International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 1 Jan 2022, pp: 847-852 www.ijaem.net ISSN: 2395-5252

Bitcoin Price Prediction Using Machine Learning and Neural Network Model

Pratiksha Patil

B.E. Student, Smt. Indira Gandhi College Of Engineering , Ghansoli ,Navi-Mumbai , Maharashtra , India

Submitted: 15-01-2022 Revised: 23-01-2022 Accepted: 26-01-2022

ABSTRACT: Bitcoin has become an investment due to its increasing popularity, and it operates on the Block chain technology, which has given rise to other crypto currencies.This makes predicting its value extremely challenging, thus this predictor is put to the test using a Machine Learning Algorithm and an Artificial Neural Network Model.We used Bitcoin data sets to test and train the machine learning and artificial intelligence models in this work.The data filtration process was completed with the help of python libraries.Python provides a fantastic capability for data visualisation and analysis.We reduce the data and employ the characteristics or attributes that are most suited for the model after we have a good knowledge of it.

The model is put into action, and the results are recorded.When compared to other Machine Learning models from relevant studies, it was determined that the linear regression model's accuracy rate is extremely high; it was shown to be 99.87 percent correct.The LSTM model, on the other hand, has a 0.08 percent minor error rate.As a result, the neural network model is superior to the machine learning model in terms of optimization.The tkinter library was used to develop a tiny GUI that allows the user to input the High, Low, and Open features values and then anticipate the coin's future value.The results of a machine learning model and an artificial neural network model are compared in this research.We selected linear regression to compare it to the LSTM model because it delivered the highest accuracy compared to the other machine learning models.

KEYWORDS: Machine Learning; Artifificial Neural Network; Bitcoin; Block chain; Crypto currency.

I. INTRODUCTION

Bitcoin is a digital crypto currency that runs on a decentralised internet network; it may be traded using an online peer-to-peer Bitcoin network that is not depending on a single administrator or central bank[1].Because it is accepted in over 40 nations throughout the world , it has prompted the birth of new alternative coins.Other crypto currencies, products, and services are also exchanged using Bitcoin[2,3].Due to block chain technology, where each electronic coin is encrypted with a unique digital signature that makes it easier to track and can be trusted, no hacker has been able to breach it since its launch in 2009.Before passing it on, each owner signs a digital hash of the previous transaction and adds the public key of the next owner[4,5].Bitcoin's value increased from $1,000 USD in January 2017 to $16,000 USD by the end of December 2017, and it is now worth 32818 USD as of July 2021[6,7].We can say that the crypto market is quite volatile, and Bitcoin, out of all the crypto currencies available, is the most popular among investors due to its anonymity and transparency[8,9].The goal of this study is to develop a Bitcoin price prediction system that employs a variety of machine learning techniques and deep learning models.The price of Bitcoin is affected by a variety of factors; in this project, we will focus on the open, close, high, and low elements. There are six libraries in this work: Pandas , Seaborn , Scikit , Tikinter , Pickle[10,11].Section 2 of the study article contains other authors' work that is similar to this project, Section 3 contains the proposed idea and technique, Section 4 has a brief summary of the data sets and outcomes, and Section 5 contains the conclusion and future work.

II. RELATED WORK

The day trading approach was used to predict the values of ETH, BTC, XEM, XRP, XLM, and LTC using the SVM and SVM-PSO methods, where they used the day trading method to predict the values of ETH, BTC, XEM, XRP, XLM, and LTC.The optimal results are shown by SVM-PSO.The accuracy of different Classifiers varies depending on the coin.However, while this article just uses a machine learning technique, the data might be improved much more by using the

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 1 Jan 2022, pp: 847-852 www.ijaem.net ISSN: 2395-5252

