Price Prediction of Cryptocurrency based on LSTM Model: Evidence from Ethereum

Diming Xu

University of California Irvine, Irvine, CA 92612, USA

dimingx1@uci.edu

Abstract. Contemporarily, under the impacts of COVID-19 and regional conflicts with radicalness fiscal policy, the prices of cryptocurrency have been fluctuated dramatically. Among various types of cryptocurrency, Ethereum is one of the most volatility assets. In order to avoid risks as well as gain extra return in the crypto market, it is necessary to construct accurate prediction approach. In this paper, the Long Short-Term Memory algorithm will be used to predict the future price of Ethereum by learning Ethereum's past price direction data. price trend by learning Ethereum's past price trend data. Based on the analysis, the predicted values of the trained model fit well with the actual data, with the regression evaluation index R2 of 97.08% and MAPE of 6.89%. According to the results, it is feasible to predict the future price trend through the past price trend data. Nevertheless, it should be noted that the stochastic process in data training might lead to the instability of model performances. Hence, it is necessary to train the data several time to select the best models. Overall, these results shed light on guiding further exploration of cryptocurrency price forecasting in terms of the state-of-art neural networks.

Keywords: Price Prediction; LSTM; Cryptocurrency; Ethereum.

1. Introduction

Contemporarily, high volatility intrinsic of cryptocurrency due to various impacts (e.g., regional conflicts, pandemic as well as regulations and lows) has attracted lots of investors. As a matter of fact, the concept of cryptocurrency is based on the techniques of blockchain, which is first proposed by Nakamoto (the white paper of Bitcoin in 2008) [1]. After development of decades, there are more than thousands of types crypto currencies with market value larger than 2 trillion dollars [2-4].

Among various of cryptos, the Ethereum is one of public databases that maintains a permanent record of digital transactions. Importantly, this database does not require any central authority to maintain and protect it. Instead, it operates as a "trustless" transaction system, i.e., a framework in which individuals can engage in peer-to-peer transactions [5, 6]. Its value comes from a decentralized consensus-based mechanism, a medium of exchange that uses cryptographic principles to secure and control the creation of units of exchange. In reality, it uses distributed ledger blockchain technology to ensure that transactions are secure and trustworthy. However again, because of the lack of regulation and other factors, Ethereum's price is completely dependent on people's consensus.

Ethereum's transactions also rely on consensus, and in Ether, all participants must agree on the order of all transactions that have occurred, regardless of whether or not the participant was involved in a particular transaction. The order of transactions is critical to the consistent state of the ledger. If the order of transactions cannot be determined, there is a risk of double spending where two parallel transactions transfer the same coins to different recipients, thus earning money out of thin air. Since the network may involve anonymous parties who do not trust each other, it is important to adopt a mechanism to protect the ledger from fraudulent or unfavorable participants attempting to double-spend. In the current implementation of Ether, this mechanism is established through mining based on a proof-of-work (PoW) scheme. All participants must agree to a common ledger, and all participants have access to all entries ever recorded. The consequence is that PoW has a negative impact on the performance of transaction processing. Regarding to the data stored on the ledger, all participants can access them even though the records are anonymous [7].

Since Ethereum is a completely overhead financial product that does not depend on the value of any real item, many stock market prediction methods would not work for it. On this basis, the exactly way to predict Ethereum's price trend is a very important question. Among various of price (a kind of time series) forecasting methods, LSTM (a type of RNN) is believed as a useful tool with high accuracy. With this in mind, this study investigates the feasibility to forecast the price of ETH based on LSTM [8-14]. The reminder of the study will be separated as follows. The Sec. 2 will describe the data collection, cleaning and processing as well as the training models and metrics. The Sec. 3 will demonstrate the results and give some descriptions accordingly. The Sec. 4 will discuss the limitations of the study as well as present the further outlooks. Eventually, a brief summary will be presented in the Sec. 5.

