Prediction of Cryptocurrency Price Dynamics with Multiple Machine Learning Techniques

Wang Zhengyang
Department of information and Communication Engineering, Xi'an University of Posts & Telecommunications
Xi'an, China
13679251160@163.com

Li Xingzhou
School of Business, University of Manchester
Manchester, UK
xingzhou.li@outook.com

Ruan Jinjin
Department of Electrical Engineering
Chongqing University
Chongqing, China
942820542@qq.com

Kou Jiaqing
Department of Mathematical Sciences
University of Liverpool
Liverpool, England
sylviasomerhalder@gmail.com

ABSTRACT

Nowadays, encrypted digital currency offers a new way of secure trading and exchanging and has become increasingly important in our financial system. However, the temporal dynamics of cryptocurrencies is highly complex, and predictions are still challenging. In this study, we establish two prevailing machine learning models, fully-connected Artificial Neural Network (ANN) and the Long-Short-Term-Memory (LSTM), to predictively model the price of several popular cryptocurrencies, including Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR). We evaluate model performance and conduct sensitivity analysis to further understand our model behaviors. We find that although LSTM seems more appropriate for time sequence prediction task, ANN, in general, outrivals LSTM in our experiments. Using price information from other different cryptocurrencies for joint training and prediction could largely facilitate the prediction of BTC. Finally, the model predictive error is highly sensitive to the time scale of interest.

CCS Concepts

Computing methodologies→Neural networks

Keywords

Cryptocurrency price prediction; matching learning models; ANN; LSTM

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
ICMLT 2019, June 21 – 23, 2019, Nanchang, China
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6323-5/19/06…$15.00
https://doi.org/10.1145/3340997.3341008

1. INTRODUCTION

Cryptocurrency is a digital currency that uses cryptography to make sure the transfer and exchange over the network could be secured. As one of the most well-known cryptocurrencies, Bitcoin builds a ledger maintained over a network based on distributed technology. One of the most important features of Bitcoin is decentralization, which is different from centralized traditional currency, meaning that transaction and the management of Bitcoin is not regulated by any central authority. As a result, people trade with each other only between buyers and sellers, not relying on a third-party trusted platform or other any banks [1]. In addition, Bitcoin is secure due to cryptographic encryption technology. All transaction information is recorded and cannot be modified, even if some information is useless. In that way that assures that a permanent character.

In the online financial system, Bitcoin and other cryptocurrencies are particularly significant. They solve the verification risk in people's transaction settlement process through the blockchain technology. Since the ledger of Bitcoin does not belong to a single person or collective management but distributed and tractable throughout the Internet, every transaction requires all members to vote on the authenticity of the "bill chain" to be recorded on the ledger. This makes it impossible to forge transactions and books, thus greatly reducing costs and risks in the settlement process of online financial system [1]. Furthermore, cryptocurrency breaks the old rules that only the authority organization, such as government, is allowed to issue currency, stipulate currency issuance procedures, and decide which deals are valid or invalid. Conversely, cryptocurrency's transaction settlement does not require verification and endorsement by any authority organization. As a matter of fact, not only in the financial system, the digital cryptocurrency mechanism is an incentive mechanism in the self-organization of the complete decentralization of the public blockchain. That's why tokens like Bitcoin have to be issued on the public blockchain, which is otherwise just a distributed ledger. In recent years, with the rapid development of blockchain technology and the application of multiple practical scenarios, the dynamics of cryptocurrency and digital currency market prices have attracted more a lot of people's attention. This makes it all the more important, therefore, to better understand, model, and predict the market values of cryptocurrencies, with e.g., advanced statistical or machine learning methods.

Given the importance and popularity and advantages of cryptocurrencies, a large amount of efforts has been allocated to better understand the characteristics, temporal patterns and predictability of cryptocurrencies. Previous research methods attempted to find factors that affect the price of cryptocurrencies. For example, Liu Y et al.[2] introduced a quantitative research method to analyze the historical data of mainstream cryptocurrencies to influence the underlying factors of cryptocurrency prices. The conclusion drawn in the study is that the crypto-money momentum and the investor's attention are the factors that influence the price. Shah et al.[3] made a "latent source model" to predict price variation of Bitcoin with the method of Bayesian regression. The "triangular pattern" and "head and shoulder pattern" in the experimental results indicated that the trading strategy generated based on this method and model was successful. Moreover, the transaction return rate can reach 89% in 50 days, and the Sharpe ratio was 4.10. Greaves et al.[4] extracted several features and used the established linear regression, logistic regression models, and neural network methods to predict the price of bitcoin. In the end, the performance of training through the neural network method performed best in this study, and the classification accuracy rate reached 55.1%. Madan et al.[5] used support vector machines, random forest algorithms, and binomial logistic regression, and selected 16 features associated with Bitcoin transactions and Bitcoin networks to predict bitcoin prices. The accuracy of predicting price changes at this interval of 10 minutes using these algorithms was 50%-55%. Indera et al.[6] presented a model that accurately predicted the price of Bitcoins based on the nonlinear regression of multilayer perceptron. This model used the particle swarm optimization method to optimize, and the regression test proves that the model can fit the dataset for prediction well.