Deep Learning approach[12].The use of a transaction graph to anticipate Bitcoin price is advocated.The Baseline, Logistic Regression, SVM, and Neural Network models are used in the experiment, and their accuracy is 53.4 percent, 54.3 percent, 53.7 percent, and 55.1 percent, respectively.This paper's feature selection is based on the Bitcoin block chain network, which is the least informative characteristic for Bitcoin price prediction[13].On ETH, BTC, and XRP, emotive analysis and machine learning principles like SVM and Random Forest were used to predict crypto currency prices, with BTC having the greatest accuracy of 0.72.Since machine learning techniques were used, this accuracy rate has been quite low, and it can be improved by testing with deep learning models[14].Xgboost Tree, Neural Network, Ensemble Learning Method (with the best accuracy of 92.4 percent), and KNN model were used to create a prediction model[15].SVM provides the best accuracy for a time-scale activity consisting of daily, 15-, 30-, and 60-minute returns, according to a prediction system that uses Log regression, SVM, ANN, and random forest.Although SVM tends to produce better results than the other four algorithms, when Deep learning concepts are used, the prediction system can still produce better outcomes[16].Using the open, low, and high costs, a linear regression model was applied to forecast the various cryptocurrency prices.The experiment yielded a 99.3 percent accuracy.Although this article has a good accuracy rate, the data set employed is insufficient for a model to work on a real-time chart.

III. METHODOLOGY

Data collection is the first step we do when beginning a project.It is defined as the process of gathering, measuring, and interpreting precise research insights utilising established validation methodologies.An analyst would then be able to evaluate their theory based on the information received.Regardless of the subject of study, information collection is by far the most important and important advance for research. For many disciplines of study, the process for gathering information varies depending on the data required.The most important goal of data collecting is to ensure that the information acquired is rich in substance and dependable for statistical analysis, allowing for efficient and effective data-driven decision-making.The data set includes day transactions from October 24, 2020, to August 15, 2020.When there are large or multiple data sets, the data is first verified using regression techniques, and then a deep learning model is deployed to provide superior accuracy than machine learning principles.

```
btc.head()
```

	Date	Symbol	Open	High	Low	Close	Volume BTC	Volume USDT
0	2020-08-05 03-PM	BTCUSDT	11617.56	11803.94	11593.01	11678.72	2108.69	24578947.18
1	2020-08-05 02-PM	BTCUSDT	11609.99	11644.65	11466.00	11617.56	5554.28	64363597.97
2	2020-08-05 01-PM	BTCUSDT	11562.86	11620.00	11542.32	11609.99	4847.99	56151259.42
3	2020-08-05 12-PM	BTCUSDT	11438.06	11584.60	11391.59	11562.86	5612.11	64457434.98
4	2020-08-05 11-AM	BTCUSDT	11393.24	11450.00	11382.21	11438.06	3069.55	35035195.65

```
btc.tail()
```

	Date	Symbol	Open	High	Low	Close	Volume BTC	Volume USDT
26023	2017-08-17 08-AM	BTCUSDT	4340.99	4377.85	4333.32	4360.69	0.9499	4139.70
26024	2017-08-17 07-AM	BTCUSDT	4324.35	4349.99	4287.41	4349.99	4.4400	19241.06
26025	2017-08-17 06-AM	BTCUSDT	4315.32	4345.45	4309.37	4324.35	7.2300	31282.31
26026	2017-08-17 05-AM	BTCUSDT	4308.83	4328.69	4291.37	4315.32	23.2300	100304.82
26027	2017-08-17 04-AM	BTCUSDT	16199.91	16199.91	4261.32	4308.83	44.5100	190952.85

Fig 1. Display of the Data collected

Now that we have all of the data we need for the project, we can go on to the following step, data segregation or feature selection.Trimming undesired data or removing unneeded info from a data set is what this method entails.This phase is crucial because we only need the attributes that can help us anticipate the outcome, while additional data can produce noise in our final result.To put it another way, we segregate data so that we may build a better model that gives us an optimum outcome, avoid over-fitting or redundancy, and shorten training time so that the system can provide output faster and more accurately. Data visualisation is a technique for better comprehending data or information by presenting it in a diagrammatic format.With the use of graphics, data visualisation allows us to communicate data relationships.These visuals are in the shape of patterns that are very simple to comprehend.This is one of the primary ways that machine learning aids with data analysis.Whether you work in finance, marketing, technical support, or design, you must be able to visualise data in order to comprehend

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 1 Jan 2022, pp: 847-852 www.ijaem.net ISSN: 2395-5252

it.As a result, data visualisation has become increasingly vital in today's environment
.

