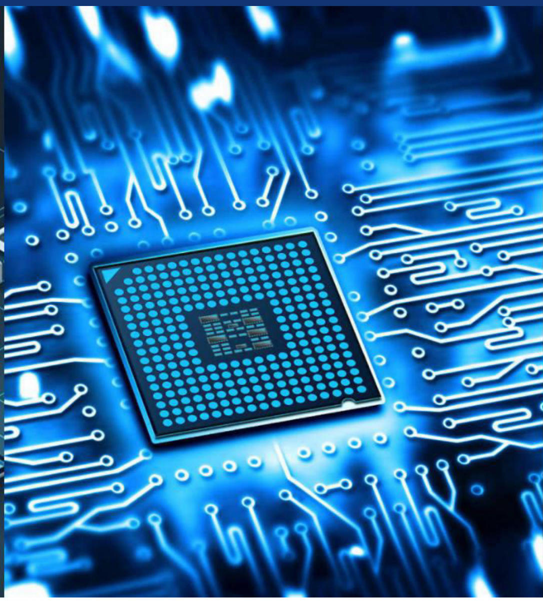
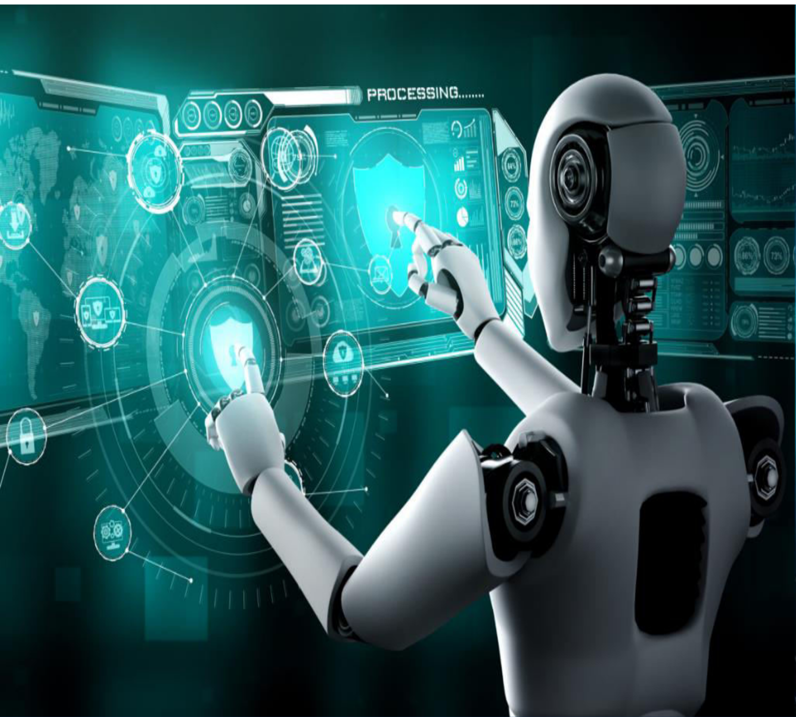
IJIRCCE

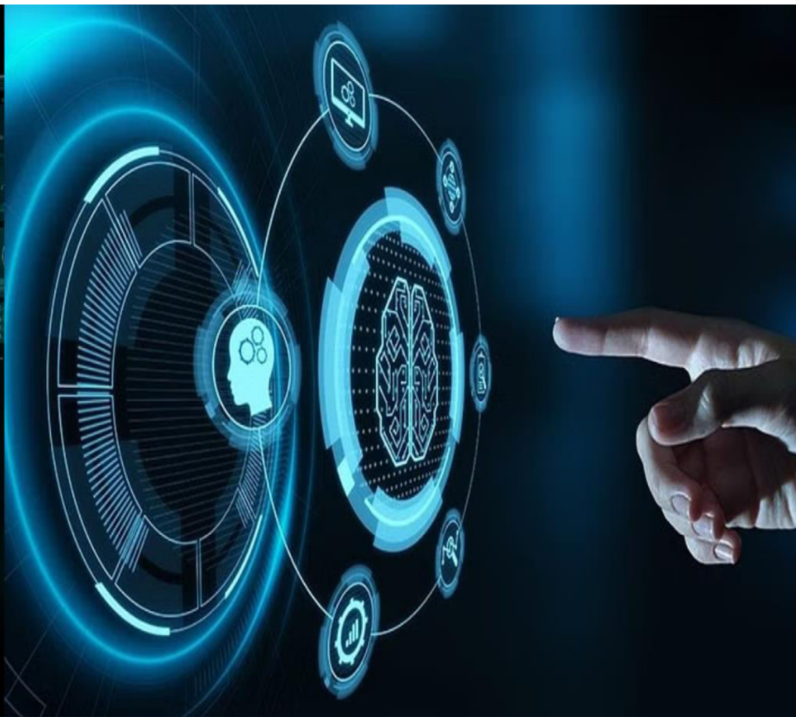
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 5, May 2026