2. Data & Method

2.1 Data Source

The data source is Historical Data for Ethereum in the Cryptocurrency section on investing.com. Since cryptocurrencies are traded 24/7, there is no need to filter the data, the date range selected is 2018/1/6 to 2021/11/6. Fig. 1 shows the price trend of ETH during the selecting period. In this study, the time division method is adopted to divide the price data into Training Set and Testing Set according to the present time interval. The time division is to use the time characteristics of financial time series, set the cut-off date according to the transaction time (Date), and set the data after the training date as the test set, and the final test set contains only some data The selected Training Data interval is from 2018/1/6 to 2020/09/18, and the Testing Data interval is from 2020/09/19 to 2021/11/6. Fig. 2 depicts the trend of price changes of ETH during the period.

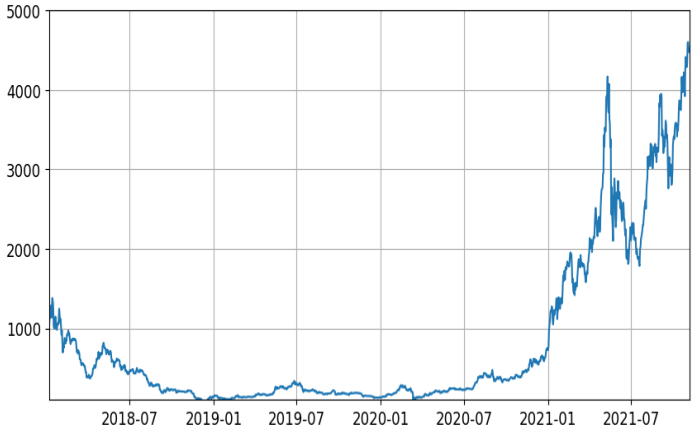

Fig 1. Price trend of ETH.

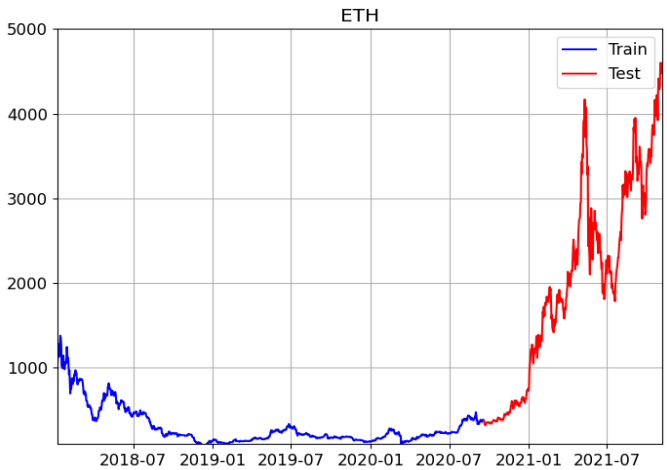

Fig 2. Trend of price change of ETH.

2.2 Models and Metrics

This paper uses the LSTM model which is a special type of RNN. To evaluate the models, various metrics and indicators including MAE, MSE, MAPE and R^2 are utilized. The smaller the value of the former three indicators, the better the performance of the mode; whereas the larger the R^2, the better the model.

3. Results & Discussion

The final model fitting results are shown in Fig. 3 and Fig. 4. To be specific, Train is the data in the training set, Test is the data in the Test set, and Result is the prediction value given by the model. Based on the figures, the data in the training set is in its "quiet period" and the fluctuations are smaller compared to the data interval in the test set, but the model still has a good fit and prediction effect even though the price trend in the test period interval is significantly different from the training time interval.

On this basis, it indicates that it is feasible to use the neural network training model to predict Ethereum's price trend. To evaluate the performances of the models, various metrics and indicators have been utilized. To be specific, the MAE of the model is 140.53, MSE is 40354.39, R2 is 97.08%, MAPE is 6.89%. According to the analysis, the fit is very good in terms of parameters, indicating the feasibility to achieve the accurate prediction for cryptocurrency based on LSTM.

