Price Prediction of Bitcoin Using LSTM Neural Network

Qilan Lin *
King's School of Professional and Continuing Education, King's College London, London, the UK.
*Corresponding author: qilan.lin@kcl.ac.uk

Abstract. Contemporarily, cryptocurrency has a high market value, and the price of cryptocurrency fluctuates dramatically. This article analyzes the parameters effects of the LSTM model on Bitcoin price prediction accuracy based on Python and modules of Numpy, Pandas, Keras, Tensorflow, and Sklern. The analysis clarifies the relationship between the accuracy of Bitcoin price prediction and different parameters in the LSTM model. It is discovered that when larger batch sizes are supplied at minor epochs, the accuracy of Bitcoin price prediction declines. Meanwhile, the number of neurons affects the accuracy. In addition, compared to lengths of 14, 30, and 60, the prediction error grows greater when a single time sequence is 7 in length. Apart from that, at present, using closing prices from the past two years rather than the past 1 year, 3 years, or 5 years can make predictions more accurate. These findings shed light on recommendations for adjusting various parameters in the development of the LSTM model for Bitcoin price prediction.

Keywords: Bitcoin; Price prediction; LSTM; Neural network.

1. Introduction

Bitcoin was invented by a man anonymized as "Satoshi Nakamoto", as "a purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution", as is written in its original white paper [1]. Bitcoin's value is inspired by properties, e.g., the money supply for Bitcoin is set and predictable. Unlike fiat currency, Bitcoin cannot be abruptly minted in vast quantities by any elected or unelected official. The final total circulation of Bitcoin is 21 million when all Bitcoins are mined. Unlike fiat money, bitcoin uses open-source software and is totally transparent. Anyone at any time can independently verify the total supply of Bitcoin, its underlying principle, and the balances of each address (account) on the decentralized global ledger. Cutting-edge cryptography and enormous energy support the security of Bitcoin. It would cost an amount of energy, specially designed computing machines, and space that any individual or organization cannot possess or utilize to undermine the core encryption of Bitcoin. Bitcoin is the world's most secure decentralized computing network [2, 3]. Unlike conventional bank accounts, the Bitcoin network is accessible to anybody whenever and wherever. Bitcoin is an uncensorable and global network for transacting value [4].

Bitcoin has become a hot investment target these years, be it out of its technical attributes of decentralized blockchain and cryptography or being hyped up. As Bitcoin gained prominence, investors and speculators became interested in it. From 2009 to 2017, cryptocurrency exchanges that facilitated bitcoin sales and purchases emerged. The price of Bitcoin started to increase, and demand climbed gradually until 2017, when the prices broke the $1,000 barrier. Many investors believed Bitcoin prices would keep rising and started buying them to hold. The market took off when traders began using cryptocurrency exchanges to make short-term trades [5]. Speculators have been drawn to Bitcoin after its prices have been rapidly acknowledged in recent years. According to Binance, Bitcoin had a price of $7,168.84 on 31st Dec. 2019, and 1 year later, it had appreciated more than 300% to $28,969.01. It continued to surge, trading at a record high of over $69,000 in November 2022 [6]. The COVID-19 pandemic caused Bitcoin to undergo an unprecedented bull market as a safe-haven asset amid persistent geopolitical concerns around the globe [7-9]. Since then, many traditional financial institutions and listed companies, e.g., Rothschild & co, Microstrategy, Tesla, and several national authorities, e.g., El Salvador, added cryptocurrencies into their portfolios [10-12].

In order to analyze the influence on Bitcoin price forecast accuracy of different parameters of the LSTM model, this paper will use the variable-controlling method to execute a series of repeated experiments. In the remaining parts, this article first introduces recurrent neural networks and long short-term memory in Sec. 2. Sec. 3 will tell the data source, the method to clean the data, and the procedures to research it. Sec. 4 presents the result and makes a discussion vis-à-vis it, together with possible explanations and suggestions on applying these results. Then, this article discusses the limitation and conclusions of the study.

2. Statistical Principle

A recurrent neural network (i.e., RNN) is a class of recursive neural networks which use sequence data as its input source, proceeding recursion in the direction the sequence proceeds, and where all recurrent units link up in the form of a chain [13]. The most significant difference between an RNN and a traditional neural network is that each time an RNN will bring the previous output results to the next hidden layer to be trained with the subsequent input.

