock Prices with a Feature Fusion LSTM-CNN Model Using Different Representations of the Same Data [J]. Pumbed, 2019, 14(2): 1–23
11. Tang D, Qin B, Feng X, et al Effective LSTMs for target dependent sentiment classification [C]. Proceedings of the 26th international Conference on computational Linguistic, 2016

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.