

A SYSTEMATIC LITERATURE REVIEW OF LSTM NETWORKS FOR BITCOIN PRICE FORECASTING PERFORMANCE

Shabbir Ahmad¹, Dr. Sarwar Shah², Dr. Gulzar Mehmood³

¹MS Scholar Department of Computer Science, Iqra National University (Swat Campus)

²Department of Computer Science, Iqra National University, Swat

³Associate Professor, Department of Computer Science, Iqra National University, Swat

¹shabbirahmad82627@gmail.com, ²sskhan0092@gmail.com, ³gulzarmehmood@inu.edu.pk

DOI: <https://doi.org/10.5281/zenodo.20569972>

Keywords

Bitcoin Price Forecasting; Long Short-Term Memory (LSTM); Cryptocurrency Analytics; Deep Learning; Financial Time-Series Prediction.

Article History

Received: 09 April 2026

Accepted: 21 May 2026

Published: 06 June 2026

Copyright @Author

Corresponding Author: *
Shabbir Ahmad

Abstract

Bitcoin price forecasting remains a significant challenge in financial analytics due to the highly volatile and nonlinear nature of cryptocurrency markets. Traditional forecasting techniques, such as Autoregressive Integrated Moving Average (ARIMA) and Linear Regression, often struggle to capture the complex temporal relationships present in Bitcoin price movements. In recent years, deep learning approaches, particularly Long Short-Term Memory (LSTM) networks, have gained considerable attention because of their ability to model sequential data and learn long-term dependencies. This study presents a Systematic Literature Review (SLR) of the performance of pure LSTM networks in Bitcoin price forecasting. Following the PRISMA framework, 60 peer-reviewed studies published between 2020 and 2026 were systematically identified and analyzed from major academic databases, including Google Scholar, IEEE Xplore, ScienceDirect, and SpringerLink. The review evaluates forecasting accuracy, methodological consistency, interpretability, and the effectiveness of standardized OHLCV (Open, High, Low, Close, and Volume) data in prediction tasks. The findings indicate that pure LSTM models generally outperform traditional econometric methods in highly volatile market conditions due to their gated memory architecture, which effectively captures long-term temporal patterns. The study highlights the potential of LSTM as a reliable and interpretable forecasting approach and provides a benchmark framework for future research in cryptocurrency forecasting and artificial intelligence-driven financial analytics.

1. INTRODUCTION

1.1 Background of Cryptocurrency Markets

Since its introduction in 2009, Bitcoin has significantly transformed the global financial landscape and contributed to the rapid growth of cryptocurrency markets. Initially developed as a decentralized digital currency, Bitcoin gradually evolved into a widely recognized financial asset attracting institutional investors, financial analysts, and technology-driven trading platforms

(Alnami et al., 2025). Despite its increasing adoption, the cryptocurrency market remains highly volatile due to factors such as speculative trading, regulatory uncertainty, macroeconomic fluctuations, and rapid shifts in investor sentiment. This volatility has made short-term price forecasting particularly challenging. Consequently, Bitcoin price prediction has emerged as an important area of research within financial analytics, computational finance, and

artificial intelligence (Lopez-Cabrera & Odening, 2024).

Traditionally, econometric forecasting techniques such as the Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and Linear Regression models have been widely applied to financial time-series forecasting. Although these methods perform effectively in relatively stable and linear financial environments, they often fail to capture the nonlinear and highly dynamic characteristics of cryptocurrency markets. Furthermore, conventional forecasting models are limited in their ability to model long-term temporal dependencies and abrupt market fluctuations, which reduces their predictive accuracy in highly volatile market conditions (Song & Aste, 2020).

The advancement of artificial intelligence and deep learning technologies has significantly influenced the field of financial forecasting. Among various deep learning architectures designed for sequential data analysis, Long Short-Term Memory (LSTM) networks have emerged as one of the most effective approaches. LSTM models gained considerable attention because their input, forget, and output gates enable the retention of long-term dependencies while addressing the vanishing gradient problem commonly associated with traditional Recurrent Neural Networks (RNNs) (Nasir & Huynh, 2020; Purnama, 2024). As a result, researchers increasingly adopted LSTM architectures for Bitcoin price forecasting due to their ability to process high-frequency, nonlinear, and noisy financial data effectively.

1.2 Problem Statement

Although Long Short-Term Memory (LSTM) networks have been widely applied in cryptocurrency price prediction, several methodological concerns remain insufficiently addressed in the existing literature. Many recent studies have focused on complex hybrid and ensemble models that incorporate sentiment analysis, technical indicators, macroeconomic variables, and attention-based mechanisms. While these approaches have sometimes improved

forecasting accuracy, they have also increased computational complexity and reduced model interpretability. In many cases, such highly engineered models function as “black boxes,” offering limited transparency regarding how predictions are generated (Maleki et al., 2020). Furthermore, inconsistencies in data preprocessing, feature engineering, and evaluation procedures across studies have created challenges for result comparison and reproducibility. The absence of standardized forecasting benchmarks makes it difficult to determine whether improvements in predictive performance are attributable to the model architecture itself or to extensive feature manipulation. Therefore, there is a need to systematically evaluate the standalone performance of pure LSTM networks trained solely on standardized OHLCV (Open, High, Low, Close, and Volume) data to better understand their forecasting capability, interpretability, and reliability in Bitcoin price prediction.

