Available online at www.HighTechJournal.org

HighTech and Innovation Journal

Vol. 6, No. 2, June, 2025

ISSN: 2723-9535

Closing Price Prediction of Cryptocurrencies BTC, LTC, and ETH Using a Hybrid ARIMA-LSTM Algorithm

Jherson S. Ruiz-López [1*], Miguel Jiménez-Carrión [1]

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

Received 28 March 2025; Revised 19 May 2025; Accepted 26 May 2025; Published 01 June 2025

Abstract

This study aims to develop a hybrid algorithm using the ARIMA model and LSTM-type recurrent neural networks to predict the closing prices of the cryptocurrencies BTC, LTC, and ETH. The methodology includes an exploratory data analysis, followed by the design, implementation, and evaluation of each individual algorithm as well as the combined hybrid algorithm. The results, after experimentation and evaluation of metrics on the test set, indicated that the ARIMA model was inefficient in predicting the closing prices of cryptocurrencies. On the other hand, the hybrid model for BTC showed significant statistical differences in the metrics, with MAE = $726.21 and MAPE = 1.75%, compared to the LSTM model, which achieved MAE = $729.35 and MAPE = 1.76%. These results indicate better performance from the hybrid model. Regarding the RMSE metric, the hybrid model scored 1157.47, while LSTM scored 1159.99; although statistically equivalent, the hybrid model was numerically better. For the remaining metrics and other cryptocurrencies, both methods were statistically equivalent. For five-day-ahead predictions, the hybrid algorithm continued to yield better results for LTC and ETH.

Keywords: LSTM Networks; ARIMA Model; Hybrid Approach; Prediction; Cryptocurrencies.

1. Introduction

In recent years, cryptocurrencies have gained a significant presence in financial markets, becoming an attractive alternative for investors seeking high returns in contexts characterized by high uncertainty and price fluctuations [1]. Their operation is based on blockchain technology—a decentralized digital architecture that guarantees integrity, security, and traceability in transactions—key aspects that have driven their rapid adoption and global expansion. Among the cryptocurrencies with the highest market capitalization and recognition are Bitcoin (BTC), Litecoin (LTC), and Ethereum (ETH), which have captured the attention of institutional players, retail investors, and academics due to their worldwide accessibility and potential for appreciation [2]. However, the high volatility that characterizes these assets—driven by exogenous factors such as government regulations, macroeconomic indicators, and market perception—makes it difficult to formulate robust investment strategies [3]. This complexity has spurred the development and implementation of advanced analytical models aimed at improving the understanding and prediction of cryptocurrency market behavior.

Traditionally, statistical models such as Moving Average, ARIMA, or logistic regression have been used for financial time series forecasting due to their ability to capture linear and seasonal patterns [4, 5]. However, these models have

* Corresponding author: 0502020015@alumnos.unp.edu.pe

http://dx.doi.org/10.28991/HIJ-2025-06-02-09

limitations when applied to the highly nonlinear data typical of the cryptocurrency market. As a result, machine learning approaches—such as recurrent neural networks (RNN), Long Short-Term Memory (LSTM), or Gated Recurrent Unit (GRU)—have proven effective in capturing long-term dependencies and nonlinear relationships [6, 7]. Despite the strong performance of machine learning models, recent research suggests that hybrid approaches offer a promising alternative. Studies by Xu et al. (2022) [8], Pierre et al. (2023) [9], and Bouteska et al. (2024) [10] show that combining statistical models like ARIMA with deep neural networks such as RNN, GRU, or LSTM can considerably improve accuracy in forecasting tasks related to demand, production, and finance.