Compared with traditional financial instruments, cryptocurrencies can be very challenging to predict due to the lack of indicator data. Coupled with government policy and other uncertainties, currency forecasts can be difficult. In fact, the results of a single-point prediction model based only on historical price data are still generally inaccurate, and there is almost no use in actual trading. In this study, we aim to predict the cryptocurrency price dynamics by using prevailing time series machine learning models.

2. METHODOLOGY

2.1 Data collection

In this research work, encrypted currency (Bitcoin or BTC) from August 7, 2015, to June 2, 2018 is used, including four features, opened currency, close, high, low and market capitalization. Figure 1 shows several important characteristics of the Bitcoin time series data. The upper left panel is the frequency distribution data of BTC, including opening, closing, highest and lowest prices. Upper right scatter plot describes the highest (x-axis) and the lowest price (y-axis) while the lower left scatter plot depicts the relationship between the opening and closing prices. The lower right panel is a scatter plot of the opening price of BTC and Ethereum-ETH.

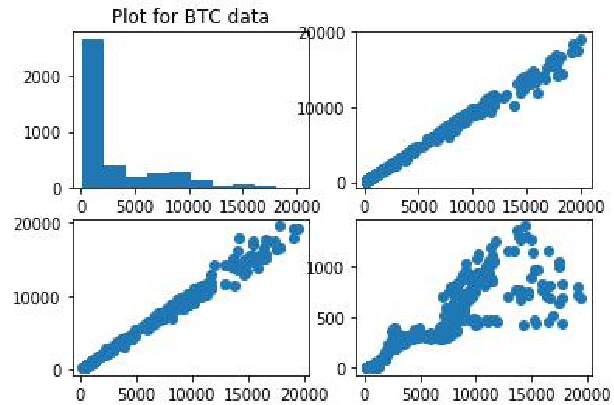


Figure 1. Characteristics of Bitcoin (BTC) data used in this study.

2.2 Machine learning models

Two advanced machine learning models are used in this study. The first on is fully-connected Artificial Neural Network (ANN)[7], which is trained by gradient based back-propagation algorithm. The mathematical equation for each node is as follows.

$$node = f\left(\sum_{i=1}^{i=n} w_i \cdot x_i\right) \quad (1)$$

$$f_{(node)} = \frac{1}{1+e^{-node}} \quad (2)$$

where w is the weight of the previous node to the next node connection, x is the input value, n is the number of input nodes, and f is the activation function. As is shown in Eqn. 2, here we employ sigmoid activation function.

The second machine learning model is called Long-Short-Term-Memory (LSTM) [8]that consists of multiple cells that propagate information through time. A Cell has three Gates (input, get, output) and a cell unit. Gate uses a sigmoid activation function, while input and cell state are usually converted using tanh. In Figure 2, $\odot$ is the element-wise product. Inner products will be represented as ·. Outer products will be represented as $\otimes$.

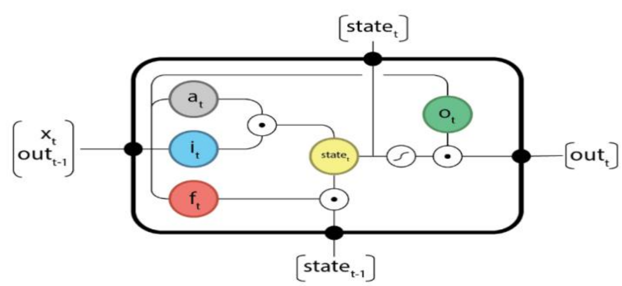


Figure 2. Architecture of Long-Short-Term-Memory (LSTM) model

First, the forget gate decides whether to pass the information learned at the last moment or pass it partially, that is, selectively filter the old information as