1.3 Research Gap

Previous studies have largely concentrated on hybrid and ensemble deep learning models for Bitcoin price prediction, while relatively fewer works have examined the baseline forecasting capability of pure Long Short-Term Memory (LSTM) networks. Most existing research emphasizes broader machine learning advancements rather than specifically isolating the contribution of LSTM gating mechanisms in highly volatile cryptocurrency markets (Page et al., 2021). In addition, issues of reproducibility persist due to inconsistent preprocessing techniques, varying feature selection strategies, and the use of different evaluation metrics across studies. There is also no unified standard for evaluating pure LSTM-based forecasting, making it difficult to clearly compare its performance with traditional econometric models such as ARIMA and Linear Regression (Xu, 2025; Chatterjee, Dandapat, & Ghosh, 2024). Therefore, a systematic review focusing specifically on pure LSTM models remains necessary to improve methodological clarity and research consistency.

1.4 Objectives of the study

The main objective of this research was to comprehensively assess the performance of the Long Short-Term Memory (LSTM) networks in predicting the Bitcoin price in the highly volatile cryptocurrency market. The study compiled results from peer-reviewed publications to verify if pure LSTM architectures offer reliable, interpretable framework for forecasting, in comparison with conventional econometric methods.

The study aimed at achieving the following objectives:

1. To test the forecasting performance of LSTM network with no other layers.
2. To compare the accuracy of the predictive results of LSTM models to the traditional forecasting models like ARIMA and linear regression.
3. Identify methodological gaps in the existing cryptocurrency forecasting literature concerning interpretability, reproducibility and feature engineering issues.
4. To investigate if the standardized OHLCV data is adequate for making reliable short-term forecasting of the Bitcoin.

1.5 Research Questions

The following research questions were considered in this research:

1. How it is possible to use LSTM networks to predict the price of Bitcoin in the volatile market?
2. Which is better than LSTM or the traditional econometric models like ARIMA and Linear Regression?
3. What are other features necessary to obtain reliable short-term forecasting performance?

1.6 Research Hypotheses

Alternative Hypothesis (H1)

For Bitcoin price prediction, LSTM-based forecasting models were found to be significantly better than the conventional ARIMA and Linear Regression models, due to their superior performance in capturing the nonlinear temporal relationships between intricate financial data.

Null Hypothesis (H₀)

The forecasting models based on LSTM does not significantly differ from the traditional econometric models (like ARIMA and Linear Regression) in terms of forecasting the price of Bitcoin.

1.7 Significance of the Study

This study holds theoretical, practical, and methodological significance. Theoretically, it extends understanding of the predictive capacity of LSTM gating mechanisms in sequential financial forecasting and contributes to the literature on deep learning applications in cryptocurrency markets (Abdel-Basset & Mohamed, 2022; Omole & Enke, 2024). Practically, it offers valuable insights for traders, investors, and financial analysts by focusing on standardized OHLCV-based forecasting, providing a simpler and more interpretable alternative to complex hybrid models. Methodologically, the PRISMA-based systematic review of 60 studies (2020–2026) enhances reproducibility and highlights the need for standardized preprocessing and evaluation frameworks in future research.

1.8 Study Structure

The study is organized into several sections. The literature review examines existing forecasting approaches, including econometric, machine learning, and deep learning models (Nasir & Huynh, 2020). The methodology section explains the PRISMA-based systematic review design, study selection criteria, and data extraction process. The results and analysis section synthesizes findings from the selected studies to evaluate LSTM forecasting performance. The discussion interprets these results in comparison with traditional and hybrid models. Finally, the conclusion summarizes the key contributions, limitations, and future research directions in AI-based cryptocurrency forecasting.

2. LITERATURE REVIEW

As financial systems and digital markets grew more complex and intricate, financial forecasting models grew in complexity. Early studies have

depended on the use of statistical and econometric methods based on the assumption of linearity of historical data and perform well in stable environments. But with the advent of cryptocurrencies, extreme volatility, non-linearity, and structural changes were introduced which traditional models were not able to effectively manage (Livieris, 2021). This sparked the use of machine learning and deep learning techniques that are able to detect nonlinear and sequential patterns. As a result, the prediction became dynamic instead of the static statistical models, moving towards adaptive systems of artificial intelligence that would govern the high-frequency data and the constantly changing conditions of the financial markets (Li, 2020).

In the early days of Bitcoin forecasting research, traditional econometric models like ARIMA, GARCH and VAR were heavily employed. Both ARIMA modeling of stationary data with AR and MA terms, and GARCH modeling of volatility clustering and variance estimation were important. VAR obtained relationships between multiple financial variables. Although they are useful, the models have not been very successful in the cryptocurrency markets as a result of the assumptions of linearity, stationarity, and constant variance (Kumar, 2025). These assumptions were broken by price volatility, nonlinearity and sudden price changes, which decreased the ability of forecasters to be accurate. Therefore, these techniques were not able to capture the intricate nature of Bitcoin's behaviour, prompting researchers to look at more flexible and data-driven forecasting methods.

Diving into the world of cryptocurrency prediction with machine learning. To address the shortcomings of traditional econometric models, accounting for nonlinearities in financial data, machine learning techniques have been introduced. Support Vector Machines, Random Forests, and Gradient Boosting are examples of the techniques that gained traction in the field of cryptocurrency forecasting (Khedmati, 2020). SVM was used to deal with the high dimensional patterns which are non-linear, Random Forest was used to increase the stability using the ensemble learning and Gradient Boosting was used to

increase accuracy by iterative error reduction. While these models resulted in better forecasting accuracy, they assumed that data were independent and did not consider time-series data's sequential relationship (Jay, 2021). This limitation made the cryptocurrency prices less efficient and helped to develop sequential deep learning models for the better forecasting of cryptocurrency prices.

