ISSN: 2723-9535

Available online at www.HighTechJournal.org

HighTech and Innovation Journal

Vol. 6, No. 1, March, 2025

Machine Learning Algorithms in Predicting Prices in Volatile Cryptocurrency Markets

Miguel Jiménez-Carrión [1*], Gustavo A. Flores-Fernandez [1]

[1] Faculty of Industrial Engineering, Universidad Nacional de Piura, Castilla-Piura, 20002, Peru.

Received 26 December 2024; Revised 21 February 2025; Accepted 26 February 2025; Published 01 March 2025

Abstract

This study aims to develop a predictive model for cryptocurrency prices in highly volatile markets. The methodology includes an exploratory data analysis, followed by designing and implementing machine learning (ML) algorithms, focusing on the Long Short-Term Memory (LSTM) neural network. The model's performance was optimized through hyperparameter tuning, and its stability was validated using an analysis of variance (ANOVA). We conducted a benchmark comparison with other ML approaches. Our LSTM model achieved an R^2 of 99.41% on the first day of prediction and maintained an accuracy above 97% up to the seventh day, demonstrating its robustness even for extended forecasts. During training, the LSTM model reached an RMSE of $1,187.14 and a MAPE of 2.20%, with the MAPE consistently remaining below 10% during the validation phase. For seven-day forecasts, the model recorded an RMSE of $5,038.46 and a MAPE of 6.83%. In comparison, alternative models such as Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), and Random Forests exhibited significantly higher error rates; for instance, XGBoost recorded an RMSE of $17,849.66 and a MAPE of 27.74%. Overall, these findings highlight the superior performance of the LSTM model in addressing the challenges of cryptocurrency price forecasting.

Keywords: Neural Networks; Criptocurrencies; Blockchain; Prediction.

1. Introduction

In recent years, the efficient use of financial resources has been significantly improved by technological advances. Among these, blockchain technology has emerged as a transformative financial alternative, offering secure and decentralized transactions and creating new opportunities for institutional and individual investors. The cryptocurrency market, driven by blockchain innovations, has grown exponentially, with Bitcoin as the most prominent digital asset, generating the largest monetary volume and influencing global financial dynamics [1, 2]. Despite the rapid expansion of the market, accurately predicting cryptocurrency prices remains a complex and dynamic challenge. Traditional statistical and econometric models, such as autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH), have been widely used for financial time series prediction [3, 4]. However, these models often struggle to handle cryptocurrency markets' high volatility and non-linear patterns [5]. As a result, machine learning (ML) approaches, particularly artificial neural networks (ANNs), have gained significant attention for their ability to model complex, nonlinear relationships, as well as pick up subtle market signals [6, 7]. Recent studies have demonstrated the effectiveness of Long Short-Term Memory (LSTM) networks in financial forecasting, especially for multi-step time series forecasting [8, 9]. However, there is still a gap in the literature regarding the optimization of

* Corresponding author: mjimenezc@unp.edu.pe

http://dx.doi.org/10.28991/HIJ-2025-06-01-017

LSTM architectures specifically tailored to the unique characteristics of the cryptocurrency market, including its extreme volatility, external market influences, and investor sentiment.

This study aims to address this gap by designing an LSTM neural network to predict cryptocurrency prices more accurately and reliably. By evaluating multiple configurations and architectures that allow optimizing the hyperparameters used, the study intents to identify the model that offers the best computational accuracy and efficiency, thus providing a valuable tool for investors looking to maximize their returns within a defined risk tolerance. This approach contributes to the advancement of knowledge regarding financial time series prediction and provides useful information for market participants dealing with the volatile and speculative nature of cryptocurrencies. The findings of this study could improve investment strategies, inform risk management practices, and encourage greater adoption of machine learning techniques in financial decision-making.

2. Literature Review

Cryptocurrencies such as Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Solana (SOL), Monero (XMR), and Oryen (ORY) are becoming increasingly relevant in the financial world and are considered an emerging market. Easy access and abundant data in the cryptocurrency market make it an ideal study subject. By applying machine learning (ML) and sentiment analysis techniques, researchers can gain insights into market behavior and address the complex task of predicting cryptocurrency values. Some studies have focused exclusively on Bitcoin's behavior. Some studies suggest using machine learning and social media data to predict the price movements of BTC, ETH, XRP, and LTC in cryptocurrency markets. This study compares the application of machine learning algorithms such as neural networks (NN), support vector machines (SVMs), and random forests (RF) by utilizing Twitter (rebrand as X) elements and market data as input features to develop a predictive cryptocurrency price model. The results showed that machine learning algorithms and sentiment analysis applied only to Twitter data can be used to develop a predictive model for specific cryptocurrency markets. Results also showed that NN outperforms the other models previously applied [10].

Time series modeling and prediction is an arduous and essential task for financial optimization procedures. Numerous studies have been carried out to reduce investor uncertainty by forecasting the price of currency and share prices. However, the emergence of a new type of currency with its own characteristics, known as cryptocurrencies, poses additional challenges. A past study suggested analyzing how social media posts reflect investor expectations and influence the coin's future value. The study objective was to forecast the daily market performances based on two components: those that define the behavior of the cryptocurrency (volume, opening value, closing value, maximum value, and minimum value) and those that affect its behavior, such as the expectations and interactions of the environment, obtained from the tweets collected. To achieve their goal, the researchers proposed the use of a type of recurrent neural network, known as "Long Short-Term Memory" (LSTM). Their method, which involved data preprocessing and time series forecasting, achieved a MAPE value of 34.92%. These results indicate that the representation of the perception variable in social networks was not relevant and, therefore, motivates additional work to model this variable using other natural language processing (NLP) techniques [11].

Recently, cryptocurrencies have become an essential and well-known component with both economic and financial potential. Unfortunately, acquiring Bitcoin is not straightforward due to uneven business and significant rate fluctuations. Traditional price forecasting methods have been less effective, as real-time predictions are now possible. Specifically, research recommends a machine learning-based alternative for a mortgage lender based on the issues highlighted in Bitcoin price forecasting. The proposed strategy includes a reinforcement learning algorithm for price estimation and forecasting and a blockchain framework for an efficient and secure environment. As a result, predictions achieved better performance compared to other systems, with respect to XMR, LTC, ORY, and BTC [12].

Cryptocurrency, a product of advancing financial technology, offers significant research opportunities with hundreds of cryptocurrencies used worldwide. Cryptocurrency price forecasting is difficult due to price volatility and dynamism. In the study by Hamayel & Owda (2021) [13], three types of recurrent neural network (RNN) algorithms were used to predict the prices of BTC, LTC, and ETH cryptocurrencies. The models show excellent predictions based on MAPE, with the neural network gated recurrent unit (GRU), a type of RNN, outperforming the LSTM and bi-LSTM models, making it the best algorithm. GRU presents the most accurate prediction for LTC with MAPE values of 0.2454%, 0.8267%, and 0.2116% for BTC, ETH, and LTC, respectively. The bi-LSTM algorithm features the lowest prediction result compared to the other two algorithms, as the MAPE values are 5.990%, 6.85%, and 2.332% for BTC, ETH, and LTC, respectively. The authors also argue that the prediction models of their research represent accurate results close to the actual prices of cryptocurrencies. The importance of having these models is that they can have significant economic ramifications by helping investors and traders identify cryptocurrency sales and purchases.

On the other hand, Yang et al. (2023) [14] manifest that cryptocurrency prices have the characteristic of high volatility that leads to resistance in predicting cryptocurrency prices. These limitations expose the need for accurate cryptocurrency methods for price prediction that can reduce investors' investment risk. To address these issues, the authors proposed using the fractional gray model (FGM (1,1)), a novel approach to predict the price of blockchain cryptocurrency. Specifically, the study established the FGM (1,1) through the closing price of three representative blockchain cryptocurrencies: BTC, ETH, and LTC. The authors adopted the Particle Swarm Optimization (PSO) algorithm to obtain the optimal order of the model. Based on these findings, we evaluated the predictive accuracy of FGM (1,1) using MAPE, the mean absolute value (MAE), and RMSE and compared it through experiments. Our results indicate that within the range of data studied, the predictive accuracy of the FGM (1.1) in the closing price of BTC, ETH, and LTC has reached a highly accurate level. Compared to FGM's previous results (1.1), our FGM (1.1) exceeds the predictive ability in experiments. The author's study provides a feasible new method for blockchain cryptocurrency price prediction. It has specific references and information for government departments, investors, and researchers in theory and practice.

