A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market

Some of the authors of this publication are also working on these related projects:

Pathfinding car racing game using dynamic pathfinding algorithm and algorithm A View project

Power Electronics and Drives View project

A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market

1st Ferdiansyah
Department of Informatics,
Universitas Bina Darma
Palembang, Indonesia
ferdi@binadarma.ac.id

2nd Siti Hajar Othman
School of Computing, Faculty of Engineering,
Universiti Teknologi Malaysia,
Johor Bahru, Johor, Malaysia
hajar@utm.my

3rd Raja Zahilah Raja Md Radzi
School of Computing, Faculty of Engineering,
Universiti Teknologi Malaysia,
Johor Bahru, Johor, Malaysia
zahilah@utm.my

4th Deris Stiawan
Faculty of Computer Knowledge,
Universiti Sriwijaya,
Bukit Besar, Palembang, Indonesia
deris@unsri.ac.id

5th Yoppy Sazaki
Faculty of Computer Knowledge,
Universiti Sriwijaya,
Bukit Besar, Palembang, Indonesia
yoppysazaki@gmail.com

6th Usman Ependi
Department of Informatics,
Universitas Bina Darma
Palembang, Indonesia
u.ependi@binadarma.ac.id

Abstract—**Bitcoin is a kind of Cryptocurrency and now is one of type of investment on the stock market. Stock markets are influenced by many risks of factor. And bitcoin is one kind of cryptocurrency that keep rising in recent few years, and sometimes sudden fall without knowing influence behind it on the stock market. Because it's fluctuations, there's a need and automation tool to predict bitcoin on the stock market. This research study learns how to create model prediction bitcoin stock market prediction using LSTM, LSTM (Long Short Term Memory) is another type of module provided for RNN later developed and popularized by many researchers, like RNN, the LSTM also consists of modules with recurrent consistency. The Method that we apply on this research, also technique and tools to predict Bitcoin on stock market yahoo finance can predict the result above $ 12600 USD for next days after prediction, in the last section we make conclusions and discuss future works.**

Index Terms—**Cryptocurrency, Bitcoin prediction, Bitcoin Stock Market Prediction, LSTM.**

I. INTRODUCTION

Cryptocurrency has been around for several years and has now become quite popular, widespread, and also surrounded and there is a lot of controversy from innovative developments.

Cryptocurrencies are a digital currency where transactions can be done by online transactions, unlike the common currency, cryptocurrency is designed based on cryptography.

Bitcoin is one kind of Cryptocurrency no regulation from any party and decentralized. The unique characteristic of Bitcoin is daily price fluctuations and always change every day. The value Bitcoin Exchange rate to (USD) is $ 12,354.73 USD on 28 June 2019 in yahoo finance stock market[1] and sometimes keep rising and sudden fall on march the value is $ 3900 USD.

The Stock markets are influenced by many uncertainties factor such as political issue, the economic issue at impacted to local or global levels. To interpretation key of success,

factor to providing accurate predictions is complicated work. For the market, we can analyze with any techniques such as technical indicator, price movements, and market technical analysis.[2]

To solve the problem above, regarding the fluctuations there's a need automation tool for prediction to help investors decide for bitcoin or other cryptocurrency market investment. Nowadays the automation tools are usually used in common stock market predictions, and we can do the same works and strategy on this domain cryptocurrencies.

LSTM (Long Short Term Memory) is another type of module provided for RNN. LSTM was created by Hochreiter & Schmidhuber (1997)[3] and later developed and popularized by many researchers. Like RNN, the LSTM) also consists of modules with recurrent consistency.

This paper studies about bitcoin and stock market predictions, method, technique, and tools from a big number of resources paper, and other available sources.

II. LITERATURE REVIEW

A. Cryptocurrency and Bitcoin

The history of cryptocurrency (Cryptographic currency) begins in the 1980s started with David chaum, In his paper, he proposed a novel of a cryptographic scheme to blind the content of the message before it is signed so that the signer cannot determine the content. These blind signatures can be publicly verified just like a regular digital signature. Chaum proposed digital cash approach in such a way that is untraceable by another party.[4]

The rise of cryptocurrency started on B-money In 1998, Wei Dai proposed b-money[5], an anonymous and distributed electronic cash system, In that method, describes two protocols based on network that cannot be traced, where senders and receivers are identified only by digital such as their public keys, and each message will be signed by its sender to receiver.

