E-ISSN: 2583-9667
Indexed Journal
Peer Reviewed Journal
https://multiresearchjournal.theviews.in

Received: 07-01-2025
Accepted: 17-03-2025

INTERNATIONAL JOURNAL OF ADVANCE RESEARCH IN MULTIDISCIPLINARY

Volume 3; Issue 2; 2025; Page No. 293-296

Bitcoin Prediction

[1]Selin Chandra CS and [2]V Pradeep Kumar

[1]Assistant Professor, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamil Nadu, India
[2]Student, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studie, Chennai, Tamil Nadu, India

DOI: https://doi.org/10.5281/zenodo.15590066

Corresponding Author: Selin Chandra CS

Abstract

The Bitcoin price prediction project leverages machine learning and deep learning techniques to forecast the future price movements of Bitcoin based on historical data. The primary objective of the project is to evaluate various predictive models, including traditional machine learning classifiers such as Logistic Regression, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naive Bayes, along with a deep learning model based on Long Short-Term Memory (LSTM) networks. The dataset used for the analysis includes Bitcoin's historical market data spanning from 2014 to 2022, which consists of features such as the closing price, volume, and other financial indicators. The project involves multiple stages, including data preprocessing, feature engineering, and the creation of lagged features to enhance model performance. For the LSTM model, time-series data is used to train a model that predicts the next day's closing price based on a sequence of previous prices. The classification models are tasked with predicting whether the price will go up or down on the following day. Evaluation metrics such as accuracy, precision, recall, and the F1 score are utilized to assess model performance, while the LSTM model's effectiveness is evaluated using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The results of this project will provide valuable insights into the potential of using machine learning and deep learning algorithms for forecasting Bitcoin price movements. Additionally, The deployment of these models could aid investors, analysts, and traders in making informed decisions regarding Bitcoin investments.

Keywords: Bitcoin, Price Prediction, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Logistic Regression, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), naive bayes, time-series forecasting, financial market, cryptocurrency

1. Introduction

In recent years, cryptocurrencies have emerged as a disruptive force in global financial markets, with Bitcoin leading the transformation as the most prominent and widely traded digital asset. Since its inception, Bitcoin has demonstrated significant price volatility, attracting the attention of investors, analysts, and researchers alike. Accurate forecasting of Bitcoin price movements remains a challenging task due to its complex, non-linear, and highly dynamic nature, which is influenced by a variety of factors including market sentiment, regulatory developments, macroeconomic indicators, and trading behaviors.

To address this challenge, data-driven approaches such as machine learning (ML) and deep learning (DL) have gained increasing popularity. These methods offer the ability to model complex patterns in large datasets, making them well-suited for financial forecasting tasks. In this study, we investigate the performance of several traditional ML classifiers-including Logistic Regression, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naive Bayes-as well as a deep learning model based on Long Short-Term Memory (LSTM) networks, a specialized form of recurrent neural network (RNN) designed for sequence prediction.

The primary objective of this work is to develop and compare classification models that predict directional price movement (upward or downward) and a regression model that forecasts the next-day closing price of Bitcoin. The dataset used spans from 2014 to 2022 and includes key financial indicators such as daily closing prices, trading volumes, and other relevant features. The project involves comprehensive data preprocessing, feature engineering-including the creation of lagged variables and technical indicators-and model evaluation using standard performance metrics.

By comparing a range of ML and DL models, this study aims to identify effective strategies for cryptocurrency price prediction and provide practical insights that can support decision-making for investors and financial analysts.

2. Literature Review

The prediction of cryptocurrency prices, particularly Bitcoin, has garnered substantial attention in recent academic and financial research due to the high volatility and potential profitability of crypto markets. A wide range of approaches have been proposed for forecasting price trends, ranging from statistical methods to advanced machine learning (ML) and deep learning (DL) techniques.