Acknowledging that cryptocurrencies are highly volatile and complex to predict as investment assets, Sung et al. (2022) [15] study uses various cryptocurrency data as features to predict the logarithmic return price of major cryptocurrencies. The study's contribution is the selection of the most influential main characteristics for each cryptocurrency using the volatility characteristics of the cryptocurrency, derived from the models of autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH), along with the closing price of the cryptocurrency. In addition, the authors sought to predict the logarithmic price return of cryptocurrencies by implementing various types of time series models. Based on the selected main features, the cryptocurrency's logarithmic return price was predicted using the ARIMA time series prediction model and the artificial neural network-based time series prediction model. As a result of logarithmic yield price prediction, neural network-based time series prediction models showed superior predictive power compared to the traditional time series prediction model.

The high volatility of cryptocurrencies has attracted significant attention, with Bitcoin being the most notable. These observations sparked Maleki et al.'s (2023) [16] interest in developing methods to predict these fluctuations, even though they are challenging. Although several investigations used traditional statistical and economic methods to uncover the determinants of cryptocurrency prices, progress in developing prediction models for decision-making tools in investment techniques is still in its early stages. Many methods of cryptocurrency price prediction, such as forecasting a one-step approach, can be performed using time series analysis, neural networks, and machine learning algorithms. However, it is necessary to realize the long-term trend of a currency. The study aimed to investigate and forecast Bitcoin prices using machine learning algorithms based on analyzing three well-known cryptocurrencies: Ethereum, Zcash, and Litecoin while assuming minimal information about Bitcoin prices. In addition, they proposed a new method to predict the price of Bitcoin by considering the prices of different cryptocurrencies. The results showed that Zcash performed best in predicting the price of Bitcoin without information on the price fluctuations of Bitcoin, among other cryptocurrencies.

Referring to the substantial volatility and non-stationarity of cryptocurrency prices, Jin & Li (2023) [17] stated that forecasting them has become a complex task within the financial time series analysis field. The authors presented the innovative hybrid prediction model, VMD-AGRU-RESVMD-LSTM, which combines the disintegration-integration framework with deep learning techniques to predict the price of cryptocurrencies accurately. The process begins by decomposing the cryptocurrency price series into a finite number of subseries, each characterized by relatively simple volatility patterns, using the variational mode decomposition (VMD) method. Next, the GRU neural network, combined with an attention mechanism, predicts the sequence of each modal component separately. In addition, the residual sequence, obtained after decomposition, undergoes further decomposition. The resulting residual sequence components serve as input to a network GRU (AGRU), which predicts the future values of the residual sequence. Ultimately, the neural network LSTM integrates modal and residual component predictions to generate the final predicted price. The empirical results obtained for the daily Bitcoin and Ethereum data show promising performance. The metrics results report the following values: RMSE of 50,651 and 2,873, MAE of 42,298 and 2,410, and MAPE of 0.394% and 0.757%, respectively. Notably, the predictive results of the VMD-AGRU-RESVMD-LSTM model outperform the LSTM and GRU models, as well as other hybrid models, confirming its superior performance in crypto price forecasting.

Virtual currencies, widely recognized as currencies of exchange, have been declared financial assets and are catching the attention of investors as they can lead to very profitable investments. However, having access to accurate price prediction is essential to optimize profits from cryptocurrency investments. Since price prediction is a time-series task, this study proposes a hybrid deep-learning model to provide price cryptocurrency predictions. The hybrid model integrates a one-dimensional convolutional neural network and a stacked gated recurrent unit (1D-CNN-GRU). Given the price data of cryptocurrencies over time, the one-dimensional convolutional neural network encodes the data into a high-level discriminative representation. Subsequently, the stacked closed recurring unit captures the long-range dependencies of the rendering. The hybrid model was evaluated on three cryptocurrency datasets: Bitcoin, Ethereum, and Ripple. The experimental results demonstrated that the proposed 1D-CNN-GRU model outperformed existing methods with the lowest RMSE values of 43.933 in the Bitcoin dataset, 3.511 in the Ethereum dataset, and 0.00128 in the Bitcoin dataset Ripple [18].

According to Aljadani (2022) [19], cryptocurrencies are digital currencies that have emerged with financial technology advancements. In 2017, cryptocurrencies showed a massive increase in market capitalization and popularity. They are employed in today's financial systems, as individual investors, corporate companies, and large institutions are investing heavily in them. However, this industry is less stable than traditional forex markets. A digital currency market can fluctuate due to legal, sentimental, and technical factors. Therefore, it is crucial to make accurate cryptocurrency price forecasts. Recently, cryptocurrency price prediction has become a trending research topic globally. The study presented machine and deep learning algorithms, including NN, GRU, LSTM, and two-way LSTM (BiLSTM) methodologies to analyze the factors influencing cryptocurrency prices and predict them accordingly. The author proposed a five-phase framework for predicting cryptocurrency prices using BiLSTM and GRU deep learning models. The author used three real-time public cryptocurrency datasets from "Yahoo Finance," Long-term bidirectional memory and closed recurring unit-based deep learning-based algorithms to forecast the prices of three popular cryptocurrencies (i.e., Bitcoin, Ethereum, and Cardano, and the Grid Search approach for the hyperparameter optimization processes. The results indicate that GRU outperformed the BiLSTM algorithm for Bitcoin, Ethereum, and Cardano. The lowest RMSE for the GRU model was found to be 0.01711, 0.02662, and 0.00852 for Bitcoin, Ethereum, and Cardano, respectively. The experimental results demonstrated the significant performance of the proposed framework that achieves the minimum values of MSE and RMSE.

Since the arrival of Bitcoin, the cryptocurrency landscape has seen the emergence of several virtual currencies that have quickly established their presence in the global market. The dynamics of this market, influenced by a multitude of factors that are difficult to predict, pose a challenge to fully understand its underlying ideas. A study suggests a methodology for determining the best times to buy or sell cryptocurrencies to maximize profits. The study indicates that based on large market and social media datasets, they used a methodology that combines different statistical, text analysis, and deep learning techniques to support a recommendation trading algorithm. In particular, the study examines the correlation between social media posts and price changes and the impact of social media sentiment on cryptocurrency prices. Several experiments were conducted with historical data to evaluate the effectiveness of the trading algorithm, achieving an overall average profit of 194% without transaction fees and 117% deducting transaction fees. Cryptocurrencies considered included high-capitalization coins, solid projects, and meme coins. Meme coins are based on memes and serve as an alternative for easy investments. Therefore, a meme coin has no intrinsic value and is rarely useful.), the trading algorithm proved to be very effective in predicting the price trends of influential meme coins, generating considerably higher profits compared to other types of cryptocurrencies [20].

Quiroga Juárez & Villalobos Escobedo (2023) [21] propose a descriptive and inferential statistical study using one hundred cryptocurrencies. Their hypothesis states that by analyzing historical data, it would be possible to generate scenarios that favor the understanding of the cryptocurrency phenomenon; in addition, it could be supportive of portfolio management. The analysis period covered April 28, 2013, to August 4, 2022. The data was obtained from the CoinGecko platform. The theoretical contribution spans studying an emerging phenomenon with social implications that has gained global momentum, influenced by technological dynamism and governmental and private agents. The analysis results provide the historical behavior of one hundred cryptocurrencies in the market, prospect scenarios, and identify correlations between cryptocurrencies, which is important for the creation of investment portfolios from a risk diversification approach. In conclusion, the study generates a framework for understanding the evolution of the cryptocurrency market from a selected sample of one hundred assets. Likewise, with cluster analysis, a classification of these was made according to correlation; this, from a portfolio theory approach, would allow risk diversification.

Currently, it has been determined that highly accurate cryptocurrency price predictions are of utmost importance to investors and researchers. However, due to the non-linearity of the crypto market, it is difficult to assess the distinctive nature of time-series data, leading to challenges in generating accurate price predictions. These scenarios motivated numerous studies on the prediction of the price of cryptocurrencies using different algorithms based on DL. Among these studies is Seabe et al. (2023) [22], who proposed using three types of networks, LSTM, GRU, and Bi-LSTM, for exchange rate predictions, applied to the top three cryptocurrencies by market capitalization: BTC, ETC, and LTC. The metric results RMSE and MAPE indicated that Bi-LSTM provided higher prediction accuracy compared to LSTM and GRU with MAPE values of 0.036, 0.041, and 0.124 for BTC, LTC, and ETH. Therefore, Bi-LSTM can be considered the best algorithm. The study suggests that its models for predicting cryptocurrency prices are accurate and can prove beneficial for investors and traders.