$$f_t = \sigma\left(W_f \cdot x_t + U_f \cdot out_{t-1}\right) \quad (3)$$

σ indicates the sigmoid activation function. W_f and U_f is the weight parameter. x_t is the current input information. out_{t-1} indicates the data output from the previous moment. In this step, the forget gate reads the current moment's input and the previous

moment's output, and then outputs a value between 0 and 1 for each number in the cell state. 1 means "completely reserved" and 0 means "completely discarded".
The second step is to determine what new information is stored in the cell state. Then, the input gate: The output value is multiplied to the output of the tanh layer to serve as a scaling. σ represents the sigmoid function. W_i and U_i is the weight parameter.

$$i_t = \sigma\left(W_i \cdot x_t + U_i \cdot out_{t-1}\right) \qquad (4)$$

The input activation generates candidates for updated values as

$$a_t = \tanh\left(W_a \cdot x_t + U_a \cdot out_{t-1}\right) \qquad (5)$$

The *tanh* activation layer will create a new candidate value vector to be added to the state. W_a and U_a is the weight parameter. Meanwhile, the internal state help generate new cell state to be used by model output. $state_t$ indicates current internal state. i_t and a_t are the values generated by the input gate and input activation in the last two steps. f_t means the forget gate, $state_{t-1}$ means internal state of the previous moment.

$$state_t = a_t \quad i_t + f_t \quad state_{t-1} \qquad (8)$$

Finally, the output gate predict result values to be:

$$o_t = \sigma\left(W_o \cdot x_t + U_o \cdot out_{t-1}\right) \qquad (6)$$

$$out_t = \tanh\left(state_t\right) \quad o_t \qquad (7)$$

σ is also the sigmoid activation function, which determines how much "information" of the cell state will be output. W_o and U_o are the weight parameters.
Finally, the cell state is processed through the tanh layer to get a value between -1 and 1 and multiplied by the output of the sigmoid gate. $state_t$ indicates the updated internal state. out_t indicates the final result, which will only output the part that determines the output.

2.3 Modeling experiments

In all experiments, the data set is split into 80% training data and the 20% test data. In ANN model, three hidden layers are considered, and other critical hyperparameters are set to follows: nodes per layer is 20, the activation function is linear Rectifier, training epoch number is 500, batch size is 100. For LSTM, the memory length is five time step. Also, we use the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) to evaluate the performance of the model training results.

3. RESULTS AND DISCUSSION

3.1 ANN vs LSTM

We first present results of Bitcoin time series prediction and compare the fully-connected Artificial Neural Network (ANN) and the Long-Short-Term-Memory (LSTM) models based on the performance metrics of Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). ANN is a fully connected neural network, and LSTM can bypass some units and remember long steps. Therefore, LSTM can solve the problem of gradient disappearance to some extent. We use the highest price of the daily Bitcoin as the target variable. As is shown in Figure 3, ANN model performs better than LSTM. In the single point prediction process for Bitcoin, the model prediction result of the ANN is RMSE 209 and the MAE value is 99. However, the RMSE and MAE values of LSTM are 3894 and 2789, respectively.

3.2 Joint prediction

We jointly model the price dynamics of Ethereum (ETH), Ripple (XRP), Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR). Our hypothesis is that by using multiple cryptocurrencies, the predictability of each cryptocurrency could be potentially enhanced since different cryptocurrency may partly share information that is helpful for predicting others. In fact, there are also views in the market that bitcoin, as a leader, has certain effects on other cryptocurrencies, attracting a large amount of money from the cryptocurrency market into bitcoin. When we look at the relationship between several popular cryptocurrencies and bitcoin, the peaks and troughs in the value of cryptocurrencies and bitcoin have the same resonance, and the resonance period can be accurate to weeks. By comparing the time series predictions with a single cryptocurrency, we confirm that using the six cryptocurrency time series for joint prediction is superior and these currencies indeed have overlapping information in the past. Therefore, the accuracy of prediction can be effectively improved. The joint predictions of the experimental results RMSE and MAE were 137 and 30, respectively, which were lower than the results of the single predicted currency price model.

