

BITCOIN PRICE PREDICTION USING LSTM

Sakshi Patil* and Gaurav Patil†

Department of Computer Engineering, Sardar Patel Institute Of Technology
Mumbai, India

Email: *sakshi.patil@spit.ac.in, †gaurav.patil@spit.ac.in

Abstract—Bitcoin is a sort of cryptocurrency that has become a popular stock market investment. The stock market is influenced by a variety of risk factors. And bitcoin is one type of cryptocurrency that has been steadily rising in recent years, with occasional sharp drops without any apparent impact on the stock market. Because of the volatility, there is a demand for an automated technique to forecast bitcoin on the stock market. LSTM (Long Short Term Memory) is another form of module supplied for RNN that was later developed and popularized by many researchers, and like RNN, the LSTM also consists of modules with recurrent consistency. The contribution of this research is the development of a new forecasting framework for bitcoin price prediction that can solve and improve the problem of input variable selection in LSTM without relying on stringent data assumptions. The method we used in this study, as well as the techniques and tools we used to predict Bitcoin on the stock market through Yahoo Finance, may predict a result of more than 12600 USD in the days after the prediction. In the final section, we draw conclusions and provide references.

Index Terms—Bitcoin, Cryptocurrency, Forecasting, LSTM, RNN, Yahoo Finance

I. INTRODUCTION

Cryptocurrencies are a type of digital currency that allows for online transactions. Unlike traditional currencies, cryptocurrency is built on encryption. Bitcoin is a type of cryptocurrency that is unregulated and decentralized. Bitcoin's one-of-a-kind feature is its daily price changes, which alter on a regular basis. Many uncertainties influence stock markets, such as political and economic issues that have an impact on local and global levels. It is a difficult task to interpret the secret to success, factor, and provide correct predictions. We can use any technique to study the market, such as technical indicators, price movements, and market technical analysis. To address the aforementioned issue, an automation tool for prediction is required to assist investors in deciding whether to participate in bitcoin or other cryptocurrency markets. Automation technologies are commonly utilized in conventional stock market predictions these days, and we can do the same job and technique in the cryptocurrency arena. Cryptocurrency is based on Blockchain, a digital ledger technology. With the installation of an access control mechanism, blockchain systems can maintain the privacy and security of user data[3]. A Blockchain is a decentralized linked data structure that is characterized by its inherent resistance to data tampering, but it suffers from poor data formatting, which makes it unsuitable for search queries. M Muzammal, et al.(2019) [1] present ChainSQL, an open source system developed by combining the Blockchain with the database, i.e. they present

a Blockchain database application platform that combines the Blockchain's decentralized, distributed, and audibility features with the distributed databases' quick query processing and well-designed data structure.

Another sort of RNN module is the LSTM (Long Short Term Memory). Hochreiter & Schmidhuber (1997) invented LSTM, which was further developed and popularized by a number of researchers. The LSTM is made up of recurrently consistent modules, just as the RNN. This article examines bitcoin and stock market predictions, methods, techniques, and tools using a variety of materials, including books, papers, and other publicly available sources.

II. LITERATURE SURVEY

Under the several factors that affect the market, Fei Qian et al. [2] developed a common stock market prediction model based on LSTM. In this study, they selected three stocks with comparable tendencies. The performance of the LSTM prediction model is excellent. The authors of [4] use the conventional Support Vector Machine and Linear Regression methods to present the Bitcoin prediction. This project considers taken into account, the time series prediction was made up of Bitcoin closing prices on a daily basis. The authors of [5] compare and contrast the existing Theil - Sen and Huber regression methods. In addition, the GRU and LSTM are compared to the above-mentioned machine learning approaches in this study. The forecast of Bitcoin process utilising empirical study is carried out in a recent research effort [6]. This study compares the Bayesian neural network to existing non-linear and linear benchmark methods, yielding useful results. The authors of [7] used machine learning techniques to create models that can anticipate Bitcoin prices. For Bitcoin price prediction, the authors presented both a multiple regression strategy that relies on highly correlated characteristics and a deep learning mechanism that uses a conjugate gradient mechanism in conjunction with linear search. The authors of [8] found daily Bitcoin market trends using the semantics and insights connected with the attributes that influence Bitcoin price. The writers have also anticipated the daily price fluctuations of Bitcoin. In the study given by these authors, Random Forest and Bayesian regression were used. Many people are conducting research into cryptocurrency forecasting. Greaves et al. [9] suggested a technique for predicting bitcoin price using Logistic Regression and SVM, which was examined using Graph. Huisu Jang et al. [10] are concerned with a study on modeling and prediction of bitcoin using Bayesian Neural