Bit Gold In 1998, Nick Szabo[6] propose models a new digital currency, the models based on cryptographic system puzzles, which after being solved, were sent to the Byzantine-fault-tolerant public registry and assigned to the public key of the solver.

Hashcash proposed by Adam Back, Haschash, a system relied on a cryptographic hash function to derive a probabilistic proof of computational work as authentification system Pow (Proof of work)[7]

And the last is RPOW Hal finney propose currency system based on a reusable proof of work (RPOW) in 2004[8].

Between 2008 and 2009, Bitcoin was made as to the first decentralized cryptocurrency by Satoshi Nakamoto. Nakamoto published the Bitcoin whitepaper in 2008[9], and after January 3rd, 2009, the genesis block of the bitcoin protocol was created. Nowadays it is most successful cryptocurrency in terms of market capitalization, beside above 700 altcoins that circulated in the world (eg. Litcoin, Etherum) based on Bitcoin have been proposed since the launch of Bitcoin.

B. Bitcoin Price Predictions Methods

There are many people doing research about the prediction of cryptocurrency. Greaves et al. [10] is a proposed technique using Logistic Regression and SVM and analyzed using Graph to predict bitcoin price. Huisu Jang et al. [11] they concern about a study on modeling and prediction bitcoin with Bayesian Neural Network and giving some knowledge about bitcoin. Edwin sin et al. [12] provide topic Bitcoin price prediction using Ensemble of Neural. Networks. Arief Radityo et al.[2] proposed a prediction of bitcoin using Artificial Neural Network Technique. They combine with market technical indicators but the results are worse of performance and training time

Papers using LSTM, Sean et al [13] They propose method the price of Bitcoin using RNN and combine Using Recurrent Neural Network and Long Short Term Memory and Ruchi Mittal et al [14] is propose an Automated cryptocurrencies prices prediction using machine learning technique based on the historical trend (daily trend). Chih-Hung et al. [15] are created new forecasting framework bitcoin price using LSTM, they proposed with two various LSTM models (conventional LSTM and LSTM with AR(2) model) with 208 records dataset, compared with MSE, RMSE, MAE, and MAPE. Fei Qian et al.[16] produced a common stock market prediction model based on LSTM under Different factor that impacts the market, in this research their selected three stock with similar trends. The LSTM prediction model is performed well.

The researches above proposed various method to prediction bitcoin. In this paper, we analyze and constructing a model to predict bitcoin using LSTM

C. Bitcoin price predictions Technique

Technically Bitcoin stock market prediction it's the same with prediction technique on the common stock exchange but in the other way, when you try another technique and strategy like sentiment analysis maybe we can't get different results orthe strategy it does not work. Because beside many factors that impact on stock exchange prediction, Bitcoin it's decentralized and not regulated by any party so it's different from common currency or the common stock market.

We can use the same algorithm to Bitcoin prediction using Machine Learning (eq. SVM, Naïve Bayes, Regression) [10][11] or any other Advance Machine learning technique to improve the results like Deep Learning using Neural Network (eq. ANN and RNN)[2]

On Prediction, we can predict Bitcoin using technique on the specific subject that we wanted. Example, we want to predict only by the signal or the price, or we can predict just for current day or next day close value based on Long Short Term Memory (LSTM),[14] historical price and other technique like regime prediction to detect current day's trend on market, to help investor to make decision to investment.

To make more accurate and enrich the result we can combine the prediction algorithm with another method or technique on prediction Bitcoin

In this section, we describe some of the technique that mentions previous research on some papers.

Market technical analysis is a method that studies price movements by looking at historical price data that occur in the market through media charts. By studying this historical data a conclusion can be drawn for making investment decisions in the market.