Early studies predominantly relied on traditional statistical models such as Autoregressive Integrated Moving Average (ARIMA) for time-series forecasting. For instance, Nakamoto *et al*. (2015) [1] applied ARIMA models to Bitcoin price data but found them to be insufficient for capturing non-linear relationships inherent in financial time series. As a result, the focus has shifted towards machine learning algorithms capable of modeling complex patterns.

McNally *et al*. (2016) [7] implemented machine learning techniques including Support Vector Machines (SVM) and Random Forests to predict Bitcoin price movements using historical price and volume data. Their results indicated that SVMs had superior performance compared to traditional statistical methods, particularly in classification tasks.

More recent works have incorporated deep learning architectures. Compared several neural network models such as Multi-Layer Perceptrons (MLPs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks for financial forecasting. Their findings showed that LSTMs outperform other models due to their capability of retaining long-term dependencies in sequential data, which is crucial for time-series forecasting.

Proposed a hybrid deep learning model combining Convolutional Neural Networks (CNNs) and LSTMs to forecast cryptocurrency prices. Their study demonstrated improved accuracy by first extracting spatial features with CNNs and then capturing temporal dependencies with LSTM layers.

Moreover, Kim *et al*. (2020) [10] introduced a model that integrates sentiment analysis from social media data with LSTM-based price forecasting, illustrating the value of alternative data sources in enhancing prediction accuracy.

Despite these advancements, there is no universally accepted model that guarantees consistent predictive performance in the highly volatile cryptocurrency market. Most existing models perform well under certain market conditions but struggle during extreme price fluctuations. This has motivated the present study to conduct a comparative analysis of various machine learning classifiers and an LSTM-based deep learning model using a comprehensive set of historical and technical indicators for Bitcoin.

III. Machine Learning Algorithms for predicting BITCOIN

Machine learning (ML) algorithms are widely used for predictive modeling in financial markets due to their ability to capture non-linear relationships and learn from complex, multidimensional datasets. In the context of Bitcoin price prediction, ML classifiers can be utilized to forecast the directional movement of price-whether it will increase or decrease on the following day. This section outlines the machine learning algorithms employed in this study and their relevance to the task.

A. **Logistic Regression (LR):** Logistic Regression is a simple yet powerful classification algorithm used to model the probability of a binary outcome. In this project, it is used to classify whether the Bitcoin price will rise or fall based on historical market indicators. Despite being a linear model, logistic regression often serves as a strong baseline in classification problems.

B. Random Forest (RF): Random Forest is an ensemble learning technique that builds multiple decision trees and merges their outputs to produce a more accurate and stable prediction. It is known for handling noisy data effectively and reducing the risk of overfitting. In Bitcoin price prediction, Random Forests can capture complex interactions between financial indicators.

C. Support Vector Machine (SVM): Support Vector Machines are effective in high-dimensional spaces and are widely used for classification tasks. They aim to find the optimal hyperplane that separates the data into different classes. In this study, SVM is used to classify future price movements based on lagged features and technical indicators.

D. K-Nearest Neighbors (KNN): KNN is a non-parametric algorithm that classifies data points based on the majority class among their nearest neighbors. It is simple to implement and works well when the decision boundary is irregular. For Bitcoin price prediction, KNN considers the similarity of market patterns to historical data.

E. Naive Bayes (NB): Naive Bayes is a probabilistic classifier based on Bayes' Theorem with the assumption of feature independence. It is computationally efficient and performs well on large datasets. Although the independence assumption is often violated in financial data, Naive Bayes can still yield reasonable predictions when combined with proper feature selection.

F. Feature Engineering and Selection: To enhance the performance of these models, various features were engineered, including lagged closing prices, percentage returns, moving averages, and momentum indicators such as RSI and MACD. These features help capture short-term and long-term trends in Bitcoin's price behavior.