A typical sketch of the compressed and unfolded basic RNN is given in Fig. 1. A problem of gradient vanishing exists in traditional RNN architecture, which obstructs the study of long-distance dependency in the training process [14]. Long short-term memory (i.e., LSTM), which is a particular type of RNN, was first put forward by Hochreiter and Schmidhuber. It introduces a forget gate, an input gate, and an output gate, studying long-distance dependency by the "gating" mechanism to reduce the problem of gradient vanishing.

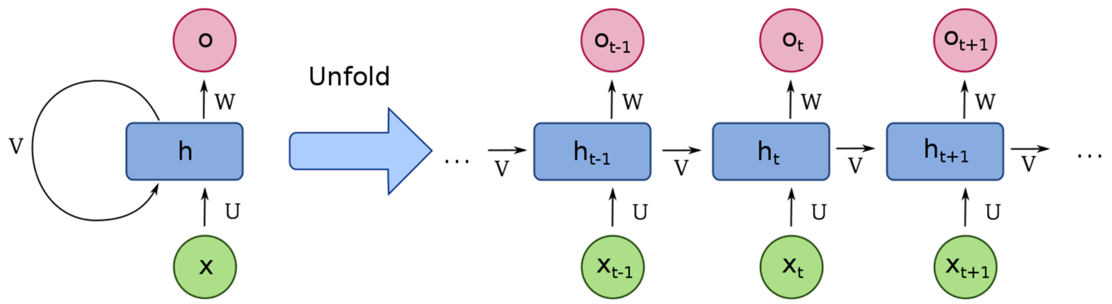

Fig. 1 Compressed (left) and Unfolded (right) Basic Recurrent Neural Network

LSTM performs outstandingly in many domains [15], where a sketch is shown in Fig. 2. Hochreiter et al. utilized LSTM to create a fast protein homology detection in which alignment is not needed [16]. In 2016, Google released a novel English-Chinese translation system, Powered by LSTM, which reduces error rates by 60 per cent [17]. In 2019, AlphaStar, an AI of DeepMind, used LSTM trained by policy gradients to sweep Starcraft II players [18]. Numeral value prediction from former time series is a popular LSTM-based research topic. Chen et al. used LSTM models to predict the return of China stocks, improving accuracy from 14.3% to 27.2% compared to the random prediction method [19]. Nelson et al. built an LSTM-based model and executed a series of experiments. The result showed an average of 55.9% accuracy, which is promising [20].

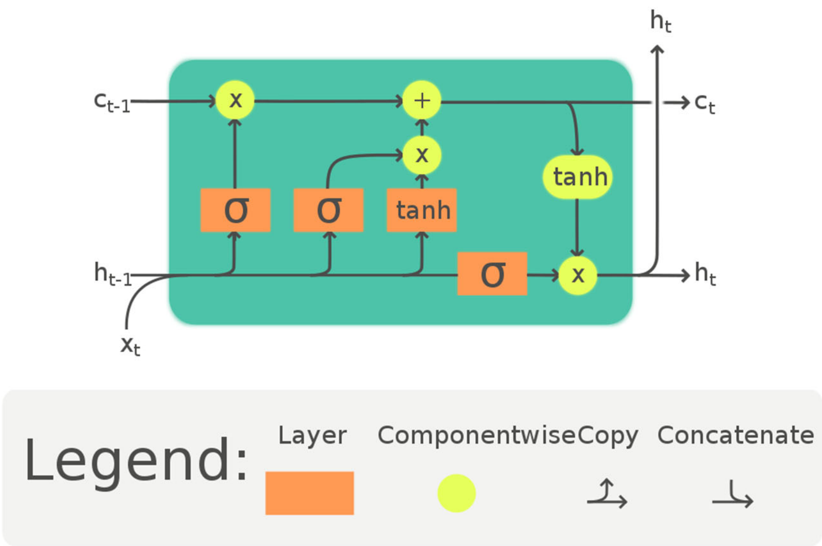

Fig. 2 LSTM Cell

3. Data & Method

Based on Python and modules of Numpy, Pandas, Keras, Tensorflow and Sklearn to establish LSTM models, this research is aimed at exploring the influence on Bitcoin price forecast accuracy of different parameters of the LSTM model. This study gains the historical prices in the form of comma-separated values (.csv) of Bitcoin on Yahoo Finance, traced every day from 18th Aug. 2017 to 17th Aug. 2022. The price trend of Bitcoin is illustrated in Fig. 3.