Features	Definition
Open	Opening value of trade at that time stamp
Close	Closing value of trade at that time stamp
High	Highest trade value in the time stamp
Low	Lowest trade value in the time stamp
Volume BTC	Total trade volume in BTC in the given timestamp
Volume USDT	Total trade volume in USDT in the given timestamp
Date	The given date and time of each bid
Symbol	Symbolic representation of coin

Fig 2. The features represented in the data

We can see the link between attributes and pinpoint the ones we need with the help of data visualisation libraries.A sample image of the correlation graph between the features in the given data set is presented below.The relationship between "Volume of USDT" and "Volume of BTC" with the other features is not evident in the presented image Figure 3, thus they are separated and removed from the training data set.

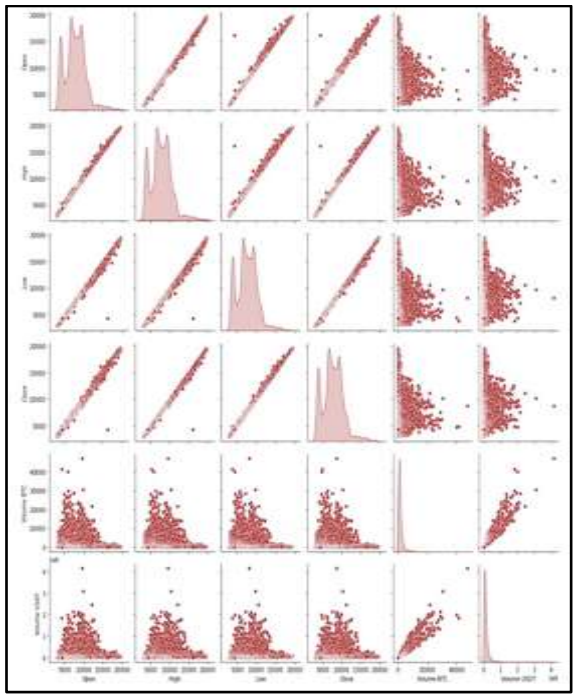

Fig 3. Correlation graph between the features.

Linear regression is a technique for determining the relationship between dependent and independent variables and predicting future events.Simple linear regression is used when only one dependent and one independent variable are used.
Multi-linear regression is used when the number of independent and dependent variables increases.

The graph is drawn using a straight line across the graph that is calculated using the least square approach to find the best fit.
y = mx+C
C = y intercept
m = slope
x, y are the points on the graph

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2$$

MSE = mean squared error
n = *number of data points*
Y_i = predicted values
$\widehat{Y}_i$ = predicted values

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 1 Jan 2022, pp: 847-852 www.ijaem.net ISSN: 2395-5252

Long Short-Term Memory, It's a deep learning notion, specifically a Recurrent Neural Network concept, that avoids the vanishing gradient issue.The fundamental advantage of this approach is that it prevents back propagation mistakes from disappearing or inflating, allowing them to flow backward via an infinite number of virtual layers unfurled in space.LSTM mostly works with time series graphs with data sets containing events that occurred hundreds or millions of discrete time steps previously.It can handle signals with a mix of low and high-frequency components and works with considerable delays between critical incidents.LSTM has been utilised by a number of researchers to predict time series related data sets for stock prediction, with greater or higher accuracy than other algorithms.Unlike prior models based on Hidden Markov Models (HMM) and other comparable concepts, LSTM is capable of understanding context-sensitive language.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

$y = prediction$
$x = true\ value$
$n = total\ number\ of\ data\ points$

V. RESULT AND DISCUSSION

Following the data analysis, we discovered that only four features were suitable for this project's testing.Only the specified features were remained after the data was reduced, as shown in Figure 4.