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Bitcoin Price Prediction Using Machine Learning

Dr. AzizKhan F Pathan[1], Keerthi M Rao[2], Keerthishree B M[2], Asha M Angadi[2], Gayatri KHM[2]

HOD & Associate Professor, Dept. of AI&ML, Jain Institute of Technology, Davangere, Karnataka, India[1]

UG Students, Dept. of ISE, Jain Institute of Technology, Davangere, Karnataka, India[2]

ABSTRACT: Among all digital assets, Bitcoin remains the most prominent cryptocurrency, yet its remarkable price volatility renders accurate forecasting a significant challenge. This study introduces a Bitcoin Price Prediction System that harnesses historical market data through a suite of machine learning algorithms. The research encompasses key phases including data preprocessing, exploratory analysis, and systematic feature engineering, followed by the implementation and comparison of three predictive models: Linear Regression, Random Forest, and Long Short-Term Memory (LSTM) networks. Model effectiveness is assessed through standard regression metrics encompassing Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the R-squared (R^2) coefficient. The final system is presented through an interactive web interface developed with Streamlit. Results confirm that machine learning methods consistently outperform conventional statistical approaches and hold substantial promise for guiding investment decisions within cryptocurrency markets.

KEYWORDS: Bitcoin Price Prediction, Machine Learning, Cryptocurrency Forecasting, Time Series Analysis, LSTM, Predictive Analytics

I. INTRODUCTION

Over the past decade, cryptocurrencies have transitioned from obscure technological experiments to mainstream financial instruments with measurable impacts on global economic systems. Bitcoin, introduced by Satoshi Nakamoto in his foundational whitepaper on a decentralized peer-to-peer payment network [1], occupies a central position within this landscape. Its decentralized architecture, transparent transaction ledger, and perceived potential for substantial returns have attracted sustained interest from investors, academic researchers, financial institutions, and technology developers worldwide.

Despite its prominence, Bitcoin exhibits pronounced price instability driven by a diverse set of factors, including shifts in investor sentiment, fluctuations in market demand, evolving regulatory frameworks, geopolitical developments, changes in trading volume, and the influence of social media discourse. The complexity and unpredictability of these variables render conventional statistical forecasting techniques inadequate in isolation. Consequently, researchers have increasingly turned to Artificial Intelligence (AI) and Machine Learning (ML) methodologies to better capture the dynamics of cryptocurrency markets.

AI and ML have emerged as powerful tools in the domain of financial forecasting, primarily owing to their capacity to process vast amounts of data, uncover latent patterns, and model complex nonlinear relationships within time-series datasets. Foundational work by Brownlee [2] and Géron [3] established the effectiveness of models such as LSTM networks, Random Forest, and Linear Regression for time-series and predictive analytics tasks. In contrast to classical forecasting methods, ML models demonstrate greater adaptability to dynamic and nonlinear financial conditions.