Networks, as well as providing some bitcoin expertise. Edwin Sin and his associates [11] discuss how to anticipate Bitcoin prices using an ensemble of neural networks. Arief Radityo et al.

[12] suggested an Artificial Neural Network Technique for bitcoin prediction. They integrate with market technical indicators, but the performance and training time suffer as a result.

LSTM-based papers Sean et al [13]present a method for calculating the price of Bitcoin that combines the use of a recurrent neural network and long short-term memory. Ruchi Mittal et al. [14] present an automated cryptocurrencies price prediction based on historical trend using machine learning techniques (daily trend). Chih-Hung et al.

[15] proposed two different LSTM models (standard LSTM and LSTM with AR(2) model) with 208 records dataset, and compared them to MSE, RMSE, MAE, and MAPE. Under the several factors that affect the market, developed a common stock market prediction model based on LSTM. In this study, they selected three stocks with comparable tendencies. The LSTM prediction model did well.

The writers of [16] found the daily patterns in the Bitcoin market using the semantics and insights connected with the elements that influence Bitcoin price. The writers have also anticipated the daily price fluctuations of Bitcoin.

In the study given by these authors, Random Forest and Bayesian regression were used. According to another study [17], the Neural Networks ensemble technique is utilized to investigate the relationship between the next day's change in Bitcoin price and Bitcoin attributes. Their method is built using a Multi-Layered perceptron model that corresponds to each and every neural network in the ensemble.

Baur, Dirk G.et al.(2015) [18] offered a paper to investigate whether Bitcoin is a money or an asset, what is its current usage, and what usage will most likely dominate in the future given its characteristics. If Bitcoin is primarily utilised as a medium of exchange for goods and services, it will face competition from legal tenders such as the US dollar, impacting its value and, in turn, monetary policy. If it is primarily utilized as an investment, however, it will compete with other assets such as government bonds, equities, and commodities, thereby disrupting the financial system and financial stability. Whether it's money or assets, the success of Bitcoin or similar alternatives in comparison to existing monetary and financial assets will determine their ability to impact the wider economy. The following are some research findings: (i)In the IASB's Conceptual Framework, an asset is defined as "a present economic resource held by the company as a result of previous events." "An economic resource is a right with the ability to provide economic benefits." Each cryptocurrency held by an entity must be evaluated to see if it qualifies as an asset. (ii) [19] estimates that over 70% of all Bitcoins are held in inactive accounts. This demonstrates that the majority of people regard Bitcoin as a valuable asset. Bitcoin and gold have a lot in common. The government or other groups have no control over supply. Bitcoin is a worldwide money that has no nationality.

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Adj Close	Volume
2	2014-09-17	465.864014	488.174011	452.421997	457.334015	457.334015	21056800
3	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
4	2014-09-19	424.102997	427.834991	384.532013	394.79599	394.79599	37919700
5	2014-09-20	394.673004	423.29599	389.882996	408.903992	408.903992	36863600
6	2014-09-21	408.084991	412.425995	393.181	398.821014	398.821014	26580100
7	2014-09-22	399.100006	406.915985	397.130005	402.152008	402.152008	24127800
8	2014-09-23	402.09201	441.557007	396.196991	435.790985	435.790985	45099900
9	2014-09-24	435.751007	436.112	421.131989	423.204987	423.204987	30627700
10	2014-09-25	423.156006	423.519989	409.467987	411.574005	411.574005	26814400
11	2014-09-26	411.428986	414.937988	400.009003	404.424988	404.424988	21460800
12	2014-09-27	403.556	406.622986	397.372009	399.519989	399.519989	15029300
13	2014-09-28	399.471008	401.016998	374.332001	377.181	377.181	23613300
14	2014-09-29	378.928009	385.210999	372.23999	375.46701	375.46701	32497700
15	2014-09-30	378.088013	390.97999	373.442993	388.944	388.944	34707300
16	2014-10-01	387.427002	391.378998	380.779999	383.61499	383.61499	26229400

Fig. 1. Sample Data

Bitcoin and gold both have value due to their scarcity and high mining costs. They don't earn any kind of revenue.