With the purpose of providing a framework that overcomes the limitations of uncertainty, volatility, and dynamism and that is capable of reproducing the predictions not only in the most common cryptocurrencies but at the same time is consistent and has the capacity for generalization. Murray et al. (2023) [23] proposed to create a comparison framework that overcomes these limitations and to use that framework to conduct extensive experiments in which the performance of statistical, ML, and DL approaches widely used in the literature to predict the price of five popular cryptocurrencies, XRP, BTC, LTC, ETH, and XMR. The researchers argue that they are the first to propose the use of the temporary fusion transformer (TFT) in their study. In addition, they expanded their research to hybrid models and sets to assess whether combining individual models increases prediction accuracy. The assessment demonstrates that deep learning approaches, particularly LSTM, serve as effective predictors for all cryptocurrencies studied, with LSTM achieving an average RMSE of 0.0222 and an MAE of 0.0173.

Samson (2024) [24] presents a relevant study in this area, evaluating the effectiveness of three machine learning (ML) algorithms—Gradient Boosting (GB), Random Forest (RF), and Bagging—in predicting the daily closing prices of six major cryptocurrencies: Binance, BTC, ETH, SOL, USD, and XRP. Unlike traditional approaches that use open, high, and low prices as predictive characteristics, the study adopted an innovative methodology by employing lagged prices as entry characteristics. The approach assumed that lagging prices better represent the temporal dynamics of cryptocurrency prices than conventional prices.

The analysis used a historical dataset that spanned from 2015 to 2024, depending on the cryptocurrency, and divided the data into a training set (80%) and a test set (20%) to evaluate the performance of the algorithms. The results showed that the GB algorithm performed the best at predicting the prices of BTC and SOL, while RF was more effective at predicting the prices of ETH, USD, and XRP. This finding highlights differences in the effectiveness of algorithms depending on cryptocurrency and market characteristics, suggesting that the use of RF may be more appropriate in certain contexts, while GB could deliver better results in others.

Fang et al. (2024) [25], inspired by the recent success of the application of ML in stock market prediction, analyzed and presented the specific characteristics of the cryptocurrency market in a high-frequency trading context. Specifically, the study showed the application of an ML approach to predict the direction of changes in the average price in the next tick. Their results indicate that there are universal features across cryptocurrencies that allow models to outperform asset-specific ones. In addition, they demonstrated that using long sequences of data points does not improve predictions, highlighting the inefficiency of feeding models with extensive datasets. They also addressed the technical challenge of designing a lightweight predictor capable of working effectively with live data from cryptocurrency exchanges. To improve the performance of the model, they presented a new method of retraining. Finally, they examined the trade-off between model accuracy and retraining frequency in the context of multi-label prediction. Overall, their findings show that promising results can be achieved with live data, as evidenced by a consistent 78% accuracy in predicting Bitcoin's average price movements against the U.S. dollar.

Kiranmai Balijepalli & Thangaraj (2025) [26] presents a major study that addresses the growing popularity of cryptocurrency markets, which, as of 2023, included more than 23,000 cryptocurrencies and a total market valuation of $870.81 billion. Although cryptocurrencies are becoming increasingly significant, they are still prone to volatility, making predicting their prices challenging for investors seeking to make informed decisions. The study aimed to develop a dynamic forecasting model using an assembly approach and assess the accuracy of predictions for the top 15 cryptocurrencies. The accuracy of statistical and econometric models is evaluated after the adjustment of hyperparameters, extracting information from these models to build an assembly model using ML algorithms. Specifically, the study employs Gradient-Boosted Regressor (GBR), Random Forest Regressor (RFR), Support Vector Regression (SVR), and Multi-Layer Perceptron (MLP), using validation curves to optimize model parameters and improve prediction accuracy. The study's findings reveal that when price movements exhibit autocorrelation, the ARIMA and the assembly model outperform other methods. Models such as ARIMA, Simple Linear Regression (SLR), Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), and Multi-Model Regression (MLR) demonstrated good performance with cryptocurrencies, suggesting that trends, seasonality, and historical price patterns play a significant role in price prediction. Notably, the MLR approach provided more accurate forecasts for cryptocurrencies with higher volatility and irregular price patterns, highlighting their potential for prediction in the unpredictable digital market.

Hossain et al. (2024) [27] explored the critical role of time series prediction in financial markets, particularly in predicting asset prices and guiding investment decisions. The volatility inherent in cryptocurrency markets, such as BTC and ETH, complicates prediction due to extreme price fluctuations driven by market sentiment, technological changes, and government regulations. Traditionally, prediction in financial markets was based on statistical methods, but as markets became more complex, the emergence of deep learning models such as LSTM, Bi-LSTM, and, more recently, FinBERT-LSTM, offered a new approach to capturing intricate patterns and dynamics within data. In response to the challenges posed by the high volatility of cryptocurrencies, Mabsur proposes a hybrid model that integrates Bi-LSTM networks with FinBERT, a model known for its sentiment analysis capabilities. This hybrid approach aims to enhance prediction accuracy by integrating advanced time series prediction models with sentiment analysis, a method that incorporates both historical price data and the emotional and psychological factors affecting market behavior. The study fills a significant gap in financial forecasting by offering a model capable of navigating the complexities and unpredictability of volatile cryptocurrency markets. The hybrid model provides valuable insights for investors and analysts, enabling them to make more informed decisions in the face of unpredictable market conditions.

Islam et al. (2024) [28] explored the dynamic and volatile nature of the cryptocurrency market, which has significantly influenced financial ecosystems globally. In their study, the authors focus on the growing importance of cryptocurrencies, which have evolved from niche digital assets to mainstream investment opportunities, such as BTC and ETH. The study aimed to investigate the effectiveness of various ML algorithms in predicting cryptocurrency prices within the volatile U.S. financial market. By identifying which ML techniques provide the most accurate and reliable predictions under different market conditions, the research contributes to understanding the strengths and limitations of these approaches. The dataset used for cryptocurrency price prediction analysis includes a wide range of data sets sourced from major cryptocurrency exchanges such as Binance, Coinbase, and Kraken, in addition to essential trading metrics to define market dynamics. The authors use aggregated data from renowned financial databases, such as CoinMarketCap, CryptoCompare, and Yahoo Finance, ensuring a solid foundation for ML models. The models considered in the study range from simpler linear methods to more complex assembly and gradient optimization algorithms. The authors evaluate the predictive performance of these models using several metrics, including Accuracy, Recall, F1-Score, MAE, RMSE, and R-squared. Among the algorithms tested, the Gradient Boosting model showed superior performance in terms of accuracy, recall, and F1-score. In addition, the three models evaluated exhibited relatively low values of MAE and RMSE, indicating their effectiveness in predicting cryptocurrency price movements.

The findings underscore the importance of ML models for cryptocurrency price prediction, particularly for investors and financial market players. The study highlights that cryptocurrencies have become key components of individual and institutional investment portfolios, as well as trading strategies. By integrating ML models into investment management, they can provide valuable insights into entry and exit points, portfolio diversification, and risk management. Lee et al. (2024) [29] suggest that the consolidation of machine learning techniques within the financial system marks a significant shift towards data-driven decision-making in cryptocurrency trading.

3. Research Methodology

The methodology followed in this study is presented in Figure 1 and is described in the following steps: 1) Data Collection: Gather data from the top cryptocurrencies based on each daily transaction volume. This approach ensures that the most active and relevant cryptocurrencies are included in the analysis. The data includes historical price information, trading volumes, and other pertinent market indicators. 2) Data Exploration and Preprocessing: Examine data to identify missing or incomplete values. Properly handling missing values is essential, as they could affect the accuracy and quality of predictive models. Select imputation techniques, such as average imputation or time-based interpolation, so the dataset is complete and ready for modeling. 3) Implementation of Predictive Models: Several ML algorithms are implemented to build predictive time-series models. These algorithms include a) Random Forest: An ensemble learning method that combines multiple decision trees to increase the accuracy of predictions. b) XGBoost: A gradient optimization algorithm known for its high performance and efficiency, especially when working with large volumes of data and complex patterns. c) Support Vector Machines (SVMs): A robust classifier that works well in high-dimensional spaces and is suitable for predicting movements in cryptocurrency prices. d) LSTM (Long Short-Term Memory) networks: A recurrent neural network that captures long-term dependencies in time series data, making it ideal for predicting volatile markets such as cryptocurrencies. 4) Hyperparameter Optimization: To determine the hyperparameters that best fit the cryptocurrency time series, the performance of each model is evaluated using metrics such as mean square error (MSE) or root mean square error (RMSE). Hyperparameter optimization uses techniques like grid search to find the best model configuration for accurate predictions. 5) Analysis and Discussion of Results: Once the models have been trained and their hyperparameters optimized, the results are analyzed by comparing the performance of different models, discussing the accuracy and reliability of predictions, and identifying patterns or trends in the cryptocurrency market that can help explain the results obtained. 6) Conclusions: This last step involves summarizing the main insights gained, discussing the implications of the results for cryptocurrency price prediction, and suggesting possible improvements or directions for future research.