There are several reasons why we should use TA, the first, not necessarily fundamental analysis can be applied in trading. I do not say Fundamental is not important, but here that needs to be underlined fundamental analysis will be far more important if our position as an Investor where the investment period can take more than one year.

Secondly, with the media chart, it will be seen how the journey of a stock price where it is very helpful for traders to analyze in anticipation of future price changes and see patterns of patterns that occur in the price movements of a stock, so traders do not trade in ' darkness or without clear direction.[17]

Another technique to interpretation trend is **Technical Indicator** [18]is a series of data points obtained by applying the formula to securities price of data. The combinations of price data, such as close, low, high, low and open can be used as data point certain period of time.

Time series Data Analysis in the terms of future price predictions popular methods is using Autoregressive integrated moving average (ARIMA) models are a popular choice for forecasting over a short term condition, it works when data exhibits consistent or stable pattern (constant) over time with least possible outliers. But this does not work always in the real-time scenario, where data fluctuated drastically and it is highly volatile.[19]

Trading Strategy, this strategy is that we maintain the position of +1 Bitcoin, 0 Bitcoin or −1 Bitcoin. [20]

III. METHODOLOGY

LSTM (Long Short Term Memory) is another type of module provided for RNN. LSTM was created by Hochreiter & Schmidhuber (1997)[3] and later developed and popularized by many researchers. Like RNN, the LSTM network (LSTM network) also consists of modules with

recurrent consistency.

LSTM is an updated version from RNN, the difference is the connection between the hidden layers of RNN. The explanation structure of RNN is shown in Figure 1. RNN & LSTM have a similar structure, the other different is memory cell of structure hidden layer. And the design of three special gates effectively solve the gradient problems. The LSTM memory structure of hidden layer shown in Figure 2. [16]
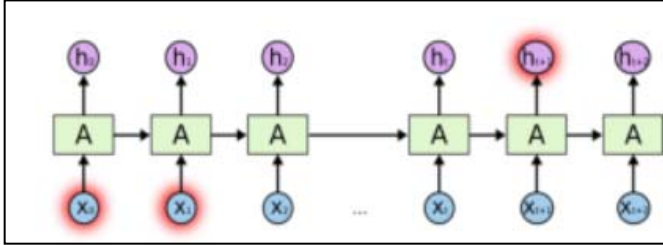

Figure 1. The Expanded Structure of RNN[21]

In Figure 1 explains the RNN has shortcomings, the shortcomings can be seen in the input X_o, X_1 has a very large range of information X_t, X_{t+1} so that when $_{t+1}$ requires information those that are relevant X_o, X_1 to RNN are unable to learn to link information because of old memory saved will be increasingly useless as time goes by because it is overwritten or replaced with new memory, this problem was discovered by Bengio, et al. (1994)[22] .

Unlike the RNN, LSTM does not have the disadvantage is that LSTM can manage the memory at each input by using memory cells and gate units.
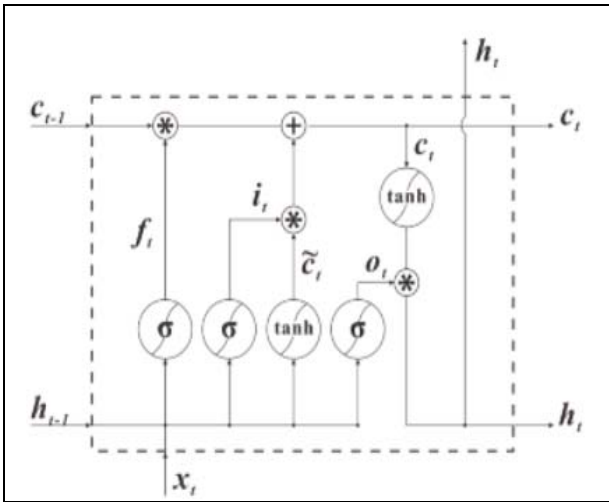

Figure 2. LSTM memory cell structure of the hidden layer[16]

- i_i = the input gate, it's means information will be updated in the cell
- f_t = the forget gate, it's means information should be dropped from the cell.
- O_t = the output gate, it means how much information is output