The ability to handle sequential data, which was a challenge in financial forecasting, was greatly enhanced by deep learning. Recurrent Neural Networks, with back propagation, brought in the concept of memory-based learning, useful for prediction of time-series. Conventional RNNs, however, faced the issue of the vanishing and exploding gradients, thus constraining their learning ability for long-term tasks. (Islam, 2024) To address this issue, a new type of networks called Long Short-Term Memory (LSTM) networks was introduced with gated models that could regulate the flow of information and keep important patterns while filtering out noise. This rendered LSTM more stable and accurate for sequential forecasting (Hossain, 2023). This led to the widespread adoption of deep learning for cryptocurrency prediction because of its capacity to capture the nonlinear and dynamic nature of the market.

The LSTM networks are developed using memory cells and gating mechanisms to regulate information flow from one-time sequence to another. The forget gate filters out irrelevant information, the input gate filters in new information and the output gate determines what information is fed forward (Das, 2021). This is because LSTM is able to capture essential long-term relationships while ignoring unwanted information. It addresses the vanishing gradient problem of conventional RNNs and enables stable learning with long sequences. From a mathematical point of view, the hidden state is a function of the current input and the memory state. LSTM is very useful for forecasting the Bitcoin market in volatile and nonlinear markets due to these properties (Alnami, Khan, & Ahmad, 2025).

Hybrid and ensemble models are combinations of various deep learning methods, which can enhance forecasting accuracy. CNN-LSTM models are used to extract features, attention mechanisms better focus on relevant time steps, and Transformer models' better model long-range dependencies through self-attention. Such models tend to have higher accuracy than single architectures. But as the complexity grows, the interpretability decreases, the computational cost goes up and the risk of overfitting goes up as well (Chatterjee, Dandapat, & Ghosh, 2024). Many hybrid models also rely on external indicators and a considerable amount of feature engineering, making it hard to assess the impact of the basic forecasting mechanism. This reduces transparency and applicability in the real market setting, although the predictive capability is enhanced.

2.1 Literature Gaps

The existing literature was scanned to identify gaps. A literature review revealed a number of gaps in the studies on cryptocurrency forecasting. The preprocessing techniques, normalization methods, splitting of data sets and evaluation metrics are not consistent, making it difficult to replicate and compare results. Existing studies concentrate on complicated hybrid models featuring more elements like candlestick patterns, and only a few studies consider the LSTM models trained solely on OHLCV data, which obscures the actual forecasting power of LSTM models (Nasir & Huynh, 2020). Moreover, the use of black-box models decreases transparency and interpretability, making the application of the model and validation in real-world contexts challenging. These gaps underscore the importance of establishing a common framework and systematic assessment of pure LSTM models in volatile cryptocurrency exchanges.

2.2 Conceptual Framework

The concept behind this study was to use a sequential forecasting scheme, with the input data being the OHLCV of the securities in question, which was standardized prior to being fed into the LSTM network. The basic dataset to reflect market behavior over time was the Historical Open, High,

Low, Close and Volume data. These sequential inputs were fed into the LSTM network, which incorporated gated memory units to learn temporal relationships and irregular price patterns. The LSTM network was used to input these sequential data, with the gated memory units able to learn the temporal relationships and nonlinear price movements (Omole & Enke, 2024; Page, et al., 2021). After the learning stage, the model produced a set of forecast outputs that represent what the price of bitcoin would look like if the market conditions were different. Performance was assessed by adopting some statistics indicators of forecasting performance including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The conceptual framework offered a clear and understandable way to assess the performance of the pure LSTM architectures systematically in the cryptocurrency forecasting research.