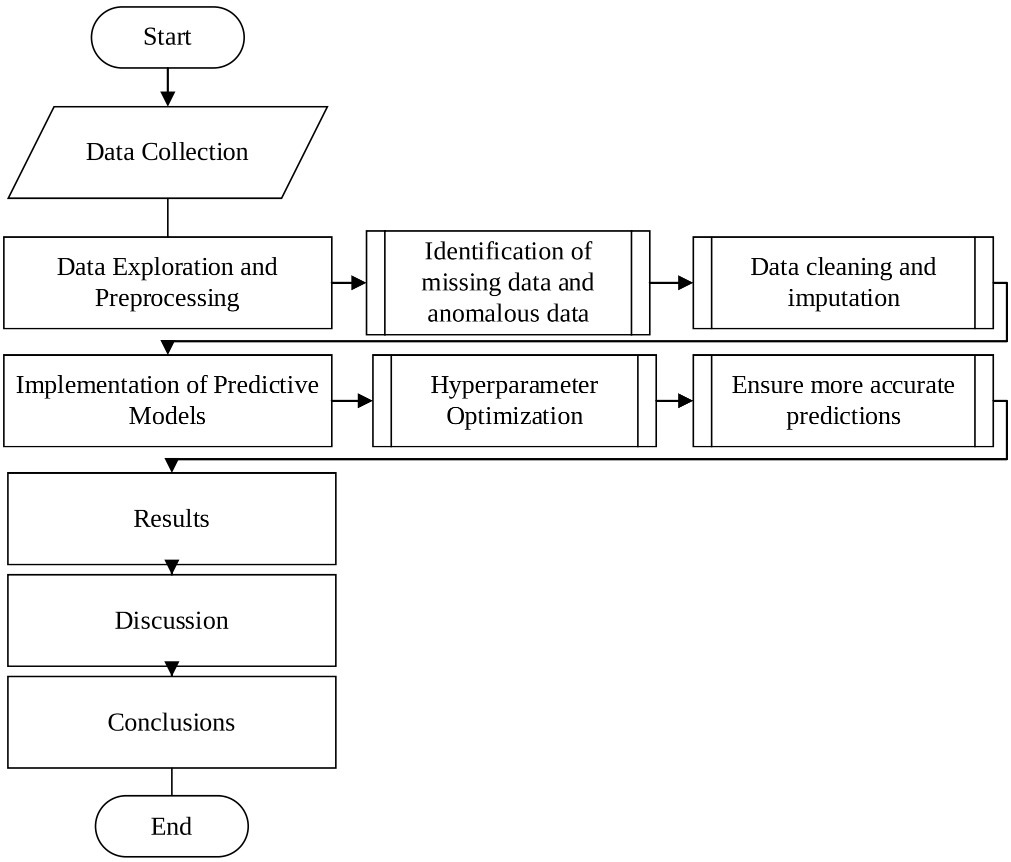


Figure 1. Flow diagram of the methodological process of the research

4. Results

4.1. Exploratory Data Analysis

In this study, 12 cryptocurrencies with the highest cryptocurrency market capitalization were considered, whose daily data was downloaded from the es.investing.com platform; this data is shown in Table 1.

Table 1. Historical data of cryptocurrencies

Cryptocurrency	Start Date	End Date	Number of Records
Bitcoin	18/07/2010	11/07/2024	5108
Ethereum	10/03/2016	11/07/2024	3046
Polkadot	08/02/2021	11/07/2024	1250
Shiba Inu	12/05/2021	11/07/2024	1157
BNB	09/11/2017	11/07/2024	2437
Avalanche	03/01/2021	11/07/2024	1286
TRX	13/06/2018	11/07/2024	2221
Dogecoin	03/06/2017	11/07/2024	2596
Cardano	31/12/2017	11/07/2024	2385
XRP	22/01/2015	11/07/2024	3458
Solana	13/07/2020	11/07/2024	1453
Litecoin	24/08/2016	11/07/2024	2879

The platform provided data on 6 variables related to the price of cryptocurrencies, as shown in Table 2.

Table 2. Variables Related to Cryptocurrencies

Last	Last cryptocurrency price value
Initial	Initial cryptocurrency price value
Maximum	Maximum cryptocurrency price value
Minimum	Minimum cryptocurrency price value
Volume	Daily volume in monetary value of cryptocurrency transactions
% of variation	Variability of cryptocurrency

We proceeded to identify the missing data for each of the cryptocurrencies in relation to their 6 characteristics. The summary is shown in Table 3.

Table 3. Percentage of Missing Data

Cryptocurrency	Percentage of Missing Data					
	Last	Initial	Maximum	Minimum	Volume	% Var.
Bitcoin	0.0%	0.0%	0.0%	0.0%	0.12002%	0.0%
Ethereum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Polkadot	0.0%	0.0%	0.0%	0.0%	17.36000%	0.0%
Shiba Inu	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
BNB	0.0%	0.0%	0.0%	0.0%	0.4103%	0.0%
Avalanche	0.0%	0.0%	0.0%	0.0%	19.67341%	0.0%
TRX	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Dogecoin	0.0%	0.0%	0.0%	0.0%	0.03852%	0.0%
Cardano	0.0%	0.0%	0.0%	0.0%	0.20964%	0.0%
XRP	0.0%	0.0%	0.0%	0.0%	5.75477%	0.0%
Solana	0.0%	0.0%	0.0%	0.0%	27.39160%	0.0%
Litecoin	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

As can be observed, only three cryptocurrencies presented more than 15% of missing data in their characteristic capitalization volume, which is above the permissible limit [30]. This limitation was considered while training the LSTM network's multivariate model, including or excluding this parameter to assess performance metrics' behavior. To address missing or incomplete data, we took into account that conventional methods such as imputation by mean or median can bias the behavior of the price of the cryptocurrency to use data that may be far from the maximum or minimum value considering the range of the price of cryptocurrencies, as well as the interpolation method, whether linear, quadratic, or cubic, are not suitable to have large ranges of empty values as could be observed in the data set, the K-Neighbors method (k=3) was considered [31] for data imputation. Concerning outliers, cryptocurrencies are highly volatile assets due to speculation; therefore, updating data using anomalous data processing methods is not considered since their nature includes many anomalous data typical of this type of asset. Boxplots showing Bitcoin data distribution are shown in Figure 2.