Among deep learning architectures, Long Short-Term Memory (LSTM) networks have proven particularly well-suited for cryptocurrency price forecasting. Their inherent ability to retain information over extended time horizons enables the detection of sequential dependencies that are critical when prior price behaviour substantially influences future market outcomes [2]. Complementing LSTM, Random Forest models offer robustness in managing complex nonlinear variable interactions, while Linear Regression serves as a useful benchmark for evaluating the added value of more sophisticated approaches [3].

The present study applies these machine learning techniques to historical Bitcoin data sourced from established cryptocurrency platforms, namely CoinGecko [4] and Binance [10]. The dataset encompasses essential market attributes such as opening, closing, high, and low prices, as well as trading volume and trend indicators. Prior to model training, the data undergoes rigorous preprocessing, including missing value treatment, normalization, and feature scaling, to maximize model accuracy and training efficiency.

The computational pipeline is constructed using Pandas [5] for structured data handling, Scikit-learn [6] for classical machine learning and evaluation tools, and TensorFlow [7] for deep learning model construction. The resulting prediction system is deployed via the Streamlit framework [8], which facilitates the creation of a responsive, user-centric web application for real-time forecasting and result visualization.

The overarching objective of the proposed Bitcoin Price Prediction System is to furnish dependable short-term forecasts that deepen users' comprehension of cryptocurrency market dynamics and support evidence-based investment strategies. By converging financial analysis with machine learning, this project highlights the transformative role that AI-driven systems can play in advancing predictive analytics within the rapidly expanding digital economy.

II. LITERATURE REVIEW

The challenge of predicting Bitcoin prices has prompted a diverse body of research spanning statistical, machine learning, and deep learning paradigms. Early investigations relied primarily on traditional time-series models such as ARIMA; however, these approaches demonstrated considerable limitations when confronted with the extreme price volatility characteristic of cryptocurrency markets [9].

Subsequent research revealed that machine learning algorithms, including Random Forest and Support Vector Regression, offer superior handling of nonlinear market patterns. Géron [3] provided a comprehensive treatment of how supervised learning methods, when combined with strategic feature engineering and principled optimization, can substantially elevate predictive performance.

Deep learning, and LSTM architectures in particular, gained considerable traction in cryptocurrency forecasting research due to their effectiveness in modelling long-range sequential dependencies within financial time series. Brownlee [2] provided compelling evidence that LSTM-based models deliver meaningful improvements in time-series forecasting accuracy relative to conventional alternatives.

Several researchers have also underscored the value of incorporating technical indicators—such as moving averages, rolling statistical measures, and volatility indices—to sharpen prediction accuracy [2][3]. Furthermore, access to high-quality, real-time market feeds from sources like CoinGecko [4] and Binance [10] has been shown to substantially improve the reliability of cryptocurrency forecasting systems.

Contemporary frameworks, notably TensorFlow [7] and Scikit-learn [6], have significantly lowered the barrier to implementing sophisticated machine learning pipelines, while Streamlit [8] has enabled researchers to package prediction systems within accessible, interactive web interfaces.

III. METHODOLOGY

A. Study Design

The architecture of the proposed Bitcoin Price Prediction System is grounded in a machine learning-driven time-series framework tailored for cryptocurrency market forecasting. Time-series methodologies are particularly suited to this domain, given their capacity to detect recurring patterns, directional trends, and cyclical variations embedded in longitudinal market data. The forecasting philosophy and evaluation conventions adopted in this work are informed by the foundational contributions of Hyndman and Athanasopoulos [9], whose research underscores the centrality of rigorous statistical modelling and systematic validation in financial prediction contexts.

The core aim of this investigation is to construct an intelligent forecasting model capable of processing historical Bitcoin price data to yield dependable near-term price estimates. The system integrates both classical machine learning and advanced deep learning methodologies to enhance predictive precision and minimize error rates. The end-to-end research

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

workflow encompasses data acquisition, preprocessing, feature construction, model training, performance evaluation, and deployment within a live operational environment.

A central component of this study is the systematic comparison of three predictive algorithms—Linear Regression, Random Forest, and Long Short-Term Memory (LSTM)—to determine which approach is best equipped to handle the high volatility and nonlinearity inherent in cryptocurrency price data. Through the fusion of machine learning with financial analytics, the system delivers a data-driven lens for interpreting Bitcoin market behaviour and supporting informed investment analysis.