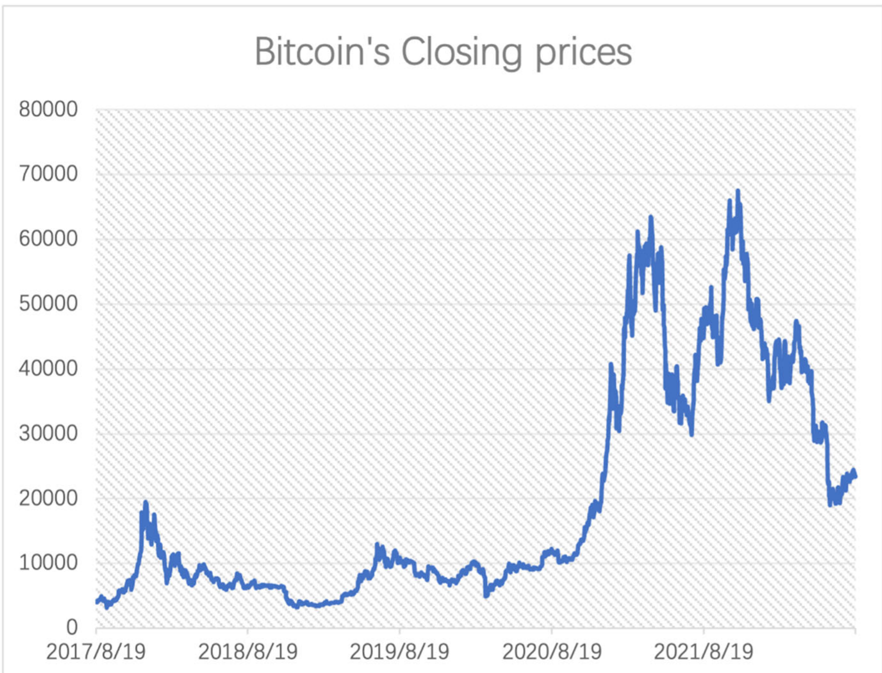

Fig. 3 Bitcoin's Closing Prices

First, the historical price data of Bitcoin are inputted in the Python code, and then the "MinMaxScaler" method from Sklearn is used to scale all the historical prices in the range of 0 and 1. Subsequently, the percentage of data is chosen for training among all the historical price data and the length of a single time sequence for training. The lengths of a single time sequence of 60, 30, 14, and 7 are tested, which are mainstream days used for moving averages, a widely used standard in various financial markets. Afterwards, the number of neurons in each layer is decided. The units of neurons in each layer in a neural network model determine the model's sophistication. The number of epochs and batch size are critical parameters that also make a difference in the model. Experimental groups are divided by the following parameters: dates of Bitcoin trading, percentage of data for training, the length of a single time sequence, the number of neurons, batch size, and the number of epochs. For each experimental group, eight repeated experiments are done. Statistics are rounded to the nearest hundredth. Average RMSEs are calculated after removing the maximum and minimum values in each experimental group. The models each contain two LSTM layers and two dense layers and are compiled by loss function 'mean_squared_error' and optimizer 'adam'. Root mean squared error (RMSE) is a measure of the difference in value between the data predicted and the data observed. It can be used as an indicator to reflect the accuracy of the model. The lower the numeric value of RMSE indicates the better performance of the model. A numeric value of 0 expresses that the predicted value of the model wholly matches the actual value in the test dataset.

4. Results & Discussion

Under the circumstance that for each experimental group, the percentages of data for training are all 70%, the length of a single time sequence is 60, the first and second LSTM layers both consist of 50 neurons, the first dense layer has 25 neurons, and the second dense layer has 1 neuron. The results of the models are presented in Table 1. When the batch size is 10, and different values are given to the batch sizes (1, 10, and 100; which are 10 to 0th, 1st, and 2nd), the average RMSEs grow from 2330.27 to 2739.74, then surge to 5140.05 as batch size rise exponentially. Nevertheless, when the batch size becomes 50, the average RMSEs decrease from 2426.68 to 1841.50 first and then bounce back to 2901.80. In this case, it can be concluded that the accuracy of Bitcoin price prediction drop as larger batch sizes are given at minor epochs. That may be attributed to the fact that some values with deviation in smaller batches help improve the adaptability of the model when offered in the training process; the training time of each epoch is shortened as well, which means the update of weight coefficients in the model will be quicker.