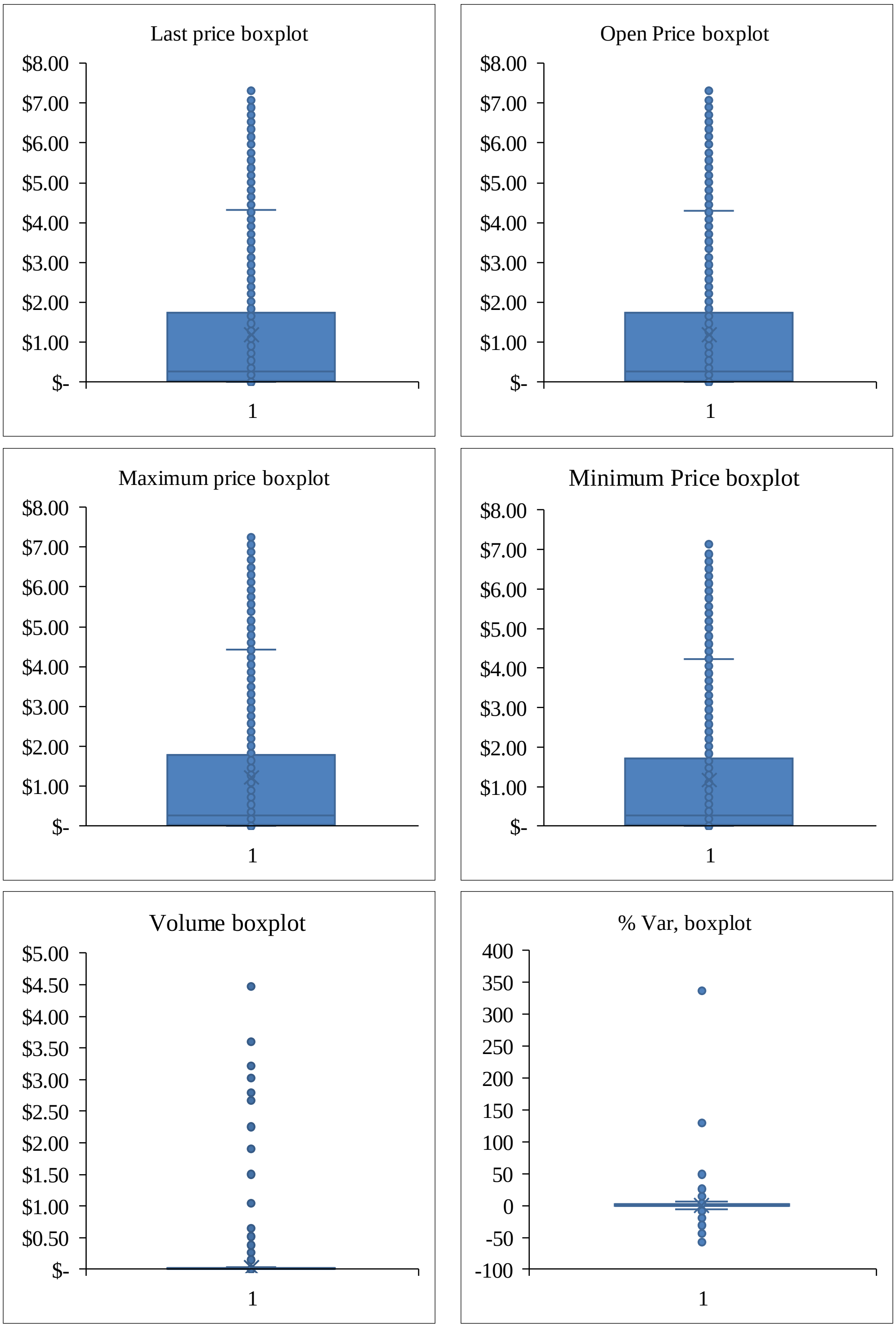


Figure 2. Bitcoin Features Distribution

4.2. LSTM Network Implementation

We considered two implementation approaches; the first uses the final value of the cryptocurrency as a prediction value; that is, it is a multistep univariate model, and the second uses the 6 characteristics provided by the investing platform, so it is categorized as a multivariate multistep model. Pseudocode 1 represents the LSTM network model, which is valid for both approaches.

Pseudocode 1. LSTM Network

```
Input: Libraries
Input: Input variables
Input: Hyperparameters
Output: Algorithm Performance Metrics and Cryptocurrency Price
Data frame ← Cryptocurrency values
tr, vl ← Training and validation sets
function DATA SCALING
    Scaled training set ← MinMax Scaler(Training set)
    Scaled validation set ← MinMax Scaler(Validation set)
end function
function MODEL TRAINING
    Model ← Sequential
 Model ← Add LSTM layer
 Model ← Add Dense layer
end function
function PERFORMANCE METRICS
    RMSE, MAE, MAPE, R² ← Model performance
end function
Output: Model performance Metrics and Cryptocurrency Price
```

As a first step, we determined the size of the training and validation sets as 80% and 20%, respectively. For example, the Bitcoin data is split into 4086 values in the training set and 1022 in the validation set. This segmentation is shown in Figure 3.

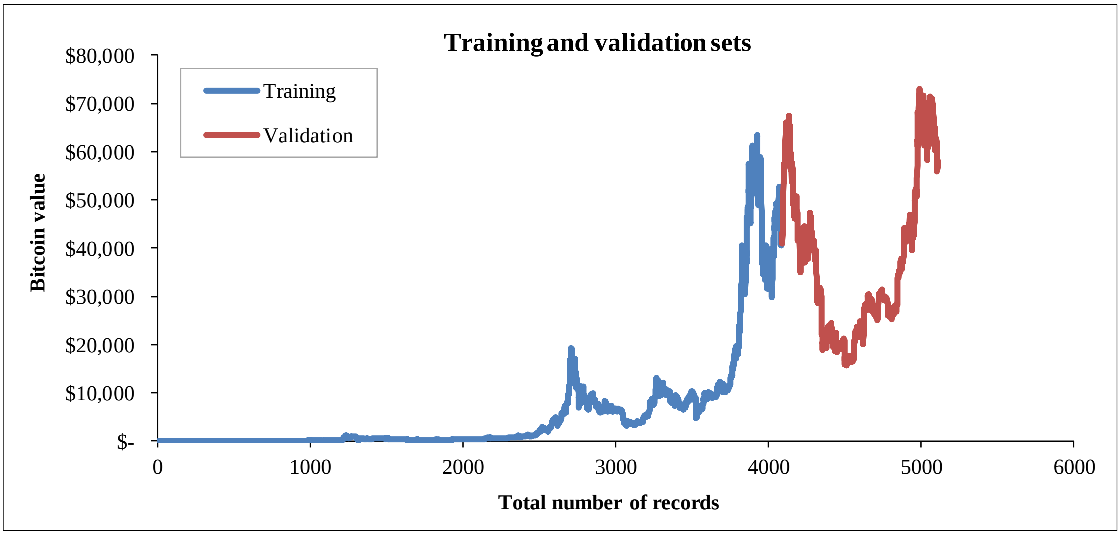


Figure 3. Bitcoin data representation

The data was normalized using the min-max scaler, which ensured that the input features were unbiased and maintained the stability of the synaptic weights throughout the training process. To determine the best-performing hyperparameters, we applied the calibration process, managing to specify some parameters constantly, such as the number of layers of the LSTM network, which were 3, and for the factors that demonstrated influence on the predictions, a factorial experimental design was made with the factors and levels, which are shown in Table 4.

Table 4. Factors and Levels of the Factorial Design

LSTM Model			
Factors	**Levels**		
Learning Rate (A)	0.0009	0.00009	0.00171
Number of Epochs (B)	50	100	150
Batch Size (C)	40	70	150

The system's stability is determined with an analysis of variance using Equation 1 to establish whether there are significant differences in the results.

$$y_{ijkn} = \mu + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + e_{ijkn}$$
$$i = 1,2,3; \; j = 1,2,3; \; k = 1,2,3; \; n = 1,2,3,4 \tag{1}$$

In Equation 1, the parameter y_{iikn} represents the predictive model response expressed in the metrics RMSE, MAE, MAPE, and R^2, referred to the validation and then training data. The ANOVA analysis is shown in Table 5, and the results show high accuracy concerning the mean. The stability of the network was evidenced by a coefficient of variability of 20.44%, which indicates a minimal dispersion from to the mean. The statistical significance of the factors and their interactions are presented in Table 5.

Table 5. RMSE Analysis of Variance

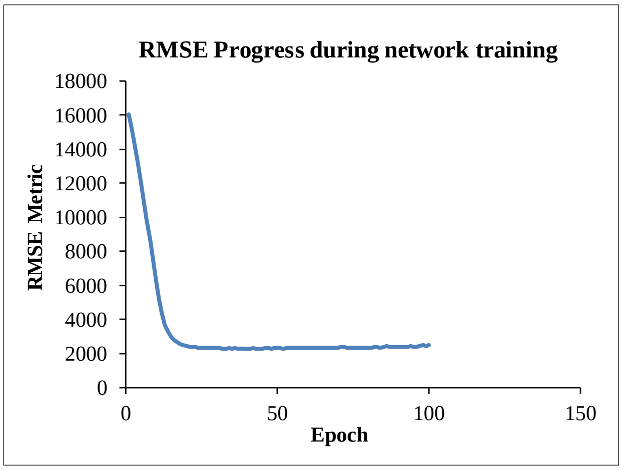
F.V.	SC	gl	CM	F	p-valor	Sig
A	9282870.02	2	4641435.01	40.24	0.00000	**
B	1501483.75	2	750741.87	6.51	0.00239	*
C	4297553.95	2	2148776.97	18.63	0.00000	**
A*B	597770.01	4	149442.50	1.30	0.27866	
A*C	1725977.17	4	431494.29	3.74	0.00763	*
B*C	645143.34	4	161285.83	1.40	0.24193	
A*B*C	1107351.13	8	138418.89	1.20	0.30943	
Error	9341877.10	81	115331.82			
Total	3284502438	168				
CV = 20.44%						

At 99.99% confidence, the results showed the statistical significance of the learning ratio factors and batch size. In comparison, at 95% confidence, there is a statistical significance of the factor number of epochs and interaction of batch size and learning ratio. These results validated the factors' relevance in the model. Next, we performed the Duncan test with an alpha of 0.05 as a parameter. These results are shown in Table 6.