B. Data Collection

Historical Bitcoin market records were obtained from two reputable cryptocurrency data providers: CoinGecko [4] and Binance [10]. Both platforms are widely regarded for furnishing accurate, continuously refreshed financial market data, making them reliable sources for research applications.

The compiled dataset includes the following market attributes:

- Open Price
- Close Price
- High Price
- Low Price
- Trading Volume
- Market Capitalization
- Date and Time Information

The dataset spans a substantial time range, facilitating analysis of both short- and long-term market dynamics. Real-time data availability strengthens the system's practical relevance and predictive robustness. Data management and structuring were carried out using the Pandas library [5], which provides efficient tooling for working with structured financial datasets. Beyond raw price data, trend-based attributes and historical trading patterns were also incorporated to augment the models' predictive capabilities. High-quality, consistent data acquisition is a prerequisite for model reliability, as prediction performance is fundamentally contingent on the integrity of the input dataset.

C. Data Preprocessing

Preprocessing constitutes a foundational phase in any machine learning pipeline, particularly given that raw cryptocurrency datasets are frequently plagued by missing entries, duplicated records, inconsistent formatting, and extraneous noise. To ensure data integrity and optimize model performance, a structured set of preprocessing operations was applied using Pandas [5] and Scikit-learn [6].

The preprocessing workflow comprised the following steps:

Data Cleaning: All null values and duplicate rows were detected and eliminated. Erroneous or inconsistent records were corrected to uphold dataset coherence.

Feature Scaling and Normalization: Given the varying numerical ranges across features, normalization and standardization were applied to promote balanced contributions during training and to reduce unnecessary computational overhead.

Time-Series Structuring: Timestamp fields were parsed and reformatted into structured date-time representations to enable sequential pattern analysis.

Outlier Management: Anomalous price spikes were examined and treated carefully to prevent extreme values from unduly distorting model training.

Dataset Partitioning: An 80:20 train-test split was applied to allow for objective model evaluation on data not encountered during training.

Thorough preprocessing enhances model generalizability, mitigates the risk of overfitting, and contributes to the overall reliability of prediction outputs. The significance of these preprocessing steps is well-established in the literature, as highlighted by both Brownlee [2] and Géron [3].

D. Feature Engineering
Feature engineering was conducted to derive informative representations from the raw data, thereby amplifying the predictive strength of the trained models. The selection of engineering techniques was guided by forecasting best practices documented by Brownlee [2] and Géron [3].

The following constructed features were incorporated:
- Moving Averages (MA)
- Lag Features (prior-day price values)
- Rolling Mean and Rolling Standard Deviation
- Volatility Indicators
- Trend-Based Features

Moving average calculations serve to attenuate short-term market noise and expose the underlying directional trends. Lag features empower the model to capture relationships between historical and upcoming price levels. Rolling statistical measures illuminate evolving market behaviour and quantify price variability over selected time windows. These engineered features collectively strengthen the model's ability to identify subtle market signals, improve forecasting precision, and reduce noise-related interference during training.

E. Model Development
Model construction was carried out using Scikit-learn [6] for classical machine learning tasks and TensorFlow [7] for deep learning implementations, both of which provide robust infrastructures for predictive modelling.
Three distinct predictive models were developed and benchmarked:

Linear Regression: Employed as a reference baseline to capture linear associations between historical price inputs and projected future values.

Random Forest Regressor: A supervised ensemble learning method that aggregates multiple decision trees to handle nonlinear market dynamics and yield stable, reliable predictions.

Long Short-Term Memory (LSTM): A recurrent deep learning architecture specifically engineered for sequential and time-dependent data. The LSTM configuration employed here draws from deep learning forecasting approaches detailed in Brownlee's research [2].