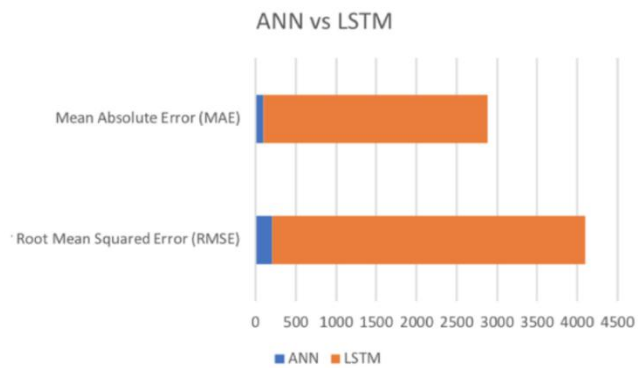


Figure 3. Model performance of Bitcoin price dynamics prediction with fully-connected Artificial Neural Network (ANN) and the Long-Short-Term-Memory (LSTM).

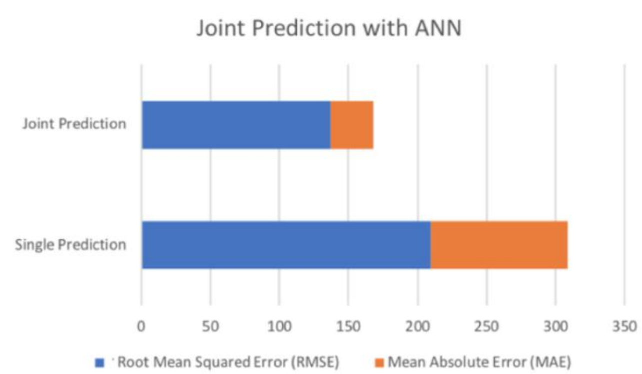


Figure 4. Performance metrics for Bitcoin ANN and LSTM predictions, when jointly using and predicting Bitcoin, Ethereum, Ripple, Stellar Lumens, Litecoin, and Monero.

3.3 Sensitivity of critical hyper-parameter

We test the sensitivity of important model parameter. In particular, for LSTM, the prescribed memory length parameter is critically important, which allows the model to "look back" several time step and learns historical knowledge for future prediction. We set the memory length parameters in the model to 5, 10, 15, and 20 time steps, respectively. And, we hypothesize that changing memory length could significantly affect LSTM model performance. However, as is shown in Figure 5, only RMSE shows some sensitivity, but MAE is insensitive. Increasing the length of memory does not necessarily mean a better model performance, while the initial guess of 5 time step is the optimal one.

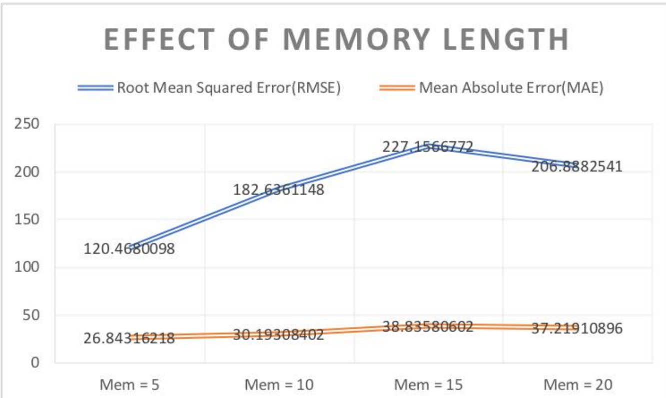
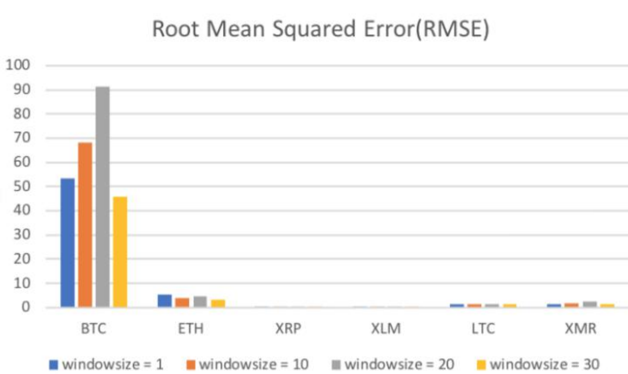


Figure 5. Sensitivity of LSTM model to the length of internal memory.