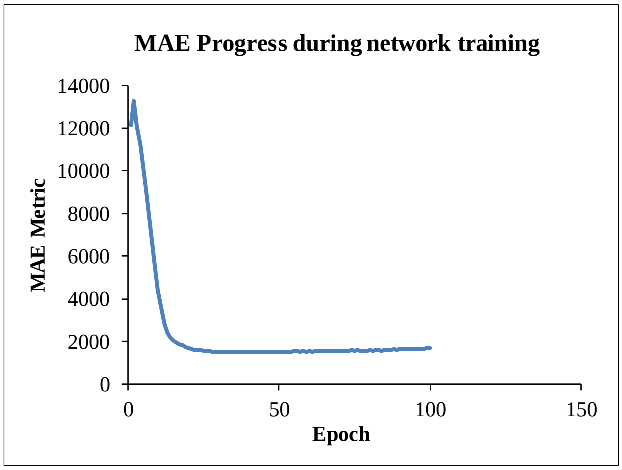
Table 6. Duncan RMSE Test

A	B	C	Means	n	E.E.						
0.00009	50	150	1187.14	4	169.80	a					
0.00009	150	150	1195.21	4	169.80	a					
0.00009	100	150	1197.99	4	169.80	a					
0.00009	50	40	1198.72	4	169.80	a					
0.00009	50	70	1206.23	4	169.80	a					
0.00090	50	150	1222.84	4	169.80	a					
0.00009	100	70	1232.89	4	169.80	a					
0.00009	150	70	1342.43	4	169.80	a	b				
0.00009	150	40	1351.08	4	169.80	a	b				
0.00009	100	40	1357.54	4	169.80	a	b				
0.00171	50	70	1413.84	4	169.80	a	b	c			
0.00171	50	150	1450.83	4	169.80	a	b	c			
0.00090	150	150	1453.41	4	169.80	a	b	c			
0.00090	50	70	1461.91	4	169.80	a	b	c			
0.00171	100	150	1486.48	4	169.80	a	b	c			
0.00090	100	150	1645.56	4	169.80	a	b	c	d		
0.00171	150	150	1864.14	4	169.80		b	c	d	e	
0.00090	150	40	1864.94	4	169.80		b	c	d	e	
0.00171	150	70	1931.41	4	169.80			c	d	e	
0.00090	50	40	2124.62	4	169.80				d	e	f
0.00090	100	70	2126.24	4	169.80				d	e	f
0.00171	100	70	2149.84	4	169.80				d	e	f
0.00171	50	40	2189.71	4	169.80				d	e	f
0.00090	100	40	2198.84	4	169.80				d	e	f
0.00090	150	70	2203.16	4	169.80				d	e	f
0.00171	100	40	2280.85	4	169.80					e	f
0.00171	150	40	2530.65	4	169.80						f

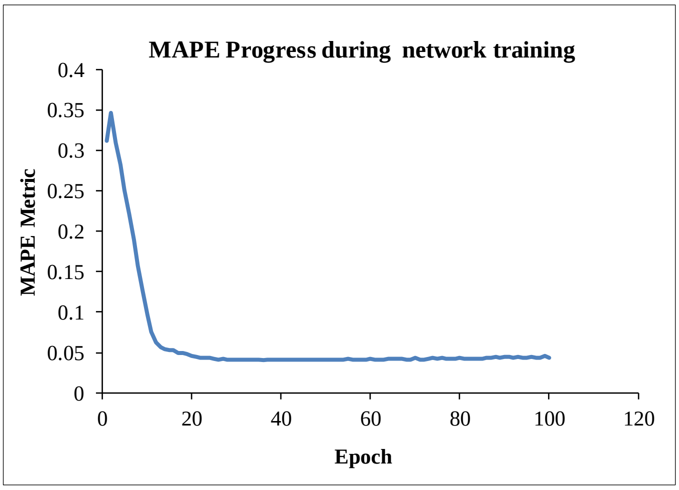
Based on Duncan's test, the learning rate of 0.00009, 50 iterations, and a batch size of 150 provided the better average RMSE of $1187.14. The stability of the error metrics is shown in Figure 4.



(a) RMSE



(b) MAE



(c) MAPE

Figure 4. Error metrics during network training

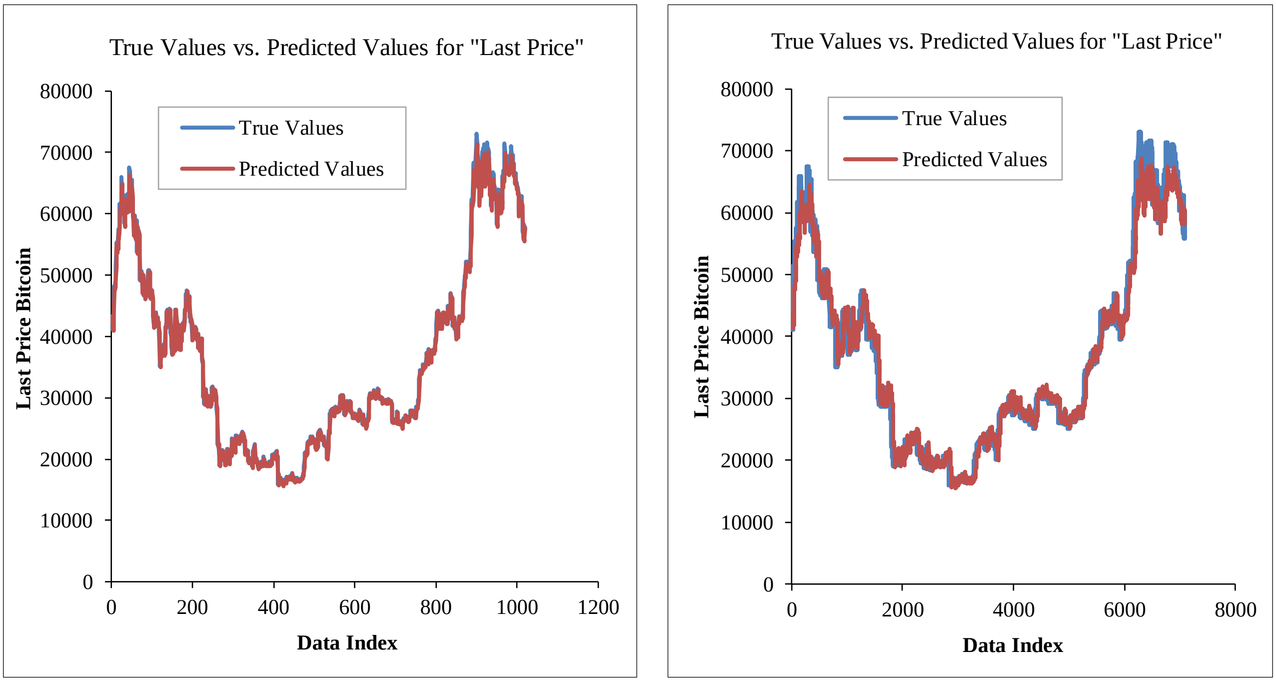
After determining the optimal hyperparameters, we found the value of the validation set metrics for each of the seven steps, which represent each predicted da, see error values in Table 7.

Table 7. LSTM Network Multi-Step Performance

Step	RMSE	MAE	MAPE	R^2
1	$ 1,187.14	$ 813.81	2.20%	99.41%
2	$ 1,445.75	$ 957.25	2.57%	99.17%
3	$ 1,625.89	$ 1,068.28	2.87%	98.95%
4	$ 1,801.73	$ 1,182.84	3.18%	98.71%
5	$ 1,982.72	$ 1,299.89	3.49%	98.44%
6	$ 2,143.41	$ 1,405.56	3.79%	98.17%
7	$ 2,497.82	$ 1,676.43	4.45%	97.52%

The results show consistent error metric values. On the first prediction day, we obtained an R^2 of 99.41%, remaining above 97% after step 7, which evidences the high predictive capacity of the LSTM network with a good fit despite the uncertainty generated by predictions greater than one day. In addition, other metrics such as MAPE remained below the threshold of 10%, which, according to the literature, is the maximum permissible [31], emphasizing the capacity of the network to generate predictions up to a week later with good performance.