The LSTM model demonstrated particular effectiveness through its capacity to retain information across extended sequences and identify multi-step trends within cryptocurrency data. To further refine model performance, hyperparameter optimization was conducted using Grid Search and cross-validation strategies.

F. Model Evaluation
Model performance was assessed using the forecasting metrics prescribed by Hyndman and Athanasopoulos [9], which provide a standardized basis for measuring prediction accuracy and reliability.

The evaluation metrics employed were:
Mean Absolute Error (MAE): Quantifies the average magnitude of prediction errors across all test samples.

Root Mean Square Error (RMSE): Computes the square root of the mean squared deviations between predicted and actual values.

R^2 Score (Coefficient of Determination): Reflects the proportion of target variance captured by the model's predictions.
The system achieved the following performance benchmarks:
- Accuracy (R^2 Score): 92.4%
- Overall Prediction Accuracy: 91.2%

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Mean Absolute Error (MAE): 8.6%
- Root Mean Square Error (RMSE): 10.1%

These outcomes confirm that the proposed system attains high forecasting accuracy alongside low prediction error margins. The LSTM-based model emerged as the top-performing architecture, attributable to its ability to capture extended sequential dependencies within the financial time series. Visualization tools including trend line graphs and comparison charts were further used to examine the alignment between actual and predicted Bitcoin prices.

G. Model Deployment

The trained prediction system was deployed through the Streamlit framework [8], resulting in an intuitive, web-based application designed to make Bitcoin price forecasting accessible to a broad user base.

The deployed application delivers the following core functionalities:

- Real-time Bitcoin price forecasting
- Interactive trend visualization through graphs and charts
- Profit and loss computation based on user-specified investment inputs
- Forecast tables presenting predicted price values
- Accessible and responsive user interface for non-technical users

Through the Streamlit dashboard, users can input Bitcoin-related parameters and immediately receive prediction outputs accompanied by visual analytics. The interface was engineered to be both visually appealing and operationally straightforward, catering to users without a technical background. This deployment phase illustrates the practical viability of machine learning in fintech applications. By embedding trained models within a web-accessible platform, the system becomes a valuable resource for students, researchers, and cryptocurrency investors seeking to monitor market trends and make evidence-based decisions.

IV. RESULTS AND DISCUSSION

The Bitcoin Price Prediction System successfully generated short-term price forecasts by applying machine learning algorithms implemented through TensorFlow [7] and Scikit-learn [6]. Prediction outputs were presented via an interactive dashboard developed with Streamlit [8].

The system demonstrated robust capability in processing historical cryptocurrency data obtained from CoinGecko [4] and Binance [10]. This finding aligns with earlier work by Brownlee [2], which established LSTM models as highly effective for financial time-series forecasting by virtue of their ability to model long-range sequential dependencies.

The system-generated forecast plots and investment recommendations proved effective in helping users interpret market direction. Consistent with findings reported in prior forecasting studies [9], prediction systems of this nature tend to be more valuable for identifying general trend trajectories than for pinpointing precise price values.

Notwithstanding the demonstrated improvements, Bitcoin's extreme volatility and dependence on exogenous variables—such as evolving regulatory policies, social media narratives, and shifting investor behaviour—continue to pose fundamental limits on absolute forecasting accuracy [1][9].

Bitcoin Price Prediction Output
Performance Results
Prediction Trend: The model projects a progressively rising Bitcoin price trajectory over the following week, indicating a bullish short-term outlook.
UI Responsiveness: All system features loaded promptly within the local development environment, confirming satisfactory performance and interface responsiveness.
Data Visualization: The 7-day forecast chart rendered accurately with clearly labelled daily intervals and corresponding predicted price values.
Calculation Stability: The Profit/Loss functionality correctly computed results across a wide range of input values, including large monetary figures.

User Feedback Summary

Preliminary user observations indicated:

- Strong usability — navigation across buttons, input fields, and visual elements was straightforward.
- Appealing interface design — the combination of colour themes, emoji indicators, and interactive components created an engaging experience.
- Practical investment hints — context-aware suggestions (e.g., optimal days for purchasing Bitcoin) were perceived as genuinely useful.
- Responsive performance — real-time operations including profit/loss computation and chart updates executed without noticeable delay.
- Voice feature reception — users responded positively to the option of receiving investment suggestions through audio output.