	Open	High	Low	Close
0	11617.56	11693.94	11593.01	11678.72
1	11609.99	11644.65	11466.00	11617.56
2	11562.86	11620.00	11542.32	11609.99
3	11438.06	11584.60	11391.59	11562.86
4	11393.24	11450.00	11382.21	11438.06

Fig 4. Attribute/Features selected are Open, High, Low, and Close

We can see the result of two models: one is a Machine Learning model, which is Linear regression, and the other is a Recurrent Neural Network model, which is Long Short-Term Model, which shows us two different outputs. The Mean Squared Equation, which shows us the correctness of the linear graph with regard to the continuous-time frame data set, is commonly used in linear regression.As shown in Figure 5, the accuracy of the training data is around 99.97 percent, while the accuracy of the testing data is also roughly 99.97 percent.Meanwhile, as demonstrated in Figure 6, the LSTM model tends to find accuracy with respect to the Mean Absolute Error, which displays an error rate of roughly 0.08 percent.The Mean Squared Equation, which gives us the correctness of the linear graph with regard to the continuous-time frame data set, is commonly used in regression.

```
In [38]: model.score(x_train, y_train)
Out[38]: 0.9997158887216999

In [39]: pred = model.predict(x_test)

In [40]: model.score(x_test, y_test)
Out[40]: 0.9997966090479169
```

Fig 5. Accuracy obtain from the training and testing data set usingLinear regression model

S.No	Open	High	Low	Close	Expected Result
1	11617.56	11693.94	11593.01	11678.72	11669.05
2	11609.99	11644.65	11466.00	11617.56	11530.74
3	11562.86	11620.00	11542.32	11609.99	11603.13
4	11438.06	11584.60	11391.59	11562.86	11525.64
5	11393.24	11450.00	11382.21	11438.06	1140.47

Fig 6. Testing of Linear regression model

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 1 Jan 2022, pp: 847-852 www.ijaem.net ISSN: 2395-5252

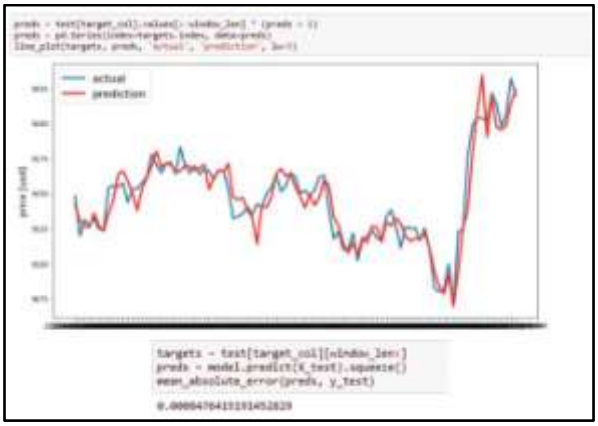

Fig 7. Final Resultant graph of LSTM and the Mean Absolute Error rate (0.08%)

The data visualisation depicts the correlation between all of the features, although only the four chosen features have a significant association.The data is then fitted into the model using the specified python commands.

These data models were trained and tested on a small number of data sets, and the results were provided.With the advancement of technology and the expansion of data sets, we may continue to test the model with a variety of alternative crypto currencies.The LSTM model has a slightly higher prediction rate than the linear regression model, but only by a little margin.

VI. CONCLUSION

According to the findings, Long Short-Term Memory has a higher accuracy rate than Linear Regression.Because this study only compares the features of open, close, high, and low, the outcome may alter if we examine additional factors.Data sets cannot be the main rationale for forecasting because the crypto market is dynamic and influenced by social media and other external factors.New data can be acquired, evaluated, and rehearsed as technology progresses, resulting in greater findings for this experiment.

VII. FUTURE WORK

1) More algorithms are being implemented in order to determine the best approach for predicting the crypto currency.
2) Implementing IOT model for smart automatic analysis.
3) To work on a better User Interface so that people can access these data easily and effortlessly.

REFERENCES

[1]. Sin E, Wang L. Bitcoin price prediction using ensembles of neural networks. In: 2017 13th International conference on natural computation, fuzzy systems and knowledge discovery. IEEE. 2017;p.666–671. doi:10.1109/FSKD.2017.8393351.