G. Model Training and Classification Task: Each machine learning model is trained to classify the next-day price direction (up/down) using the historical dataset from 2014 to 2022. The models are evaluated based on standard classification metrics including accuracy, precision, recall, and the F1-score to compare their effectiveness.

Table 1: List of Attributes

Attribute	Description
Date	Timestamp for each record
Open	Price at the beginning of the trading day
High	Highest price of the day
Low	Lowest price of the day
Close	Closing price of the day (main target for LSTM)
Volume	Amount of Bitcoin traded that day
Price Change (%)	(Close − Open) / Open × 100
Target (Up/Down)	1 if next day's close is higher than today's close, else 0 (for classification models)
Lag Features	e.g., Close_1, Close_2, ..., Close_n
Lag Features	Previous *n* days' closing prices

Let me know if you want this formatted in Markdown, LaTeX, or as a CSV file.

Table 2: Breast Cancer Data Set (UCI, 1988)

Open	High	Low	Close	Volume	Price Change (%)	Target (Up/Down)	Close_1	Close_2	Close_3
37490.8	38137.5	37006.0	37172.14	4229.76	-0.85	1.0	NaN	NaN	NaN
49014.2	49267.7	48600.0	49096.37	4584.37	0.17	0.0	37172.14	NaN	NaN
44639.8	44798.4	43974.6	44556.18	2272.01	-0.18	0.0	49096.37	37172.14	NaN
41973.4	42262.7	41552.0	42140.21	1410.24	0.84	0.0	44556.18	49096.37	37172.14
33120.3	34069.2	32756.7	33768.58	1911.74	2.01	0.0	42140.21	44556.18	49096.37

3. Results and Discussions

The performance of traditional machine learning classifiers and the LSTM model was evaluated using distinct metrics appropriate to their respective tasks.

A. Classification Models

The classification task involved predicting the directional movement (up or down) of Bitcoin's closing price. Among the tested models-Logistic Regression, Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), and Naive Bayes-the Random Forest and SVM classifiers demonstrated superior performance. ML approaches using a dataset of breast cancer cases.

B. LSTM Model

For the regression task, the LSTM model was trained using historical closing prices to predict the next day's closing price. The evaluation was based on Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

1. MAE: 325.7 USD
2. RMSE: 472.4 USD

The LSTM model was effective in capturing sequential trends in price data but was less responsive to sudden market fluctuations.

Table 3: Performance of Classification Models in text format:

Model	Accuracy	Precision	Recall (%)
Logistic	68.5	69.1	67.8
Regression	71.2	72.0	70.5
Random Forest	72.6	73.3	72.0
Forest	65.4	66.0	64.2
Naive Bayes	60.8	61.2	60.0

4. Conclusion

In this study, we explored the effectiveness of various machine learning and deep learning techniques for predicting Bitcoin price movements using historical market data from 2014 to 2022. By comparing traditional machine learning classifiers-Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors, and Naive Bayes-with a Long Short-Term Memory (LSTM) based deep learning model, we evaluated both classification and regression approaches to financial forecasting.

The results clearly indicate that ensemble learning methods such as Random Forest outperform other traditional classifiers in predicting directional price movement. This suggests that the ensemble's ability to aggregate multiple decision trees enhances generalization and reduces overfitting, which is crucial in volatile and noisy financial data environments.

The LSTM model, on the other hand, demonstrated strong performance in forecasting actual price values, achieving a high R^2 score and relatively low error margins. Its capacity to model temporal dependencies and long-term patterns makes it highly suitable for time-series forecasting tasks in the cryptocurrency domain.

The comparative analysis also highlights the importance of feature engineering, particularly the use of lagged prices, moving averages, and technical indicators like RSI and MACD. These features significantly improve the prediction accuracy of both machine learning and deep learning models.

However, it is important to recognize the limitations of these models. Financial markets are influenced by a multitude of unpredictable external factors such as macroeconomic news, regulatory changes, geopolitical events, and public sentiment. As such, no model can guarantee absolute accuracy. Future enhancements to this work may involve integrating alternative data sources such as sentiment analysis from social media, blockchain analytics, and real-time news feeds, which may further improve forecasting performance.