However, because Bitcoin also includes fiat currency features, it was determined that Bitcoin is a hybrid of assets and currency.

III. PROPOSED METHODOLOGY

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This is a behavior required in complex problem domains like machine translation, speech recognition, and more. LSTMs are a complex area of deep learning. It can be hard to get your hands around what LSTMs are, and how terms like bidirectional and sequence-to-sequence relate to the field

The link between the hidden layers of RNN is different in LSTM, which is an upgraded form of RNN. Figure 1 depicts the RNN explanation structure. The structure of RNN and LSTM is similar; the only difference is the memory cell of the structure hidden layer. The gradient difficulties are efficiently solved by the construction of three unique gates.

A. Assembling Data

According to Figure 2, the method begins with data collection, the dataset is sourced from yahoo finance stock market and is based on the USD exchange rate. This research employed time-series data on this dataset, with 2713 entries in CSV format, during a period from 2014-09-17 to 2022-02-19, which is historical data pricing. Fig. 1 depict sample dataset data.

B. Data Analysis and Visualization

To analyze and further understand the collected data, we visualized it using scatter and bar plots. Bitcoin price has drastically fluctuated from 200 dollar in year 2014 to 15000 dollar in year 2018 to 3000 dollar in year 2019(these values are apporx) so we have just considered 1 Year to avoid this type of flucation in the data.