When the batch sizes are all 10, the epochs are all 50 for each experimental group, and the neuron units in the first and second layers are modified to 32 and 32, respectively. The numbers of neurons do not change in the first and second dense layers, and the average RMSE increases a little from 1841.50 to 1970.77. Then, when the numbers of neurons in the first and second layers are changed from 25,1 to 16,1, the average RMSE falls from 1970.77 to 1907.04. At the time that all batch sizes are 10, all epochs are 50 for each experimental group, and the percentages of data for training all changed from 70% to 90%, the average RMSEs all go down, whether the dates of Bitcoin trading are 5 years, 3 years, 2 years, and 1 year. It is also found that the average RMSEs of group 5 y, group 3 y, and group 2 y declined by 32.78%, 27.65%, and 22.70%, respectively, which seems to indicate that the longer the dates of Bitcoin trading are, the more significant the decrements of the average RMSEs are. However, in group 1 y, when the percentage of data for training is 70%, its prediction error is far higher than other groups; the average RMSE plunges by 69.08%, to be the lowest among all groups when the percentages of data for training turns to 90%.

As mentioned above, the length of a single time sequence of 7, 14, 30, and 60 are typical days used for moving averages when trading financial indices. In the case where batch sizes are all 10, the epochs are all 50 for each experimental group of 5 y. Besides, the percentages of data for training are all 70%, and the length of a single time sequence is adjusted to 7, 14, 30, and 60. It is found that if a single time sequence input consists of 7 closing prices, the prediction error increases more compared to the length of 14, 30, and 60.

In the condition that batch sizes are all 10, the epochs are all 50, and the percentages of data for training are all 70% for each experimental group, whether the length of a single time sequence is 7, 14, 30, or 60, the RMSEs all first be on a downward trend when the dates for Bitcoin training vary from 5 years to 3 years, and 2 years, but when the dates vary from 2 years to 1 year, the RMSEs soar to a number more significant than that of group 5 y.

Table 1. The Prediction Results of the Models

Dates of Bitcoin Trading	Percentage of Data for Training	Length of a Single Time Sequence	Number of neurons in Each Layer	Batch Size	Number of Epochs	Average RMSE
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	1	50	2426.68
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1841.50
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	100	50	2901.80
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	1	10	2330.27
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	10	2739.74
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	100	10	5140.05
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=32, LSTM L2=32, Dense L1=25, Dense L2=1	10	50	1970.77
5 y (19th Aug. 2017~18th Aug 2022)	70%	60	LSTM L1=32, LSTM L2=32, Dense L1=16, Dense L2=1	10	50	1907.04
5 y (19th Aug. 2017~18th Aug 2022)	90%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1237.94
3 y (19th Aug. 2019~18th Aug 2022)	90%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1146.55
2 y (19th Aug. 2020~18th Aug 2022)	90%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1091.92
3 y (19th Aug. 2019~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1584.82
2 y (19th Aug. 2020~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1412.63
1 y (19th Aug. 2021~18th Aug 2022)	90%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	980.00
1 y (19th Aug. 2021~18th Aug 2022)	70%	60	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	3169.88
5 y (19th Aug. 2017~18th Aug 2022)	70%	30	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1959.09
3 y (19th Aug. 2019~18th Aug 2022)	70%	30	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1657.37
2 y (19th Aug. 2020~18th Aug 2022)	70%	30	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1524.70
1 y (19th Aug. 2021~18th Aug 2022)	70%	30	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	2521.61
5 y (19th Aug. 2017~18th Aug 2022)	70%	14	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1821.94
3 y (19th Aug. 2019~18th Aug 2022)	70%	14	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1629.88
2 y (19th Aug. 2020~18th Aug 2022)	70%	14	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1437.74
1 y (19th Aug. 2021~18th Aug 2022)	70%	14	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	2679.33
5 y (19th Aug. 2017~18th Aug 2022)	70%	7	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	2187.88
3 y (19th Aug. 2019~18th Aug 2022)	70%	7	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1595.92
2 y (19th Aug. 2020~18th Aug 2022)	70%	7	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	1398.16
1 y (19th Aug. 2021~18th Aug 2022)	70%	7	LSTM L1=50, LSTM L2=50, Dense L1=25, Dense L2=1	10	50	2869.25

5. Limitations & Prospects

Although LSTM, an improved RNN model, is used in this study and has solved the problem of gradient disappearance to some extent, LSTM will still be troubled by gradient explosion. This paper has limited changes in variables. In terms of the financial index, the derivatives of Bitcoin can also be investigated. Adding a regularized penalty term in the loss function, adding dropouts, setting learning rates and other methods introduce new variables for the experiments, from which one may find new relationships between the accuracy of prediction and these new variables. In terms of data processing, introducing Bitcoin trading volume, converting the closing prices into MA line, ball line, RSI, KDJ and other traditional financial indicator values or introducing the Hash rate of Bitcoin mining difficulty for model training may help explore the difference between its accuracy and that of using the closing prices only to train the model.