We performed model behavior comparisons during training by predicting the 1-step and 7-step validation sets; see results in Figure 5.



(a) 1 Step (b) 7 steps

Figure 5. Comparison between the 1 and 7-step prediction models

From Figure 5 we can observe a small difference between the actual and predicted data for the 1-step model. On the other hand, the differences are more noticeable for the 7-step model's results. However, both graphs show quite similar behavior, which evidences the fit of the model. The metrics results support this conclusion, with the coefficient of determination remaining above 97%.

The same behavior is observed when extending the analysis beyond 07/11/2024, finding that in the analytical comparison between the prediction of 7 steps following the validation set and the actual data as of 07/18/2024 for the 12 cryptocurrencies (see Table 8), there are good results for the LSTM network. In addition, through the experimental analysis, the results show that the univariate model had better metrics in 75% of the cryptocurrencies; likewise, when making the predictions, they had less variation compared to the actual data, a situation that is not reflected with the multivariate model for the predictions because the metrics are higher than in the univariate model.

Table 8. Comparison of Actual vs Predicted Data with Univariate Model

Variation between actual and predicted values							
Cryptocurrency	**Average**	**Cryptocurrency**	**Average**	**Cryptocurrency**	**Average**	**Cryptocurrency**	**Average**
Bitcoin	2.66%	BNB	6.68%	Solana	10.38%	Shiba Inu	6.38%
Ethereum	5.44%	Dogecoin	8.59%	Litecoin	2.16%	Avalanch	7.26%
Polkadot	2.69%	TRX	3.56%	Cardano	10.45%	XRP	17.57%

As shown in Table 8, there was a weighted average variation of 6.92% depending on the volume of operations between a minimum of 2.16% and a maximum value of 17.57% with respect to the price corresponding to the last characteristic of cryptocurrencies.

Lastly, we compared regression algorithms to demonstrate the efficiency of the methodology proposed in this study using Bitcoin for the following 7 days after the last date of the data set used for training and validation; see results in Table 9.

Table 9. Comparison of regression models

Algorithm	RMSE	MAE	MAPE
LSTM Networks	$5,038.46	$4,425.35	6.83%
XGBoost	$17,849.66	$17,687.05	27.74%
Random Forest	$19,492.41	$19,343.61	30.35%
SVM	$7718.16	$7192.06	11.32%

Overall, the results reflect a comparative analysis of the performance of different algorithms applied to Bitcoin price prediction. The error metrics used, RMSE, MAE, and MAPE, allow accuracy evaluation for each step model. Among the evaluated algorithms, the LSTM networks stand out as the best model, with an RMSE of $5038.46, an MAE of $4425.35, and an MAPE of 6.83%. These values reflect significantly lower errors than the other models. The XGBoost model offers a competitive performance compared to Random Forest, with a MAPE of 27.74%, which demonstrates its robustness as a generalist model. However, it falls short of reaching the accuracy of LSTM because it is not designed to handle temporary dependencies. Similarly, the Random Forest has the highest relative error, with a MAPE of 30.35%, highlighting its main strength in static predictions rather than sequential data. The SVM model, with an MAPE of 24.38%, is ranked as the second-best option. While it may not match the performance of LSTMs, it surpasses the other models evaluated due to its capability to capture nonlinear relationships, given appropriate configuration.

5. Discussion

The results obtained in this study show better performance of the LSTM networks in predicting cryptocurrency prices compared to traditional machine learning algorithms, standing out their ability to model volatile and non-linear time series. We acknowledge the alignment and differences between our results and those reported in previous studies.

Han et al. (2023) [6] reported LSTM networks achieving an R^2 of 93.5% in predicting cryptocurrencies during periods of high volatility, with an average MAPE of 18.3%. Although this result validates the effectiveness of LSTM networks, our findings provide improved results, achieving an R^2 of 99.41% and a MAPE of 2.2% for a one-step trained model. For one week using the multistep model, the MAPE remains below 10%. These results show that our methodology has adequate optimization of hyperparameters and data preprocessing, significantly improving predictions accuracy.

Seabe et al. (2023) [22] identified that recurrent neural networks, particularly LSTMs and variants such as Bi-LSTM, are effective in predicting cryptocurrency prices due to their ability to capture complex temporal dependencies; the authors reported a MAPE of 3.6%, which is higher than that obtained in our research, highlighting that the proposed algorithm has better performance than the one described by the authors and highlights the capacity of LSTM networks to maintain an MAPE below the theoretical limit of 10%.

On the other hand, Jin & Li (2023) [17] used decision tree models as well as the XGBoost model in the prediction of Bitcoin cryptocurrency, obtaining on average an RMSE greater than $50,651 and an MAPE of 0.394% in the short term. Our results, with an RMSE of $1,187.14 and an MAPE of 2.2% for the LSTM network of one step, show that our model provides better results comparing the RMSE values. However, the MAPE in their model is lower than 1%, suggesting that the model proposed by Jin & Li (2023) [17] may be overfitted.

The R^2 of 99.41% obtained during the first day of prediction in our LSTM model demonstrates high predictive capacity. In addition, the model maintained more than 97% performance until the seventh day, highlighting its consistency over time. This behavior also reflects the benefits of deep learning-based methodologies observed in research such as Murray et al. (2023) [23], which argued that LSTM networks outperform other approaches, including hybrid algorithms, in terms of accuracy and generalizability in predicting the prices of multiple cryptocurrencies, even though they argue that their proposal lacks generality, as the solutions are too complex and challenging to reproduce in practice.

The error values obtained in this study, with an RMSE of $1187.14 and an MAPE of 2.20% during training, are within the acceptable margins reported in the literature. In particular, the superiority of LSTM over other models discussed in our research, such as Random Forest with ($19,492.41 RMSE, 30.35% MAPE), shows the effectiveness of LSTMs in handling the non-linear and high-volatility characteristics of cryptocurrency markets. In this regard, Belcastro et al. (2023) [20] emphasized the importance of using advanced algorithms capable of capturing complex relationships, such as the correlation between social sentiment and prices, to improve prediction accuracy.

In our study, optimizing hyperparameters was crucial to obtain superior performance; with a learning ratio of 0.00009, 50 epochs, and a batch size of 150, the LSTM model achieved an optimal balance between accuracy and convergence speed. These results coincide with the quantitative and prospective methodology adopted by Quiroga Juárez & Villalobos Escobedo (2023) [21], who stressed that precise adjustments in predictive models are essential to improve the ability to generate reliable and valuable scenarios for portfolio management. In addition, the results reinforce the relevance of applying innovative and targeted approaches to address the particularities of the cryptocurrency market. For its part, Belcastro et al. (2023) [20] highlighted the effectiveness of algorithms in identifying trends in meme coins. The authors demonstrated that LSTMs have a broader scope, adapting to significant assets like Bitcoin and more complex scenarios. These results positioned LSTM networks as versatile and reliable tools in predictive cryptocurrency analytics.

Finally, studies like Mahdi et al. (2021) [32] used MSV models, reporting an MAPE of 10.5% in traditional financial series. Comparatively, in our study, the LSTM model achieved a MAPE of 6.83%, lower than the 11.32% obtained by the MSV, reaffirming its ability to adapt to aggressive fluctuations in cryptocurrency prices.

Conclusively, LSTM networks proved to be highly effective tools for predicting cryptocurrency prices, outperforming traditional models and machine learning algorithms. These results suggest that this methodology could be applied to other volatile financial assets, offering new analysis and decision-making opportunities in complex markets.

6. Conclusion

This study showed that LSTM networks are highly effective in predicting cryptocurrency prices, specifically for Bitcoin. During the first day of prediction, the model reached an R^2 of 99.41%, remaining above 97% in the following days until the seventh day. The results also indicate consistent and accurate predictive capability, even for predictions more than a week in advance. The LSTM model achieved an RMSE of $1187.14 and a MAPE of 2.20% during the validation process, both values within the acceptable limits according to the literature. The MAPE remained below 10% at all steps, supporting the network's ability to make accurate and reliable predictions in the short to medium term.