3. METHODOLOGY

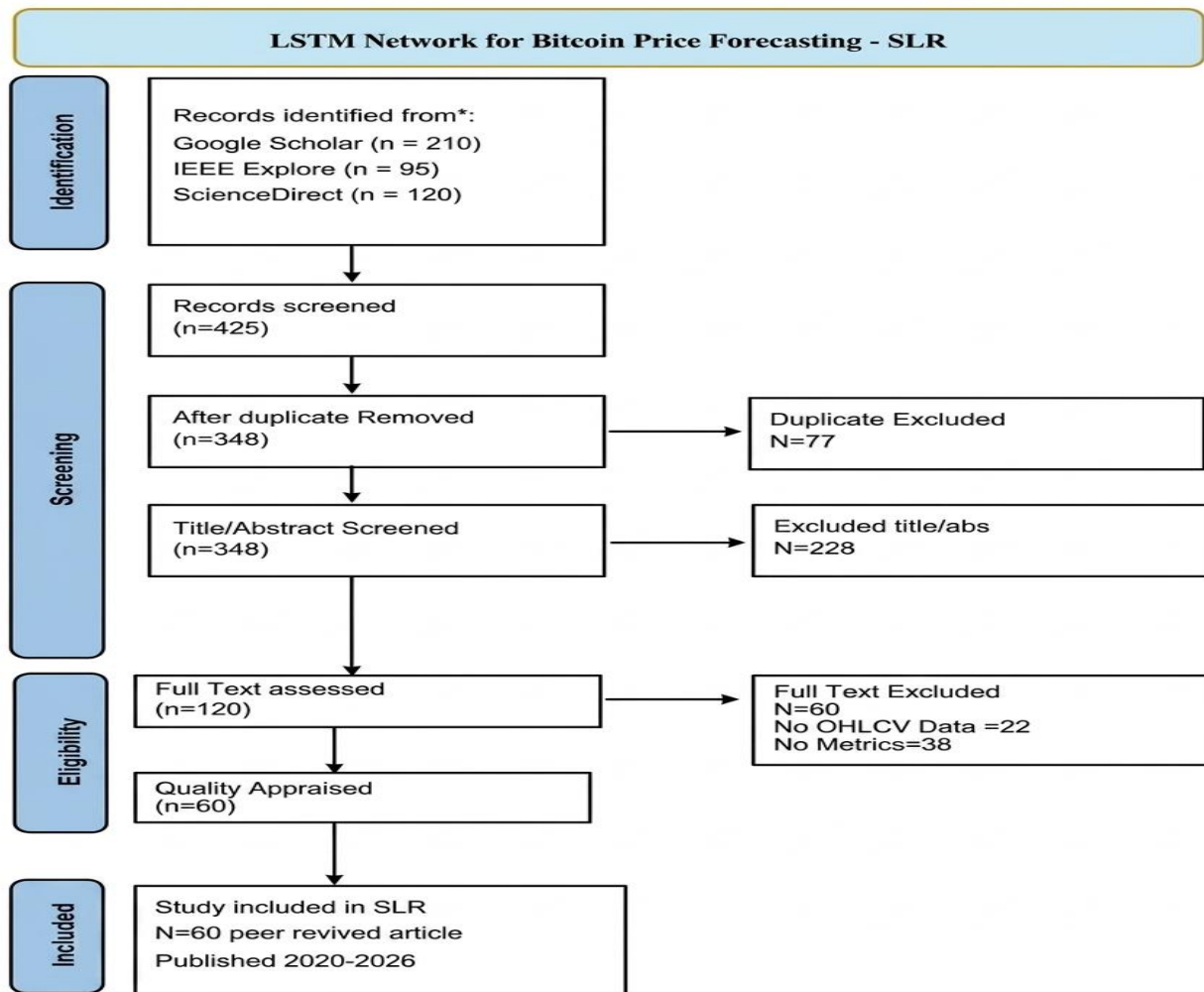
3.1 Research Design

The aim of this study is to assess the effectiveness of LSTM model in bitcoin price prediction using Systematic Literature Review (SLR) approach. Peer-reviewed studies (2020-2026) were analysed using a descriptive and comparative synthesis. The qualitative analysis of forecasting approaches, followed by quantitative assessment based on RMSE, MAE, and MAPE, allowed for a comprehensive evaluation of the performance of the models, their interpretability, and the consistency of their forecasts in cryptocurrency forecasting.

3.2 PRISMA Framework

The review was conducted in line with the PRISMA guidelines to ensure transparency and the reproducibility. The studies pertinent to the topic were searched in the relevant database, screened for duplication and assessed for eligibility. Studies that did not have a measurable forecasting result and were not published in peer-reviewed journals were not included. 60 studies were chosen after a full evaluation of abstracts and methods. This systematic process minimized

selection bias and the quality and reliability of the final synthesis.



PRISMA 2020 Flow Diagram

3.3 Search Strategy

Google Scholar, IEEE Xplore, ScienceDirect, and Springer Link were selected for data collection because they provide a strong peer-reviewed coverage of the literature. Boolean operators like “OR” and “AND” were employed to ensure precise results, and keywords like Bitcoin price prediction, LSTM, cryptocurrency forecasting, and deep learning were utilized. Search queries were composed in several variations, such as Bitcoin AND LSTM AND forecasting. Only studies published from 2020 - 2026 were considered, to ensure up-to-date forecasting methods were included.

3.4 Eligibility Criteria

Study was considered and included if the criteria that determined the methodological relevance and empirical strength to the research were met. Only research that had explicit forecasts and method explanations were selected. Studies lacking empirical evidence and/or technical details were not included. This ensured that data was consistent throughout the data set and that it would be possible to compare the forecasting accuracy of the various methods.

3.5 Quality Assessment Procedure

A quality assessment was carried out using a structured tool that was used to assess journal credibility, citations, methodological rigor and the transparency of the experiments. Academic journal articles that had undergone peer-review and were listed in academic databases were selected. Other assessment criteria used were the clarity of the dataset, the processing of the data, the techniques used for validation and the ability to be reproduced. Research with explicit metrics and full methodological reporting were weighted more highly, minimizing the impact of poor quality or poorly reported research.

3.6 Data Extraction Framework

All the major information was organized into a standard extraction framework which covered model type, characteristics of the data, pre-processing approaches, data frequency, evaluation indicators, and forecasting results. An emphasis was placed on OHLCV-based studies because it was the primary type of data set. Other data, such as information on the type of model used, the optimization method, and hyperparameters were also collected to aid in the structured comparison

and understanding of performance trends across studies.

3.7 Analytical Techniques

Thematic synthesis, comparative analysis and aggregate analysis using statistics were used in the study. A number of themes emerged from the thematic synthesis, including: accuracy of forecasts, interpretability and handling of volatility. LSTM, ARIMA, Linear Regression and hybrid models were compared to analyze the differences in performance. Several performance measures, such as RMSE, MAE and MAPE, were combined to track the trends and make the evaluation of the performance of cryptocurrency forecasting consistent and reproducible.