- $\tilde{c}_t$ = the candidate value for the states of the memory cell at time t.
- c_t = the state of the current memory cell at time t, which calculated by the combination of i_{t} and $\tilde{c}_t$ f_t and c_{t-1} through element-wise multiapplication
- h_t = it's mean output value filtered by output gate
-
- σ = is denoted denotes sigmoid function with the range 0 to 1, the function is used to put the value between -1 and 1.
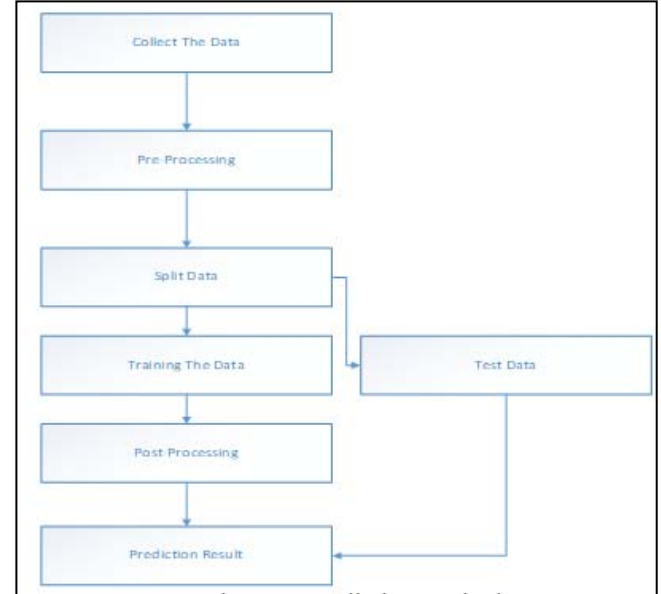
A. Propose Method


Figure 3. Prediction Method

Based on Figure 3 The process start from collect the data, dataset is collected from yahoo finance stock market based on the USD Exchange rate and collected from CCC - CryptoCompare. Currency in USD, with 5 years period from 27-06-2014 until 27-06-2019, this is historical data prices, so in this research used time-series data on this experimental, with 1829 number of datasets, on CSV format. Figures 3 show the sample data of datasets.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2014-06-27	602.229980	608.760010	593.969971	597.559998	597.559998	1778752
1	2014-06-28	597.559998	604.280029	593.190002	601.059998	601.059998	1736172
2	2014-06-29	601.059998	637.280029	596.549988	635.140015	635.140015	2394861
3	2014-06-30	635.140015	652.539978	634.090027	643.190002	643.190002	2675353
4	2014-07-01	643.190002	652.099976	604.580017	645.719971	645.719971	3099761

Figures 4. Sample data

IV. EXPERIMENTAL AND RESULTS

Figures 4 below shown pre-processing result to load the dataset into machine and algorithm, and then shown the last day close price data of bitcoin before we train and test and predict the results.
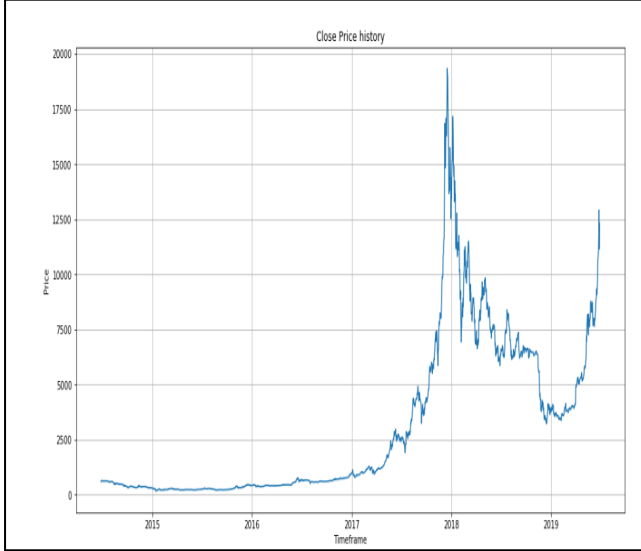
Figures 5. Close price History data bitcoin

For Step Split and training the data, we divided 4 years period of data for training and 1 year for testing. Split at 1462 data in. While splitting the data into train and validation, we cannot use random splitting since that will destroy the time component[23]. So we set the last year's data into test and the 4 years' data before that into train the dataset.