3.4 Predictability at different time scale

In this experiment, we show the trend of changes in currency price predictions at different time scales by conducting a moving average on the time series data. We use the ANN model as an example, to predict the encryption currency prices of Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR) in different periods and find that the prediction errors increase with the lengthening of time scale. Figure 6 shows that the prediction errors for the six cryptocurrencies after 1 day, 10 days, 20 days, and 30 days, respectively. Taking the predicted bitcoin price as an example, when the smoothing window size is set to 1, 10, 20, 30, the corresponding RMSE is 53.30, 67.99, 91.41, 45.71; the MAE was 28.93, 39.21, 60.63, 23.43, respectively. In general, model predictability at a longer time scale (30 days) performs much better than an intermediate time scale (e.g., 20 days), while is comparable of that with the shortest time scale (1 day). It implies that the unpredictive error first increases and then decreases through time.



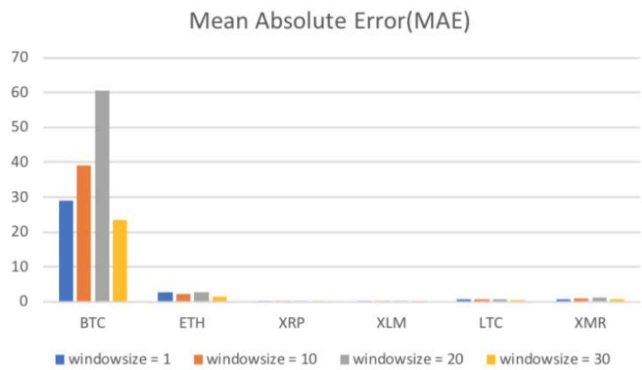


Figure 6. RMSE (upper) and MAE (lower) of ANN model predictions at different time scales (1 day ~ 30 days) for Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR).

3.5 Limitation and future work

This study is limited in many aspects including, for example, computational resources and model optimization efforts. Future work could leverage the power of cloud computation or super computer computation (cluster) to facilitate and accelerate the training of ANN and LSTM models. Furthermore, Model architectures in terms of the number of hidden layers and the number of neurons each hidden layer could be further improved and potentially upgrade the predictability of model. Finally, collecting more and better relevant datasets of digital cryptocurrencies may also improve the prediction accuracy of cryptocurrencies use in this study.

Our future work will include (1) systematically assess the relationships among different cryptocurrencies of interest; (2) improve the parameterization of the machine learning models; (3) apply the established framework to a wider range of cryptocurrencies datasets and improve the predictability of the cryptocurrency time series dynamics, therefore, better understand the future fate of certain cryptocurrency.

4. CONCLUSIONS

In this study, we establish two advanced machine learning models, fully-connected Artificial Neural Network (ANN) and the Long-Short-Term-Memory (LSTM), to understand and predict cryptocurrency price dynamics including Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR). Our modeling results show that ANN in general outperforms LSTM. In addition, the modeling of a single cryptocurrency could be improved by using several other cryptocurrencies time series for jointly training and prediction, which indicating shared “information” among various different cryptocurrencies. It also implies that there are potential lead-lag relationships among different cryptocurrencies market price dynamics that warrant future study. Theoretically, LSTM is more suitable than ANN in terms of modeling time series dynamics, given the assumption that future state of a time series is highly dependent on its history evolution. However, our results indicate the opposite, which could be explained by a limited self-dependency of cryptocurrencies time series that considered in this study. Moreover, our sensitivity analysis show that modeling results are also sensitive to critical model parameters and the time scale of interest. In general, model error increases when the predictive time scale is larger.

5. REFERENCE

[1] Rauchs, Michel. "Cryptocurrencies Meeting Business Ecosystems: The Case of Bitcoin." Available at SSRN 3095875 (2016).

[2] Liu Y, Tsyvinski A. Risks and returns of cryptocurrency[R]. National Bureau of Economic Research, 2018.

[3] Shah, Devavrat, and Kang Zhang. "Bayesian regression and Bitcoin." 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE, 2014.

[4] Greaves A, Au B. Using the bitcoin transaction graph to predict the price of bitcoin[J]. No Data, 2015.

[5] Madan I, Saluja S, Zhao A. Automated bitcoin trading via machine learning algorithms[J]. URL: http://cs229. stanford. edu/proj2014/Isaac% 20Madan, 2015, 20.

[6] Indera, N. I., et al. "Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators." Journal of Fundamental and Applied Sciences9.3S (2017): 791-808.

[7] Jain, Anil K., Jianchang Mao, and K. M. Mohiuddin. "Artificial neural networks: A tutorial." Computer 3 (1996): 31-44.

[8] Selvin S, Vinayakumar R, Gopalakrishnan E A, et al. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]// International Conference on Advances in Computing. 2017.