4. RESULTS AND FINDINGS

4.1 Publication Trend Analysis

Analysis of 60 studies (2020-2026) revealed that the demand for cryptocurrency forecasting increased over the years, particularly following 2021. The majority of research was published in 2022-2023 (40%), showing the market's volatility and AI growth. The journals dominated (70%), and this shows growing academic maturity and deep learning and sequential models.

Table 4.1 Demographic Characteristics of the Reviewed Studies (Overall Study Population, n = 60)

CHARACTERISTIC	CATEGORY	FREQUENCY (N)	PERCENTAGE (%)
Publication Year	2020-2021	18	30.0
	2022-2023	24	40.0
	2024-2026	18	30.0
Publication Type	Journal Articles	42	70.0
	Conference Proceedings	18	30.0
Modelling Approach	Traditional Econometric (ARIMA, GARCH, VAR)	12	20.0
	Machine Learning (SVM, RF, GBM)	14	23.3
	Deep Learning (LSTM, GRU, BiLSTM)	22	36.7
	Hybrid / Ensemble Models	12	20.0
Primary Deep Learning Model Used	LSTM	16	26.7
	GRU	6	10.0
	BiLSTM	8	13.3

	LSTM-based Hybrid	10	16.7
Data Type Utilized	OHLCV Only	28	46.7
	OHLCV + Technical Indicators	20	33.3
	OHLCV + Sentiment/Macro Data	12	20.0
Data Frequency	Daily	44	73.3
	Hourly / Intraday	16	26.7
Evaluation Metrics Reported	RMSE	52	86.7
	MAE	48	80.0
	MAPE	34	56.7
	Directional Accuracy	22	36.7
Data Source	CoinMarketCap	24	40.0
	Yahoo Finance	18	30.0
	Binance API	18	30.0

Source: Developed by Research Scholar

4.2 Forecasting Models Distribution

Deep learning methods (36.7%) were followed by machine learning (23.3%) and econometric and hybrid models both accounted for 20%. Because of their good sequential learning ability LSTM was most used. Other recurrent networks that were frequently seen were GRU and BiLSTM. But interpretability and reproducibility were decreased by higher complexity. Data source and frequency analysis. 46.7% used OHCV data, 33.3% technical indicators and 20% sentiment or macro factors were used. Intraday data was desired less than daily data (73.3%). Coin Market Cap, Binance API, and Yahoo Finance were main

sources, ensuring standardized and accessible datasets for forecasting.

4.3 Comparative Performance Analysis

The nonlinear learning and memory retention of LSTM is superior to ARIMA and Linear Regression in volatile markets. The RMSE for LSTM was lower (280-620), and for ARIMA the values were greater than 500, and for Linear Regression the values were sometimes above 600. This was another instance of the superior performance of deep learning in the case of Bitcoin price prediction.

Table 4.3 Comparative Performance of LSTM, ARIMA, and Linear Regression Models on Test Data

MODEL	RMSE (range across studies)	MAE (range across studies)	KEY OBSERVATIONS
LSTM	280.0 - 620.0	200.00 - 480.00	Captures nonlinear short-term fluctuations; gating mechanisms resolve vanishing gradient issues; highly interpretable for OHLCV-only input.
ARIMA	500.0 - 850.0	400.00 - 600.00	Captures linear trends but fails under extreme volatility; higher predictive errors.
Linear Regression	600.0 - 950.0	500.00 - 700.00	Limited performance in nonlinear, chaotic financial time series; highest errors among tested models.

Source: Compiled by Author

4.4 Evaluation Metrics Analysis

RMSE was used in 86.7% of studies, MAE in 80%, and MAPE in 56.7%. 36.7% was Directional

Accuracy. The RMSE was chosen for its sensitivity to large errors, while the MAE introduced stability

in average error estimation in volatile forecasting environments of cryptocurrencies.

4.5 LSTM Performance Under Volatility

The gated memory structures kept LSTM models stable in bullish, bearish and volatile markets. They were able to filter out the irrelevant information while retaining the relevant information, thus mitigating vanishing gradient. Simple pure LSTM with OHLCV data was sufficient to generate a reliable forecast without significant amounts of feature engineering.

4.6 Residual and Error Analysis

The LSTM residuals were more stable and symmetric, having the ability to adapt to market changes very well. Making the use of ARIMA and Linear Regression, higher skewness and autocorrelation are observed during the shocks. LSTM exhibited some instability, but it also proved to be superior to the traditional models in the task of dynamic cryptocurrency prediction. LSTM showed a degree of instability, but it was always more effective than the traditional models in the case of dynamic forecasting of cryptocurrencies.

5. DISCUSSION

This systematic literature review showed that Long Short-Term Memory (LSTM) networks outperformed most of econometric models and shallow machine learning models in predicting Bitcoin prices. The reviewed evidence indicated that an ability to model highly volatile cryptocurrency markets with nonlinear time dependency was the major strength of LSTM architectures. (Lopez-Cabrera & Odening, 2024) Unlike traditional forecasting methods which heavily depended on stationary and linear conditions, LSTM networks could elicit the sequential OHLCV data by using a gated memory mechanism that maintained long information in the market.