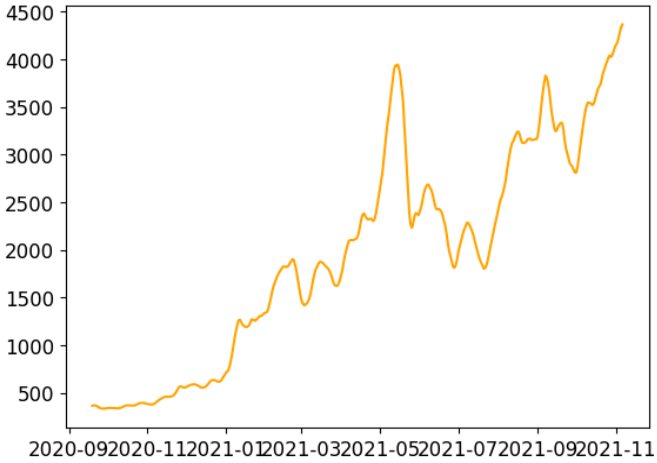

Fig 3. Model fitting results.

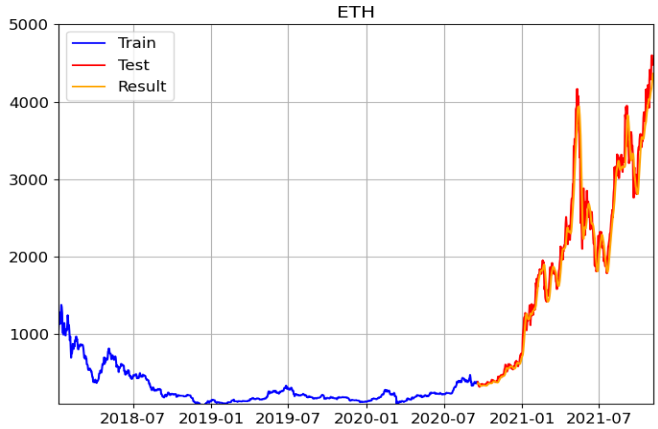

Fig 4. Comparison of model fitting results with test results.

4. Limitation & Prospects

Since random numbers are introduced in the model training, the results of each training will have deviations, and the figure shows the test results after several retraining. As illustrated in Fig. 5, the fit

rate of the model is always fluctuating randomly, where the average value of R2 is 94.58% and the average value of MAPE is 8.80%. In the worst case, the value of R2 is 86.53% and the value of MAPE is 14.14%, while in the best case the value of R2 is 97.08% and the value of MAPE is 6.89%. The maximum deviation value, R2 is 10.55% and MAPE is 7.25%. It shows that there is a large difference in the fit of the model in the different test results. If the training data set is increased, the randomness of the fit should be improved to some extent, or more independent variables should be added to integrate the training. In order to obtain the best performance model, it is necessary to carry out more data scanning in the future. Besides, other complex models can be integrated together to enhance the accuracy, including multi-factorial CNN or boosting tree scenarios and stochastic process.