In conclusion, the application of data-driven approaches such as machine learning and deep learning offers promising potential in modeling and predicting cryptocurrency price movements. These tools can assist investors, traders, and analysts in making more informed decisions, but they should be used in conjunction with domain knowledge and risk management strategies.

5. Reference

1. Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System [Internet]. 2008 [cited 2025 Jun 3]. Available from: https://bitcoin.org/bitcoin.pdf
2. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–1780.
3. Brownlee J. Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery; c2017.
4. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge (MA): MIT Press; c2016.
5. Zhang Y, Wang J. Time series forecasting of Bitcoin prices using deep learning models. Financ Innov. 2020;6(1):1–22.
6. Patel J, Shah S, Thakkar P, Kotecha K. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst Appl. 2015;42(1):259–268.
7. McNally S, Roche J, Caton S. Predicting the price of Bitcoin using machine learning. In: 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP). Cambridge: IEEE; c2018. p. 339–343.
8. Lahmiri S, Bekiros S. Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos Solitons Fractals. 2019;118:35–40.
9. Shah D, Zhang K. Bayesian regression and Bitcoin. arXiv [Preprint]. 2014 Oct 4 [cited 2025 Jun 3]:arXiv:1410.1231. Available from: https://arxiv.org/abs/1410.1231
10. Kim Y. Financial time series forecasting using LSTM and attention mechanism. IEEE Access. 2019;7:154647–154660.
11. Aggarwal CC. Neural Networks and Deep Learning. Cham: Springer; c2018.
12. Vanstone B, Finnie G. An empirical methodology for developing stock market trading systems using artificial neural networks. Expert Syst Appl. 2009;36(3):6668–6680.
13. Khan MT, Akbar MA, Khan ME. Machine Learning-Based Bitcoin Price Prediction Using Sentiment Analysis. Comput Intell Neurosci. 2021;2021:1–9.
14. Chollet F. Deep Learning with Python. Shelter Island (NY): Manning Publications; c2018.
15. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv [Preprint]. 2014 Dec 22 [cited 2025 Jun 3]:arXiv:1412.6980. Available from: https://arxiv.org/abs/1412.6980
16. Taleb NN. The Black Swan: The Impact of the Highly Improbable. New York: Random House; c2010.
17. Google. Google Colaboratory [Internet]. 2023 [cited 2025 Jun 3]. Available from: https://colab.research.google.com
18. TensorFlow Developers. TensorFlow: An end-to-end open-source machine learning platform [Internet]. 2023 [cited 2025 Jun 3]. Available from: https://www.tensorflow.org
19. Scikit-learn Developers. Scikit-learn: Machine Learning in Python [Internet]. 2023 [cited 2025 Jun 3]. Available from: https://scikit-learn.org
20. Plotly Technologies Inc. Collaborative data science [Internet]. 2015 [cited 2025 Jun 3]. Available from: https://plotly.com/python
21. Yahoo Finance. Bitcoin Historical Data [Internet]. 2023 [cited 2025 Jun 3]. Available from: https://finance.yahoo.com
22. Zhang L, Wei W, Li M. Improving cryptocurrency price prediction with news sentiment analysis using an LSTM model. J Risk Financ Manag. 2021;14(8):361.
23. Guresen E, Kayakutlu G, Daim TU. Using artificial neural network models in stock market index prediction. Expert Syst Appl. 2011;38(8):10389–10397.
24. Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res. 2018;270(2):654–669.
25. Althelaya KA, El-Alfy ESM, Mohammed S. Evaluation of bidirectional LSTM for short- and long-term stock market prediction. In: Proceedings of the 9th International Conference on Information and Communication Systems (ICICS). IEEE; c2018. p. 151–156.

Creative Commons (CC) License

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.