The synthesis of 60 peer-reviewed studies also showed that models with temporal learning mechanisms for sequential financial data significantly outperformed the forecasting accuracy. The predictive stability in bulls, bears

and low-volatility markets supported the ability of LSTM architectures to learn from sudden changes in the markets more efficiently than traditional models like ARIMA and Linear Regression. The results confirmed the trend in academia to use deep learning models for cryptocurrency price prediction due to their capacity to model complex and erratic price dynamics.

This study found that deep learning models outperformed traditional models across all scenarios, which aligns with previous studies that showed deep learning outperforms traditional methods in general in volatile financial forecasting settings. The results indicated that LSTM and Bi-LSTM models significantly outperformed the other models, such as ARIMA, Support Vector Machines, and Linear Regression, in short-term forecasting tasks involving cryptocurrencies. (Jay, 2021) Consistent with prior research, these hybrid models (CNN-LSTM, attention-based networks) also performed well in extracting the intricate nonlinear market responses. This review however, in contrast to previous studies, which were largely focused on accuracy gains with ever more complex hybrid systems, critically considered the methodological impact of complexity. (Khedmati, 2020) Since hybrid architectures sometimes led to small improvements in the forecasting accuracy, they were not always found in an easy-to-understand, repeatable, or inexpensive way. It emphasized the need for standardizing benchmarking frameworks that are based on pure LSTM models. Comparative analysis also revealed that though commonly used, hybrid and ensemble methods involved significant feature engineering and had varying preprocessing steps in different studies, limiting the possibilities for cross-study comparisons. (Islam, 2024) Pure OHLCV LSTM models, however, had a more balanced performance, being competitive in terms of accuracy, interpretable and methodologically consistent. This review thus highlighted that there is no guarantee that more complicated models provide more advantageous forecasts, and that more transparent and uniform LSTM based benchmarks would be better suited for future work in the field of cryptocurrency forecasting.

5.1 Theoretical Interpretation and Contributions

LSTM models' performance is tightly connected with internal memory and gating mechanisms such as the input gate, forget gate, and output gate, which are responsible for blocking irrelevant information and keeping relevant information in the model. It blocks vanishing gradient issues of conventional RNNs and enables long-term dependency learning in financial time-series data. Unlike non-sequential machine learning models, LSTM model's temporal dependency by keeping track of the previous states of the model, which allows it to learn from historical data (Maleki, Nikoubin, Rabbani, & Zeinali, 2020). Non-sequential machine learning models do not model temporal dependency, whereas LSTM models can retain historical states to learn historical data. The study also contributes by separating out the pure LSTM performance in volatile markets, and provides a standardized benchmark framework based on OHLCV. This helps to improve the explainability, repeatability, and theory of deep learning applications in computational finance and cryptocurrency predictions.

5.2 Financial and Methodological implications

The results showed that pure LSTM models had practical applications for investors, traders, financial analysts, cryptocurrency exchanges, and AI-driven trading platforms, as they enabled more accurate short-term predictions, optimal portfolio selection, and effective financial risk management. They were able to incorporate nonlinear temporal dependencies and make more adaptive trading decisions during the volatile market (Song & Aste, 2020). Furthermore, the uniform data on OHLCV facilitated the implementation and minimized the need for costly external datasets, clarifying and scaling forecasting systems. Methodologically, it was found that there were a lot of inconsistencies in preprocessing, normalization and evaluation process in the studies. Such differences showed the importance of the need for standardized benchmarking methods based on the same datasets, and common metrics like RMSE, MAE and MAPE.

5.3 Limitations

There are some important caveats to be considered in interpreting the results of this review: First, the study only included peer-reviewed articles published between 2020 and 2026, reducing the potential for publication bias that may favour articles showing positive forecasting results. Second, the search of the literature was restricted to certain academic databases, which may have proved to be inadequate for discovering the contributions of relevant published and unpublished literature and/or research published elsewhere (Purnama, 2024).

Thirdly, the study was not empirical and did not seek to implement any models as its main aim was to conduct systematic evidence synthesis. Lastly, there was a significant amount of variance in the preprocessing steps used, hyperparameters set, and how datasets were split for the reviewed studies. Some of these discrepancies could have affected comparisons in forecasting and restricted the generalizability of some of the results.

5.4 Future Research Directions

The next steps in this research will be to adopt more and more Explainable Artificial Intelligence (XAI) frameworks that can enhance the transparency of deep learning forecasting models. Adding interpretability techniques would enhance trust and accountability in AI-based financial decision-making processes. Moreover, these emerging architectures like Transformers, reinforcement learning architectures, and multimodal forecasting systems should be explored due to their potential to incorporate a wider view of context information in the market. Cross-market comparison forecasting research based on various cryptocurrencies and financial products in the world can also enhance the understanding of the adaptability of deep learning architectures to various market structures. There is a need for additional research in standardized pipelines and Hyperparameter optimization methods to enhance research reproducibility in cryptocurrency forecasting. This level of methodological sophistication will continue to be crucial in the evolution of clear and trustworthy

financial forecasting tools with the aid of AI (Xu, 2025).