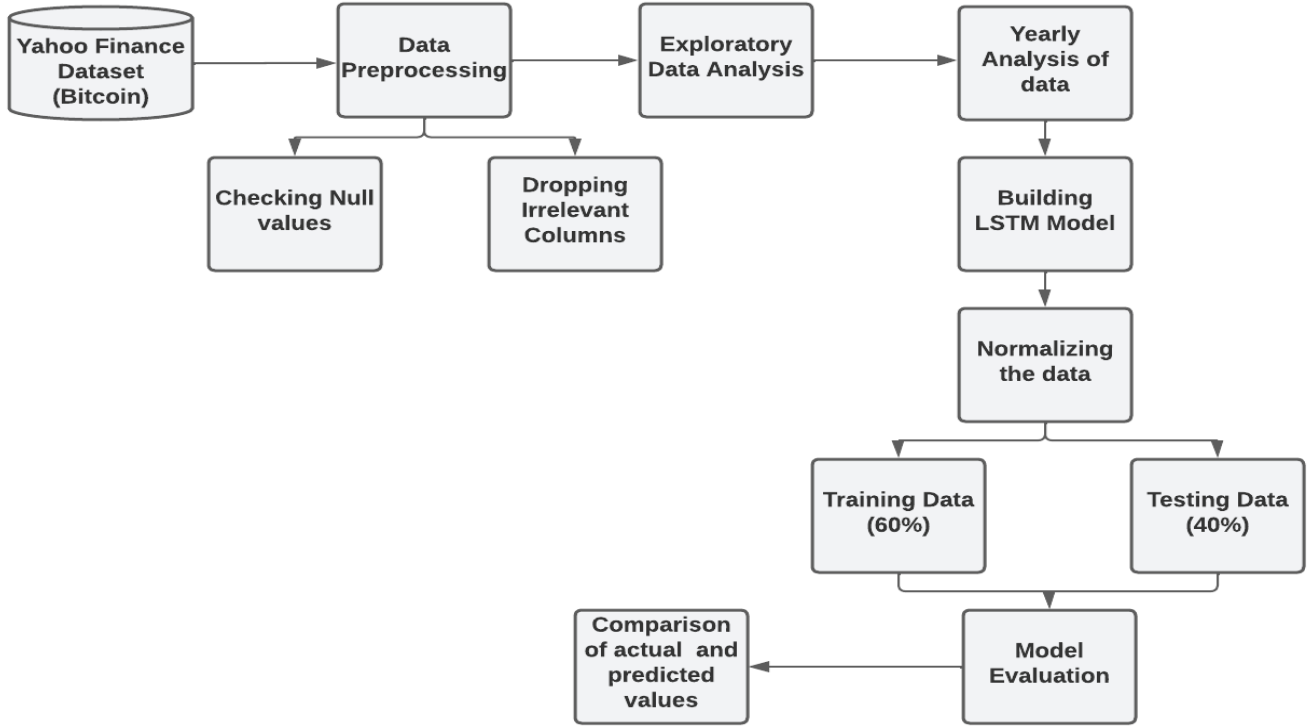


Fig. 2. System Diagram Of Bitcoin Price Prediction

C. Preprocessing Data

We preprocessed the data by dropping columns that were not relevant for our system. We have dropped every other column besides date and close price. Our current data is normalized using min-max scalar. In MinMaxScaler, for each value in a feature, MinMaxScaler subtracts the minimum value in the feature and then divides by the range. The range is the difference between the original maximum and original minimum. This preserves the shape of the original distribution.

D. Building Model

We used the preprocessed data to perform the forecasting using LSTM. Since Bitcoin price has drastically fluctuated from 200 dollar in year 2014 to 15000 dollar in year 2018 to 3000 dollar in year 2019 (these values are approx) so we have just considered 1 Year to avoid this type of fluctuation in the data. We have used tensorflow keras library for building the model. 'Relu' was used as activation function and for loss function we used 'mean squared error'. For optimization we have used 'adam', then we evaluated the model.

IV. IMPLEMENTATION & RESULTS

To implement the proposed methodology for forecasting the price of bitcoin, the data of a particular year is taken into consideration since the price of bitcoin has been drastically fluctuating through these years such that data fluctuation can



Fig. 3. Considered Period For Prediction Of Bitcoin

be avoided. Considered period to predict bitcoin close price is shown in Fig 3

For predicting the close price, two parameters are taken into account. i. Close and Date. Date column is deleted followed by normalization using min max scaler and slicing of data into training and testing data wherein normalization is a data preparation technique that is frequently used in machine learning and is the process of converting the values of numeric columns in a dataset to a similar scale without distorting the ranges of values or losing information. Shape of the training data is (219,1) and testing data is (146,1).

Dataset of close price is transformed based on Time-series analysis forecasting requirement. Initial step is to convert an array of values into a dataset matrix, wherein timestep is set to be 15 and final dataset is reshaped into [samples, time steps, features] which is required for LSTM.

Comparison between original close price vs predicted close price



Fig. 4. Forecasted Values

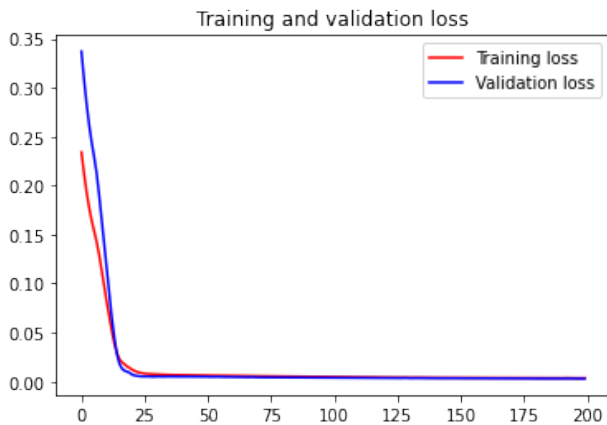


Fig. 5. Training and Validation Loss

As mentioned above, LSTM model is used for predicting the prices, sequential model is built with additional layers of LSTM with relu as an activation function followed by addition of dense layer and adam optimizer. Model is trained for batch size of 32 and epochs 200, and it has been observed that the loss was decreasing with every epoch as shown in Fig 5 with final loss as 0.0035 and validation loss as 0.0030. The quality of a statistical or machine learning model is measured using evaluation metrics. Any project requires evaluating machine learning models or methods. There are a variety of evaluation measures that may be used to test a model. Evaluation metrics used to test the sequential model are RMSE, MSE and MAE. MAE: The Mean absolute error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in

the dataset. MSE: Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals. RMSE: Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals. Train data RMSE: 2221.74, Train data MSE: 4936159.86, Train data MAE: 1753.644, Test data RMSE: 2057.12678, Test data MSE: 4231770.5901, Test data MAE: 1601.856.