[2]. Shankhdhar A, Singh AK, Naugraiya S, Saini PK. Bitcoin Price Alert and Prediction System using various Models. IOP Conference Series: Materials Science and Engineering. 2021;1131(1):012009. Available from: https://dx.doi.org/10.1088/1757899x/1131/1/01200

[3]. Mittal R, Arora S, Bhatia MP. Automated cryptocurrencies prices prediction using machine learning. ICTACT Journal on Soft Computing. 2018;8(4):1758– 1761. Available from: http://ictactjournals.in/paper/IJSC_Vol_8_Iss_4_Paper_8_1758_1761.pdf.

[4]. Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. 2019. Available from: https://git.dhimmel.com/bitcoin-whitepaper/.

[5]. Sebastião H, Godinho P. Forecasting and trading cryptocurrencies with machine learning under changing market conditions. Financial Innovation.

[6]. 2021;7(1):1–30. Available from: https://dx.doi.org/10.1186/s40854-020-00217-x.

[7]. Phaladisailoed T, Numnonda T. Machine learning models comparison for bitcoin price prediction. 10th International Conference on Information Technology and Electrical Engineering. 2018;p. 506–511. Available from: 10.1109/ICITEED.2018.8534911.

[8]. Jaquart P, Dann D, Weinhardt C. Short-term bitcoin market prediction via machine learning. The Journal of Finance and Data Science. 2021;7:45–66. Available from: 10.1016/j.jfds.2021.03.001.
[9]. Rane PV, Dhage SN. Systematic erudition of bitcoin price prediction using machine learning techniques. 5th International Conference on Advanced Computing & Communication Systems (ICACCS). 2019;p. 594–598. Available from: 10.1109/ICACCS.2019.8728424.
[10]. Roy R, Roy S, Hossain MN, Alam MZ, Nazmul N. Study on nonlinear partial differential equation by implementing MSE method. Global Scientific
[11]. Journals. 2020;8(1):1651–1665.
[12]. Mckinney W. Pandas: a foundational Python library for data analysis and statistics. Python for High Performance and Scientific Computing. 2011;14(9):1–9. Available from: https://www.dlr.de/sc/portaldata/15/resources/dokumente/pyhpc2011/submissions/pyhpc2011_submission_9.pdf.
[13]. Waskom M. Seaborn: statistical data visualization. Journal of Open Source Software. 2021;6(60):3021. Available from: https://dx.doi.org/10.21105/joss. 03021.
[14]. An optimized support vector machine (SVM) based on particle swarm optimization (PSO) for cryptocurrency forecasting. Procedia Computer Science. 2019;163:427–433. Available from: https://doi.org/10.1016/j.procs.2019.12.125.
[15]. Greaves A, Au B. Using the bitcoin transaction graph to predict the price of bitcoin. 2015. from:http://snap.stanford.edu/class/cs224w-2015/ projects_2015/Using_the_Bitcoin_Transaction_Graph_to_Predict_the_Price_of_Bitcoin.pdf.
[16]. 14) Valencia F, Gómez-Espinosa A, Valdés-Aguirre B. Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning. Entropy. 2019;21(6):589. Available from: https://dx.doi.org/10.3390/e21060589.
[17]. Chowdhury R, Rahman MA, Rahman MS, Mahdy MRC. An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning. Physica A: Statistical Mechanics and its Applications. 2020;551(124569). Available from: https://dx.doi.org/10.1016/j.physa.2020.124569.
[18]. Akyildirim E, Goncu A, Sensoy A. Prediction of cryptocurrency returns using machine learning. Annals of Operations Research. 2021;297(1-2):3–36.from:https://dx.doi.org/10.1007/s10479-02003575-y.
[19]. Huang X, Zhang W, Huang Y, Tang X, Zhang M, Surbiryala J, et al.. 2021.
[20]. Muniye TA, Mohanty MR, Satapathy S. Bitcoin Price Prediction and Analysis Using Deep Learning Models. In: Communication Software and Networks. 2020. Available from: 10.1007/978-981-15-5397-4_63.