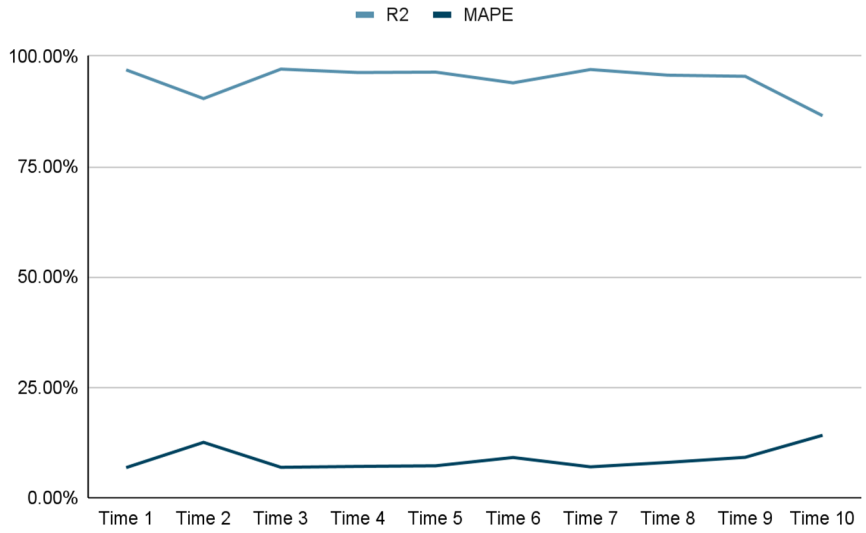

Fig 5. Changes in regression metrics with different random srand

5. Conclusion

In summary, this paper investigates the arability of price prediction of Ethereum based on the state-of-art machine learning scenarios (i.e., the LSTM model). To be specific, the LSTM model is implemented. To be specific, this study collected time series transaction data of Ethereum price between 2018/1/6 to 2021/11/6, constructed a prediction model of cryptocurrency market price based on Long Short-Term Memory algorithm, combined with statistical indicators for quantitative analysis of accuracy, and have the following conclusions. According to the analysis, it is feasible to perform cryptocurrency price prediction by the Long Short-Term Memory algorithm. Based on the evaluations, the training results may vary widely from one training model to another due to the introduction of random numbers in the training model. In future research, more price indicators can be introduced to train the model, which may reduce the randomness of the training results. Overall, these results offer a guideline for crypto assets evaluations in terms of deep learning approaches.

References

[1] Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Decentralized Business Review, 21260 (2008).

[2] Qin, W.: Research on Spillover effect of Cryptocurrencies, Precious Metals and Chinese Stock Market (Master Dissertation, Yunnan University of Finance and Economics), 2022.

[3] Gu, C., Lu, B., Liu, Y.: The Impact of the unlimited quantitative easing monetary Policy of the United States on cryptocurrencies and digital RMB. China Circulation Economy 9,117-128 (2021).

[4] Zhou, W., Li, Y., Tan, J.: Research on the correlation between cryptocurrency and stock market risk. China Soft Science, S1,116-126 (2021).

[5] Tikhomirov, S.: Ethereum: state of knowledge and research perspectives. In International Symposium on Foundations and Practice of Security (pp. 206-221). Springer, Cham (2017, October).

[6] Preethi, K.: How does Ethereum work, anyway. Medium, 1, 2017.

[7] Martin, V., Philipp, S.: Comparison of ethereum, Hyperledger fabric and corda. Frankfurt School Blockchain Center, 8, 1-8, (2017).

[8] Huang, X., Zhang, W., Tang, X., et al. Lstm based sentiment analysis for cryptocurrency prediction. In International Conference on Database Systems for Advanced Application. Springer, Cham 617-621 (2021, April).

[9] Chen, C.: Research on multivariate LSTM Price prediction based on herd effect in cryptocurrency market. Master's Thesis, Nanjing University of Information Science and Technology, 2022.

[10] Liu, C.: Research on Risk measurement of digital cryptocurrency based on LSTM-J-C model. Master's Thesis, Nanjing University of Information Science and Technology, 2022.

[11] Han, R.: Bitcoin Price prediction based on Machine Learning. Master's Thesis, Shandong Normal University, 2022.

[12] Chen, Y.: High-frequency financial time series Prediction and anomaly Detection based on LSTM and Autoencoder. Master's Thesis, Sichuan University, 2021.

[13] Jing, P.: A comparative study of Bitcoin price prediction based on multiple models. Master's Thesis, Shanxi University, 2021.

[14] Ren, H.: An empirical study on the prediction of cryptocurrency price trend by time-series Deep Learning Model. Master's Thesis, Southwest University of Finance and Economics, 2020.