The variance of observed values is a measure of how far they deviate from the average of anticipated values, or how far they range from the predicted value mean. The goal is to achieve a low value. The r2 score measures how low something is. The r2 score can range from 0 to 100 percent. It is similar to, but not identical to, the MSE. If the entire variance explained by the model is equal to the total variance, the two variables are completely linked, i.e., there is no variance at all. A low number indicates a low level of correlation, implying a regression model that isn't always valid. Train data explained variance regression score: 0.9434 , Test data explained variance regression score: 0.9485. Train data R2 score: 0.9432, Test data R2 score: 0.9481.

Comparison between original close price and predicted closed price is shown in Fig 4 which consists of Original close price, Train predicted close price and Test predicted close price. It has been observed that the predicted close price is close to the original close price which concludes that the LSTM model performed well with a regression score of 0.948 and R2 score of 0.9432.

V. CONCLUSION

Our proposed methodology was successful in predicting bitcoin prices using the Yahoo Finance stock market. Our

model can produce results using time series approaches, and the results can estimate the price for the future days using the data split to train and test that we said in the article above. The disadvantage is that the outcome isn't good enough in terms of RMSE, perhaps in the thousands or near one thousand. Following that, as previously said in the essay, the stock market is influenced by a variety of uncertainty factors. Many uncertainties influence stock markets, such as political and economic issues that have an impact on local and global levels. As a result, using LSTM to predict bitcoin price isn't good enough to make the decision to invest in bitcoin, it is another side for taking the decisions.

REFERENCES

- [1] M. Muzammal, Q. Qu, and B. Nasrulin, "Renovating blockchain with distributed databases: An open source system," *Future Generation Computer Systems*, vol. 90, pp. 105-117, 2019.
- [2] 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.
- [3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [4] Karasu, S., Altan, A., Saraç, Z., Hacıoğlu, R. (2018, May). Prediction of Bitcoin prices with machine learning methods using time series data. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- [5] Phaladisailoed, T., Numnonda, T. (2018, July). Machine Learning Models Comparison for Bitcoin Price Prediction. In *2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE)* (pp. 506-511). IEEE.
- [6] Jang, H., Lee, J. (2018). An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *IEEE Access*, 6, 5427-5437.
- [7] Saad, M., Mohaisen, A. (2018, April). Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 704-709). IEEE.
- [8] Velankar, S., Valecha, S., Maji, S. (2018, February). Bitcoin price prediction using machine learning. In *Advanced Communication Technology (ICACT)*, 2018 20th International Conference on (pp. 144-147). IEEE.
- [9] A. Greaves and B. Au, "Using the bitcoin transaction graph to predict the price of bitcoin," *No Data*, 2015.
- [10] 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.
- [11] 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.
- [12] 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.
- [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] Velankar, S., Valecha, S., Maji, S. (2018, February). Bitcoin price prediction using machine learning. In *Advanced Communication Technology (ICACT)*, 2018 20th International Conference on (pp. 144-147). IEEE.
- [17] Sin, E., Wang, L. (2017, July). Bitcoin price prediction using ensembles of neural networks. In *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 666-671). IEEE.
- [18] D. G. Baur, K. J. Hong, and A. D. Lee, "Bitcoin-Currency or Asset?," presented at the 2016 Financial Institutions, Regulation Corporate Governance (FIRCG) Conference, Melbourne Business School, 2015.
- [19] F. Glaser, K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering, "Bitcoin - Asset or Currency? Revealing Users' Hidden Intentions," presented at the ECIS 2014 (Tel Aviv). Available at SSRN: <https://ssrn.com/abstract=2425247>, 2014.