A. Scenario analysis results using RMSE

Before finalizing the results we try to Measure the results using RMSE, RMSE is (the Root Mean Square Error). The RMSE will always be larger or equal to the MAE. The RMSE metric evaluates how well a model can predict a continuous value. The RMSE units are the same units as your data's dependent variable/target (so if that's dollars, this is in dollars), which is useful for understanding whether the size of the error is meaningful or not. The smaller the RMSE, the better the model's performance.[24]

This is the formula $\textbf{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2}$ [25]

Where N is the total number of observations, x_i is the actual value; whereas, $\hat{x}_i$ is the predicted value. The main benefit using RMSE is that it penalizes large errors. It also scales the scores in the same units as the forecast values).

We combine the epoch and model dropout to search the best results.The number of epoch we used: 10, 100, 1000, 200, 400, 800, 2000, and 5000. With the number of model dropout we combined with 0, 0,1 and 0,5.

No	Epoch	Model Dropout	RMSE Results
1	10	0	631.74963
2	100	0	455.98107
3	1000	0	825.37505
4	200	0	360.64511
5	400	0	354.18368
6	500	0	288.59866
7	800	0	292.78967
8	2000	0	477.91428
9	5000	0	474930575
10	500	0,1	602.140637
11	500	0,5	313.6623

Table 1. RMSE Results

Based on Table 1 The smaller results is using 500 epoch with model dropout 0. With RMSE results 288.59866. as RMSE mention on above, The smaller the RMSE, the better the model's performance, after that we use to produce the results.
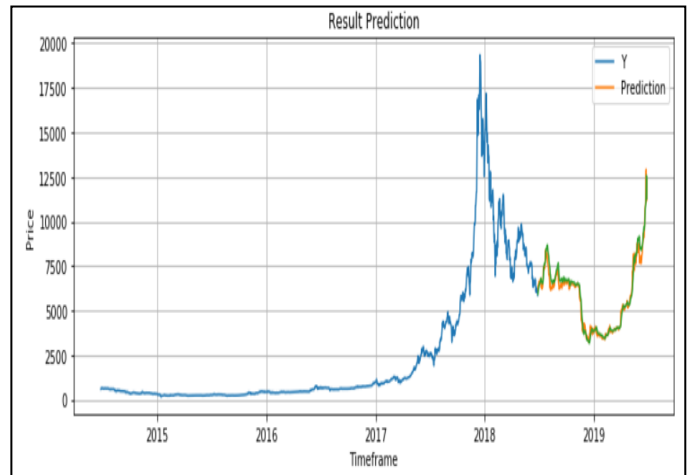
Figures 6. Prediction result price of bitcoin

Based on figures 6 that result from prediction using LSTM shown by the graph with epoch 500, model dropout 0, and, Yellow Line is a result for the close prediction, the Blue and Green line is from Data Training. The price result is above $12600 for next days based on the model.

V. CONCLUSION AND FUTURE WORKS

Our Proposed model has been succeeded to provide the result prediction bitcoin from yahoo finance stock market. Our model with time series techniques can build produce the results and the results can predict the price for the next days with split the data to train and test that we mention in the article above. But the disadvantage is the result it's not good enough regarding the RMSE, maybe under the hundreds or

near close to 50 Score RMSE. Afterward, as we mentioned before in the article, the stock market is influenced by many uncertainty factors. The Stock markets are influenced by many uncertainties factor such as political issue, the economic issue at impacted to local or global levels. So prediction price bitcoin using LSTM can't good enough to make the decision to invest in bitcoin, it is another side for taking the decisions.

Future research will focus on, modified LSTM layers, adding dropout and modified number of epochs, and using different instability dataset to test how good the prediction results or try to use sentiment analysis combined with LSTM method to see the impact of the uncertainty in value bitcoin.