6. Conclusion

In summary, this study discusses the feasibility of the price prediction for bitcoin based on LSTM neural network scenarios. Using the variable-controlling method to investigate the impact on price forecast accuracy of different parameters of the LSTM model, the accuracy of Bitcoin price prediction drop as larger batch sizes are given at minor epochs. The accuracy of the forecast is relevant to the number of neurons. To be specific, when changing the percentage of data for training from 70% to 90%, the accuracy goes up regardless of the years of Bitcoin trading. According to the analysis, when a single time sequence input is only 7 closing prices, the prediction error increases more compared to the length of 14, 30, and 60. According to the results, when the dates for Bitcoin trading is 2 y in LSTM model training, the accuracy of the forecast is higher than the dates of 1 year, 3 years or 5 years. In this case, when establishing an LSTM model for Bitcoin price prediction now, the accuracy is higher when using closing prices from the past 2 years than from the past 1 year, 3 years, or 5 years.

Nevertheless, the parameters experimented on are still limited. In the future, more parameters could be investigated, which might help find more influence on the accuracy of Bitcoin price prediction. All the findings make suggestions for the adjustment of different kinds of parameters in the establishment of the LSTM model for Bitcoin price prediction.

References

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

[2] Vranken H. Sustainability of bitcoin and blockchains. Current opinion in environmental sustainability, 2017, 28: 1–9.

[3] Nadarajah S, Chu J. On the inefficiency of Bitcoin. Economics Letters, 2017, 150: 6–9.

[4] Böhme R, Christin N, Edelman B, et al. Bitcoin: Economics, technology, and governance. Journal of economic Perspectives, 2015, 29(2): 213-38.

[5] Bitcoin. Investopedia. Retrieved from: https://www.investopedia.com/terms/b/bitcoin.asp.

[6] Binance. Bitcoin. Retrieved from: https://www.binance.com/zh-CN/trade/BTC_BUSD

[7] Huang Y, Duan K, Mishra T. Is Bitcoin really more than a diversifier? A pre-and post-COVID-19 analysis. Finance Research Letters, 2021, 43: 102016.

[8] Chen C, Liu L, Zhao N. Fear sentiment, uncertainty, and bitcoin price dynamics: The case of COVID-19. Emerging Markets Finance and Trade, 2020, 56(10): 2298–2309.

[9] Sarkodie S A, Ahmed M Y, Owusu P A. COVID-19 pandemic improves market signals of cryptocurrencies–evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin. Finance Research Letters, 2022, 44: 102049.

[10] One Of The Biggest Ever Bets On Bitcoin, Worth A Massive $250 Million, Is Already Paying Off. Forbes. Retrieved from: https://www.forbes.com/sites/billybambrough/2020/08/13/one-of-the-biggest-ever-bets-on-bitcoin-worth-a-massive-250-million-is-already-paying-off

[11] Rothschild Investment Corp More Than Tripled Its Bitcoin Position. Nasdaq. Retrieved from: https://www.nasdaq.com/articles/rothschild-investment-corp-more-than-tripled-its-bitcoin-position-2021-07-20.

[12] El Salvador, primer país del mundo en reconocer al Bitcoin como moneda de curso legal. Asamblea. Retrieved from: https://www.asamblea.gob.sv/node/11282.

[13] Goodfellow I, Bengio Y, Courville A. Deep learning (Vol. 1). Cambridge, MIT Press, 2016, pp. 367–415.

[14] Recurrent neural network unfold. Wikimedia. Retrieved from: https://commons.wikimedia.org/wiki/File:Recurrent_neural_network_unfold.svg.

[15] LSTM Cell. Wikimedia. Retrieved from: https://commons.wikimedia.org/wiki/File:LSTM_Cell.svg.

[16] Hochreiter S, Heusel M, Obermayer K. Fast model-based protein homology detection without alignment. Bioinformatics, 2007, 23(14): 1728–1736.

[17] Google claims AI breakthrough machine translation. Wired. Retrieved from: https://www.wired.com/2016/09/google-claims-ai-breakthrough-machine-translation/

[18] Deepminds AI alphastar showcases significant progress towards Agi. Medium. Retrieved from: https://medium.com/@towardai/deepminds-ai-alphastar-showcases-significant-progress-towards-agi-93810c94fbe9

[19] Chen K, Zhou Y, Dai F. A LSTM-based method for stock returns prediction: A case study of China stock market. 2015 IEEE international conference on big data (big data). IEEE, 2015: 2823–2824.

[20] Nelson D M Q, Pereira A C M, de Oliveira R A. Stock market's price movement prediction with LSTM neural networks. 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1419–1426.