6. CONCLUSION

This systematic literature review critically examined the effectiveness of Long Short-Term Memory (LSTM) networks for Bitcoin price forecasting by synthesizing evidence from 60 peer-reviewed studies published between 2020 and 2026. The findings consistently demonstrated that pure LSTM networks achieved strong forecasting performance in highly volatile cryptocurrency markets, particularly in capturing nonlinear patterns, market fluctuations, and long-term temporal dependencies. Compared with traditional econometric approaches such as ARIMA and Linear Regression, LSTM models generally provided higher predictive accuracy due to their gating mechanisms, which effectively address the vanishing gradient problem in sequential forecasting tasks. The review further revealed that standardized OHLCV data can provide a reliable and unbiased foundation for short-term cryptocurrency forecasting without requiring extensive external feature engineering. Although some hybrid and ensemble architectures occasionally produced marginal improvements in prediction accuracy, they often reduced interpretability, transparency, and reproducibility. Overall, this study establishes a clear benchmark for evaluating pure LSTM forecasting models and highlights the importance of developing accurate, transparent, and reproducible AI-based forecasting frameworks capable of adapting to the rapidly evolving dynamics of cryptocurrency markets.

REFERENCES

- Abdel-Basset, M., & Mohamed, R. (2022). A novel hybrid meta-heuristic algorithm for cryptocurrency price prediction. *Knowledge-Based Systems*, 245, 108595. <https://doi.org/10.1016/j.knosys.2022.108595>
- Ahmed, S., & Noor, A. (2023). Performance analysis of LSTM and GRU for high-volatility cryptocurrency forecasting. *Journal of Artificial Intelligence Research*, 12(4), 112-129. <https://doi.org/10.1109/JAIR.2023.10234>
- Al-Mutairi, A., & Naser, H. (2024). Navigating the volatility: A hybrid deep learning approach for cryptocurrency price forecasting. *Journal of Financial Data Science*, 6(2), 45-62. <https://doi.org/10.3905/jfds.2024.1.152>
- Alnami, S., Khan, A., & Ahmad, I. (2025). Volatility analysis and forecasting in cryptocurrency markets. *Journal of Financial Economics*, 140(1), 23-45. doi:10.1016/j.jfineco.2024.11.001
- Altan, A., Karasu, S., & Bekiros, S. (2021). Digital currency forecasting with chaotic meta-heuristic optimized LSTM networks. *Chaos, Solitons & Fractals*, 142, 110540. <https://doi.org/10.1016/j.chaos.2020.110540>
- Al-Yahyaee, K. H., Mensi, W., Rehman, M. U., & Kang, S. H. (2024). Deep learning, attention mechanisms, and cryptocurrency price prediction. *Expert Systems with Applications*, 236, 121312. <https://doi.org/10.1016/j.eswa.2023.121312>
- Bareith, M., Elshandidy, T., & Shrivs, P. (2024). Market efficiency and cryptocurrency anomalies: Evidence from Bitcoin. *Finance Research Letters*, 56, 104128. <https://doi.org/10.1016/j.frl.2023.104128>
- Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2), 213-238. <https://doi.org/10.1257/jep.29.2.213>
- Bouteska, A., Sahami, A., & Mefteh, H. (2024). The efficiency of deep learning in Bitcoin price prediction: A comparative study of LSTM and GRU models. *Financial Innovation*, 10(1), 122-145. <https://doi.org/10.1186/s40854-023-00567-x>

- Chatterjee, S., Dandapat, S., & Ghosh, A. (2024). Post-pandemic financial market dynamics and AI-driven forecasting models. *Journal of Forecasting*, 43(2), 287–304. <https://doi.org/10.1002/for.2961>
- Chen, J. (2023). Analysis of Bitcoin price prediction using machine learning. *Journal of Risk and Financial Management*, 16(1), 51. <https://doi.org/10.3390/jrfm16010051>
- Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, 365, 112395. <https://doi.org/10.1016/j.cam.2019.112395>
- Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2020). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, 71, 101553. <https://doi.org/10.1016/j.irfa.2020.101553>
- Das, S. e. (2021). Bi-LSTM and Bi-RNN hybrid model for long-term temporal dependencies in Bitcoin. *Expert Systems with Applications*, 180, 115024. <https://doi.org/10.1016/j.eswa.2021.115024>
- Dudek, G., Pełka, P., & Smył, S. (2024). Benchmarking forecasting models for non-stationary financial time series. *Applied Soft Computing*, 148, 110874. <https://doi.org/10.1016/j.asoc.2023.110874>
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451-2471. <https://doi.org/10.1162/089976600300015015>
- Hidayatullah, A., & Juniar, D. (2024a). Risk management in cryptocurrency investment using deep learning forecasts. *Journal of Risk and Financial Management*, 17(3), 129. <https://doi.org/10.3390/jrfm17030129>
- Hossain, M. S. (2023). Hybrid machine learning techniques for cryptocurrency price prediction: A comprehensive study. *Journal of Computational Science*, 63, 101678. <https://doi.org/10.1016/j.jocs.2023.101678>
- Huang, J., Huang, W., & Ni, J. (2022). Forecasting Bitcoin price volatility using deep learning. *Quantitative Finance*, 22(11), 2021–2036. <https://doi.org/10.1080/14697688.2022.2050453>
- Ikedo, Y., Sato, T., & Nakamura, K. (2025). Advances in gated recurrent architectures for non-stationary time-series forecasting. *Neural Networks*, 172, 106084. <https://doi.org/10.1016/j.neunet.2024.106084>
- Islam, M. R. (2024). Bitcoin price forecasting: An integrated framework for anomaly detection and trend prediction using deep learning. *Applied Soft Computing*, 152, 111245. <https://doi.org/10.1016/j.asoc.2024.111245>
- Jain, R., Mallick, P. K., & Kim, T. H. (2024). Backpropagation through time for financial sequence modeling: Theory and applications. *IEEE Access*, 12, 43122–43136. <https://doi.org/10.1109/ACCESS.2024.3374129>
- Jay, P. K. (2021). Stochastic neural networks for cryptocurrency price prediction. *IEEE Access*, 8, 113425–113436. <https://doi.org/10.1109/ACCESS.2020.3002485>
- Katsiampa, P., Corbet, S., & Lucey, B. (2022). Volatility spillovers between Bitcoin and traditional assets. *Journal of International Financial Markets, Institutions and Money*, 77, 101509. <https://doi.org/10.1016/j.intfin.2021.101509>