REFERENCES

[1] B. U. (BTC-USD) and C.-C. C. in USD, "Yahoo Finance," 2019. [Online]. Available: https://finance.yahoo.com/quote/BTC-USD?p=BTC-USD. [Accessed: 14-Mar-2019].

[2] A. Radityo, Q. Munajat, and I. Budi, "Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods," in *Advanced Computer Science and Information Systems (ICACSIS), 2017 International Conference on*, 2017, pp. 433–438.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[4] A. Judmayer, N. Stifter, K. Krombholz, and E. Weippl, "Blocks and Chains: Introduction to Bitcoin, Cryptocurrencies, and Their Consensus Mechanisms," *Synth. Lect. Inf. Secur. Privacy, Trust*, 2017.

[5] W. Dai, "B-money proposal," *White Pap.*, 1998.

[6] N. Szabo, "Bit gold," *Website/Blog*, 2008.

[7] A. Back, "Hashcash." 1997.

[8] Hal Finney, "RPOW - Reusable Proofs of Work," *Agosto 15, 2004*. 2004.

[9] S. Nakamoto and others, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[10] A. Greaves and B. Au, "Using the bitcoin transaction graph to predict the price of bitcoin," *No Data*, 2015.

[11] H. Jang and J. Lee, "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information," *IEEE ACCESS*, vol. 6, pp. 5427–5437, 2018.

[12] E. Sin and L. Wang, "Bitcoin Price Prediction Using Ensembles of Neural Networks," in *2017 13TH INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION, FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY (ICNC-FSKD)*, 2017,

pp. 666–671.

[13] S. McNally, J. Roche, and S. Caton, "Predicting the price of Bitcoin using Machine Learning," in *Parallel, Distributed and Network-based Processing (PDP), 2018 26th Euromicro International Conference on*, 2018, pp. 339–343.

[14] R. Mittal, S. Arora, and M. P. S. Bhatia, "AUTOMATED CRYPTOCURRENCIES PRICES PREDICTION USING MACHINE LEARNING," 2018.

[15] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, "A New Forecasting Framework for Bitcoin Price with LSTM," in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2018, pp. 168–175.

[16] F. Qian and X. Chen, "Stock Prediction Based on LSTM under Different Stability," in *2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2019, pp. 483–486.

[17] J. J. Murphy, "Technical Analysis Of The Financial Markets," *Pennsylvania Dental Journal*. 1999.

[18] E. Kristensen, S. Østergaard, M. A. Krogh, and C. Enevoldsen, "Technical Indicators of Financial Performance in the Dairy Herd," *J. Dairy Sci.*, 2008.

[19] C. Scheier and W. Tschacher, "Appropriate algorithms for nonlinear time series analysis in psychology," in *Nonlinear dynamics in human behavior*, World Scientific, 1996, pp. 27–43.

[20] D. Shah and K. Zhang, "Bayesian regression and Bitcoin," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014, pp. 409–414.

[21] M. W. P. Aldi, J. Jondri, and A. Aditsania, "Analisis Dan Implementasi Long Short Term Memory Neural Network Untuk Prediksi Harga Bitcoin," *eProceedings Eng.*, vol. 5, no. 2, 2018.

[22] Y. Bengio, P. Simard, and P. Frasconi, "Learning Long-Term Dependencies with Gradient Descent is Difficult," *IEEE Trans. Neural Networks*, 1994.

[23] AISHWARYA SINGH, "Stock Prices Prediction Using Machine Learning and Deep Learning Techniques (with Python codes)," *OCTOBER 25, 2018*, 2018. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price-machine-learningnd-deep-learning-techniques-python/. [Accessed: 20-Jun-2019].

[24] Squark, "ROOT MEAN SQUARE ERROR OR RMSE." [Online]. Available: https://squarkai.com/root-mean-square-error-or-rmse/#.XWVCLegzZPY. [Accessed: 28-Aug-2019].

[25] J. Brownlee, "Time series prediction with lstm recurrent neural networks in python with keras," *Available Mach. com*, p. 18, 2016.