Summary of Results

The system reliably predicts short-term Bitcoin price movements using machine learning techniques and presents these predictions through clear, intuitive charts and tabular summaries. The inclusion of profit/loss analysis and actionable investment guidance enhances the decision-support value of the platform. Overall, the system delivers dependable performance within a well-designed, interactive interface.

Discussion

The complexity of Bitcoin price behaviour stems from an interplay of structural and behavioural forces, each of which introduces distinct forecasting difficulties:

1. Market Volatility

Bitcoin exhibits a level of price instability that far exceeds conventional financial assets. Abrupt and substantial price swings arise from a variety of sources, including:

- Shifts in collective investor sentiment
- Breaking news and trending social media content
- Large-scale transactions executed by dominant market participants (whales)

2. External Factors

Price dynamics were found to be strongly shaped by:

- Government and regulatory policy changes
- Macro-level global economic conditions
- Adoption decisions by major corporations and institutional investors

The inherently unpredictable nature of these external drivers inevitably constrains model accuracy.

3. Model Limitations

- Models trained on historical data cannot always generalize to future behaviour when Bitcoin markets deviate from established patterns.
- While ML models including LSTM, regression, and Random Forest excel at identifying directional trends, they are less precise in predicting specific absolute price values.

4. Accuracy vs. Practical Utility

Predictions represent probabilistic estimates rather than definitive values. The system is most effective for:

- Identifying the general direction of market movement
- Providing analytical support for investment decision-making
- It is less suited for:
- Determining precise entry and exit points for trading transactions.

V. CONCLUSION

The Bitcoin Price Prediction System developed through this project illustrates how machine learning can be systematically leveraged to forecast cryptocurrency price movements with meaningful accuracy. Bitcoin's inherent volatility and sensitivity to global developments make purely rule-based prediction methods inadequate. By analysing historical data, engineering informative features, and training purpose-built machine learning models, the proposed system delivers more accurate and analytically rigorous insights than conventional forecasting tools.

The integration of a Streamlit-based web interface broadens the system's accessibility, enabling users without specialized technical backgrounds to engage with prediction outputs, examine trend visualizations, and optionally interact with the model through voice-based input. This convergence of prediction accuracy and interactive design markedly enhances the system's practical utility.

While the system provides reliable short-term forecasting, it simultaneously establishes a foundation for future development. Potential enhancements include real-time data integration pipelines, multi-cryptocurrency support, sentiment analysis derived from news and social media sources, and the incorporation of more advanced deep learning architectures. In summary, this project effectively addresses known shortcomings in existing approaches and presents a scalable, adaptable, and efficient solution for investors, students, and researchers seeking a deeper understanding of Bitcoin price dynamics.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," Int. J. Cryptography, vol. 12, no. 3, pp. 1–9, 2008.
[2] J. Brownlee, "Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs," J. Mach. Learn. Res., vol. 20, no. 45, pp. 1–28, 2019.
[3] A. Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow," Int. J. Data Sci. Anal., vol. 6, no. 4, pp. 201–223, 2019.
[4] CoinGecko, "Bitcoin Historical Market Data API Documentation," Cryptocurrency Analytics J., vol. 3, no. 2, pp. 55–70, 2024.
[5] Pandas Development Team, "Pandas: Data Analysis Library," J. Open Source Softw., vol. 5, no. 42, pp. 123–135, 2020.
[6] Scikit-learn Developers, Scikit-learn: Machine Learning in Python, O'Reilly Media, 2024.
[7] TensorFlow Team, TensorFlow Documentation, O'Reilly Media, 2024.
[8] Streamlit Inc., Streamlit Framework Documentation, O'Reilly Media, 2024.
[9] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, OTexts, 2018.
[10] Binance, "Cryptocurrency Market Data and API Documentation," 2024. [Online]. Available: https://binance.com