The LSTM model outperformed other ML algorithms predicting the seven days following the training and validation set. This model achieved a lower RMSE of $5038.46 compared to the MSV's RMSE of $7718.16 and the XGBoost's $17,849.66. These results highlight the LSTM model's ability to handle nonlinear and volatile time series more effectively than the other models. In addition, the LSTM obtained a MAPE of 6.83%, lower than the 11.32% of the MSV algorithm, the 27.74% of the XGBoost model, and the 30.35% of the Random Forest. This performance highlights the LSTM's accuracy in predicting cryptocurrency prices, significantly outperforming the other algorithms evaluated. The optimal hyperparameters for the LSTM model were a learning ratio of 0.00009, 50 epochs, and a batch size of 150. These values allowed for obtaining the best performance in terms of error, achieving a significant reduction in the RMSE, and maintaining a low relative error in the predictions. These results suggest that hyperparameter optimization is crucial to improving model accuracy.

7. Declarations

7.1. Author Contributions

Conceptualization, M.J.C. and G.A.F.F.; methodology, M.J.C.; software, G.A.F.F.; validation, G.A.F.F. and M.J.C.; formal analysis, M.J.C.; investigation, M.J.C.; resources, M.J.C. and G.A.F.F.; data curation, G.A.F.F. and M.J.C.; writing—original draft preparation, M.J.C.; writing—review and editing, M.J.C.; visualization, M.J.C. and G.A.F.F.; supervision, M.J.C.; project administration, M.J.C.; funding acquisition, M.J.C. All authors have read and agreed to the published version of the manuscript.

7.2. Data Availability Statement

The data presented in this study are available in the article.

7.3. Funding and Acknowledgments

This research is being funded by the National University of Piura-Peru, specifically from the funds of the Basic and Applied Research Projects Competition - 2024 call.

7.4. Institutional Review Board Statement

Not applicable.

7.5. Informed Consent Statement

Not applicable.

7.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. References

[1] Squarepants, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. SSRN Electronic Journal. doi:10.2139/ssrn.3977007.

[2] Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? Finance Research Letters, 20, 192–198. doi:10.1016/j.frl.2016.09.025.

[3] Box, G.E.P. and Jenkins, G.M. (1970) Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, United States.

[4] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. doi:10.1016/0304-4076(86)90063-1.

[5] 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. doi:10.1016/j.cam.2019.112395.

[6] Han, D., Liu, P., Xie, K., Li, H., Xia, Q., Cheng, Q., Wang, Y., Yang, Z., Zhang, Y., & Xia, J. (2023). An attention-based LSTM model for long-term runoff forecasting and factor recognition. Environmental Research Letters, 18(2), 24004. doi:10.1088/1748-9326/acaedd.

[7] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669. doi:10.1016/j.ejor.2017.11.054.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735.

[9] Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. Neural Computing and Applications, 32(23), 17351–17360. doi:10.1007/s00521-020-04867-x.

[10] Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. Entropy, 21(6), 589–589. doi:10.3390/e21060589.

[11] Regal, A., Morzán, J., Fabbri, C., Herrera, G., Yaulli, G., Palomino, A., & Gil, C. (2019). Cryptocurrency price projection based on tweets using LSTM. Ingeniare. Revista Chilena de Ingeniería, 27(4), 696–706. doi:10.4067/s0718-33052019000400696.

[12] Mariappan, L. T., Pandian, J. A., Kumar, V. D., Geman, O., Chiuchisan, I., & Năstase, C. (2023). A Forecasting Approach to Cryptocurrency Price Index Using Reinforcement Learning. Applied Sciences (Switzerland), 13(4), 2692–2692–2704. doi:10.3390/app13042692.

[13] Hamayel, M. J., & Owda, A. Y. (2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. AI (Switzerland), 2(4), 477–496. doi:10.3390/ai2040030.

[14] Yang, Y., Xiong, J., Zhao, L., Wang, X., Hua, L., & Wu, L. (2023). A Novel Method of Blockchain Cryptocurrency Price Prediction Using Fractional Grey Model. Fractal and Fractional, 7(7), 547–565. doi:10.3390/fractalfract7070547.

[15] Sung, S. H., Kim, J. M., Park, B. K., & Kim, S. (2022). A Study on Cryptocurrency Log-Return Price Prediction Using Multivariate Time-Series Model. Axioms, 11(9), 448. doi:10.3390/axioms11090448.

[16] Maleki, N., Nikoubin, A., Rabbani, M., & Zeinali, Y. (2023). Bitcoin price prediction based on other cryptocurrencies using machine learning and time series analysis. Scientia Iranica, 30(1 E), 285–301. doi:10.24200/sci.2020.55034.4040.

[17] Jin, C., & Li, Y. (2023). Cryptocurrency Price Prediction Using Frequency Decomposition and Deep Learning. Fractal and Fractional, 7(10), 708–708–736. doi:10.3390/fractalfract7100708.

[18] Kang, C. Y., Lee, C. P., & Lim, K. M. (2022). Cryptocurrency Price Prediction with Convolutional Neural Network and Stacked Gated Recurrent Unit. Data, 7(11), 149–149–161. doi:10.3390/data7110149.

[19] Aljadani, A. (2022). DLCP2F: a DL-based cryptocurrency price prediction framework. Discover Artificial Intelligence, 2(1). doi:10.1007/s44163-022-00036-2.

[20] Belcastro, L., Carbone, D., Cosentino, C., Marozzo, F., & Trunfio, P. (2023). Enhancing Cryptocurrency Price Forecasting by Integrating Machine Learning with Social Media and Market Data. Algorithms, 16(12), 542. doi:10.3390/a16120542.

[21] Quiroga Juárez, C. A., & Villalobos Escobedo, A. (2023). Cryptocurrency market scenarios based on a statistical study. Revista CEA, 9(20), e2530. doi:10.22430/24223182.2530.

[22] Seabe, P. L., Moutsinga, C. R. B., & Pindza, E. (2023). Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach. Fractal and Fractional, 7(2), 203. doi:10.3390/fractalfract7020203.

[23] Murray, K., Rossi, A., Carraro, D., & Visentin, A. (2023). On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles. Forecasting, 5(1), 196–209. doi:10.3390/forecast5010010.

[24] Samson, T. K. (2024). Comparative Analysis of Machine Learning Algorithms for Daily Cryptocurrency Price Prediction. Information Dynamics and Applications, 3(1), 64–76. doi:10.56578/ida030105.

[25] Fang, F., Chung, W., Ventre, C., Basios, M., Kanthan, L., Li, L., & Wu, F. (2024). Ascertaining price formation in cryptocurrency markets with machine learning. European Journal of Finance, 30(1), 78–100. doi:10.1080/1351847X.2021.1908390.

[26] Kiranmai Balijepalli, N. S. S., & Thangaraj, V. (2025). Prediction of cryptocurrency's price using ensemble machine learning algorithms. European Journal of Management and Business Economics, 244. doi:10.1108/EJMBE-08-2023-0244.

[27] Hossain, M. F. B., Lamia, L. Z., Rahman, M. M., & Khan, M. M. (2024). FinBERT-BiLSTM: A Deep Learning Model for Predicting Volatile Cryptocurrency Market Prices Using Market Sentiment Dynamics. arXiv Preprint, arXiv:2411.12748. doi:10.48550/arXiv.2411.12748.

[28] Islam, M. Z., Islam, M. S., Montaser, M. A. A., Rasel, M. A. B., Bhowmik, P. K., Dalim, H. M., & pant, L. (2024). Evaluating the Effectiveness of Machine Learning Algorithms in Predicting Cryptocurrency Prices Under Market Volatility: A Study Based on the Usa Financial Market. The American Journal of Management and Economics Innovations, 06(12), 15–38. doi:10.37547/tajmei/volume06issue12-03.

[29] Lee, K., Lim, H., Hwang, J., & Lee, D. (2024). Evaluating missing data handling methods for developing building energy benchmarking models. Energy, 308, 132979. doi:10.1016/j.energy.2024.132979.

[30] Núñez Sánchez, N. D. (2023). Analysis of Algorithms for Data Imputation and Prediction Models. Ph.D. Thesis, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. (In Spanish).

[31] Camarillo-Peñaranda, J. R., Saavedra-Montes, A. J., & Ramos-Paja, C. A. (2013) Recommendations for Selecting Indexes for Model Validation. TecnoLógicas, 109. doi:10.22430/22565337.372.

[32] Mahdi, E., Leiva, V., Mara'beh, S., & Martin-Barreiro, C. (2021). A new approach to predicting cryptocurrency returns based on the gold prices with support vector machines during the COVID-19 pandemic using sensor-related data. Sensors, 21(18), 6319. doi:10.3390/s21186319.