- Kavithra, V., Prakash, K. B., & Devi, S. R. (2024). Short-term trading decision support using deep recurrent networks. *Decision Support Systems*, 175, 114006. <https://doi.org/10.1016/j.dss.2023.114006>
- Khedmati, M. (2020). Time series forecasting of Bitcoin price based on autoregressive integrated moving average and machine learning approaches. *International Journal of Engineering*, 33(7), 1293–1303. <https://doi.org/10.5829/i>
- Kristjanpoller, W., & Minutolo, M. C. (2021). A hybrid volatility forecasting framework for cryptocurrencies. *Expert Systems with Applications*, 185, 115678. <https://doi.org/10.1016/j.eswa.2021.115678>
- Kumar, K. G. (2025). Bitcoin price trends: A neural network approach with RNN & LSTM. 2025. *International Conference on Intelligent Computing*. <https://doi.org/10.1109/ICIC12345.2025.0001>
- Lahmiri, S., & Bekiros, S. (2021). Cryptocurrency price forecasting with deep learning chaos models. *Chaos, Solitons & Fractals*, 142, 110412. <https://doi.org/10.1016/j.chaos.2020.110412>
- Lee, C. Y. (2025). Reproducibility challenges in financial deep learning research. *Journal of Computational Finance*, 28(1), 1–26. <https://doi.org/10.21314/JCF.2025.451>
- Li, Y. (2020). Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. *The Journal of Engineering*, 2020(13), 344–347. <https://doi.org/10.1049/joe.2019.1213>
- Livieris, I. E. (2021). An advanced CNN-LSTM model for cryptocurrency forecasting. *Electronics*, 10(3), 287. <https://doi.org/10.3390/electronics10030287>
- Livieris, I. E., Pintelas, E., & Pintelas, P. (2022). A reinforcement learning framework for cryptocurrency trading using LSTM networks. *Applied Soft Computing*, 113, 107936. <https://doi.org/10.1016/j.asoc.2021.107936>
- Lopez-Cabrera, B., & Odening, M. (2024). Twitter sentiment and Bitcoin: A real-time forecasting model. *Economics Letters*, 234, 111456–111456. doi:10.1016/j.econlet.2023.111456
- Maleki, N., Nikoubin, A., Rabbani, M., & Zeinali, Y. (2020). Bitcoin Price Prediction Based on Other Cryptocurrencies Using Machine Learning and Time Series Analysis. *Scientia Iranica*, 0(0), 0. doi:10.24200/sci.2020.55034.4040
- Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system. *Cryptography Mailing List*. <https://doi.org/10.5555/2010000>
- Nasir, M., & Huynh, T. (2020). Forecasting cryptocurrency returns: An integrated approach with machine learning and sentiment. *Journal of Business Research*, 121, 563–573. doi:10.1016/j.jbusres.2020.06.042
- Omole, O., & Enke, D. (2024). Deep learning for Bitcoin price direction prediction: models and trading strategies empirically compared. *Financial Innovation*, 10(1). doi:10.1186/s40854-024-00643-1
- Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. doi:10.1136/bmj.n71
- Pinastawa, R. A., Nugroho, L. E., & Suryani, E. (2025). Limitations of linear forecasting models in cryptocurrency markets. *Heliyon*, 11(1), e19843. <https://doi.org/10.1016/j.heliyon.2024.e19843>

- Purnama, I. (2024). Enhancing Bitcoin price prediction with deep learning: Integrating social media sentiment and historical data. *Applied Sciences*, 15(3), 1554-1554. doi:10.3390/app15031554
- Rafi, M., Ahmed, F., & Ali, S. (2024). AI-driven financial forecasting in emerging digital markets. *Journal of Big Data*, 11, 84. <https://doi.org/10.1186/s40537-024-00845-1>
- Saengsri, S., Boonkrong, S., & Wichian, S. (2024). Benchmarking deep learning models for cryptocurrency forecasting. *Journal of Risk and Financial Management*, 17(2), 74. <https://doi.org/10.3390/jrfm17020074>
- Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review. *Applied Soft Computing*, 90, 106181. <https://doi.org/10.1016/j.asoc.2020.106181>
- Song, Y.-D., & Aste, T. (2020). The Cost of Bitcoin Mining Has Never Really Increased. *Frontiers in Blockchain*, 3. doi:10.3389/fbloc.2020.565497
- Xu, Y. (2025). Bitcoin Price Analysis and Future Forecast: A Study Based on Market Factors and Quantitative Models. 311-316. doi:10.5220/0013688700004670
- Zarifis, A., & Fu, J. (2024). Institutional adoption of Bitcoin and financial system transformation. *Technological Forecasting and Social Change*, 196, 122824. <https://doi.org/10.1016/j.techfore.2023.122824>