Numerous studies have attempted to predict the prices of various cryptocurrencies such as those mentioned above. What is evident, however, is that no robust algorithmic method has yet been established for predicting the prices of these assets, which, although potentially profitable, are highly risky. Efforts have largely focused on using traditional tools such as the ARIMA model, alongside other machine learning algorithms. In this regard, López (2023) [3] asserts that temporal patterns are more representative than spatial patterns when estimating the future closing price of Bitcoin. Lambis et al. (2023) [6] state that the reliability of forecasts should be evaluated by analyzing the autocorrelation of the errors. Despite this, it is frequently observed that hybrid models incorporating LSTM yield better results than individual models. However, current literature lacks analyses that demonstrate statistically significant differences between the algorithms used. Additionally, the datasets in many studies are relatively limited, often excluding the full time horizon and failing to make multi-step predictions—let alone predictions beyond the dataset under study.

In this context, the present research proposes a hybrid ARIMA-LSTM model to predict the closing prices of BTC, LTC, and ETH. The objective is to compare the performance of the hybrid model against the individual models and to demonstrate statistical significance in the metrics used for both the individual methods and the proposed hybrid algorithm. This study aims to contribute to the theoretical and practical development of predictive models in a market characterized by speculation, high liquidity, and low interest rates—an environment where robust predictive tools are essential for investment planning, risk management, and improved resource utilization [11].

The rest of the article is structured as follows: Section 2 presents a general review of the previous literature; Section 3 outlines the methodology applied in the research; Section 4 discusses the results obtained from the individual models and the hybrid approach; Section 5 analyzes and discusses the results; and Section 6 presents the conclusions of the study.

2. Literature Review

In the last decade, cryptocurrencies have aroused notable interest in the financial sphere, both for their usefulness as a medium of exchange and for their potential to generate high returns. In 2017, the Bitcoin boom occurred, reaching a value of $19,000. This boom further fueled the creation and adoption of new cryptocurrencies, as well as the entry of investors and businesses into the market [12]. However, the inherent volatility of these assets has posed significant challenges in predicting their prices, increasing investment risk [13, 14].

Price forecasting in financial markets has traditionally been approached using statistical models such as ARIMA. In the context of cryptocurrencies, studies such as the one by Azari (2019) [4] have explored the applicability of the ARIMA model for forecasting the price of Bitcoin. Although this approach showed some predictive ability in single-trended intervals, significant limitations were identified—especially in capturing the abrupt fluctuations and high volatility characteristic of the cryptocurrency market. Mangiwa et al. (2025) [15], in their research, explore the use of the ARIMA model for the prediction of the ETH cryptocurrency. After the identification, estimation, and comparison of models, a MAPE of 15.01% and an RMSE of 649.702 was achieved. The authors emphasize that although the metrics obtained are reasonably good, the ARIMA model performs better with short-term predictions, managing to identify trends in the series; however, its performance decreases over longer time periods.

Due to the limitations of traditional models in capturing the volatility of cryptocurrency markets, several authors have chosen to address machine learning models, particularly those designed for processing sequential and temporal data, such as recurrent neural networks (RNN) and their variants LSTM and GRU. These models have gained relevance due to their ability to model complex and nonlinear relationships. Research such as that of Lambis et al. (2023) [6] uses Deep Learning models—RNN, LSTM, GRU, and a combination of convolutional networks with LSTM (CNN-LSTM)—for forecasting the closing prices of cryptocurrencies BTC and ETH, achieving in their research MAPE values of 2.64% and 3.31%, and R^2 values of 98.73% and 98.85% for RNN with BTC and LSTM with ETH, respectively.

Zhang (2024) [16] applied a hybrid CNN-LSTM approach for BTC price prediction. After experimentation with different hyperparameter settings, the hybrid approach achieved a minimum MAPE value of 2.75%. On the other hand, Tumpa & Maduranga (2024) [17] explored three recurrent networks—LSTM, GRU, and Bi-LSTM—for the prediction of BTC, ETH, and LTC. All three models showed good performance in predicting the three cryptocurrencies. For BTC, the best model was Bi-LSTM, with a MAPE of 1.94%, while for ETH and LTC, the best model was GRU, with MAPE values of 1.85% and 4.25%, respectively.

Pirkhedri (2025) [18] used LSTM and GRU networks to predict the price of the Solana cryptocurrency, studying the impact of the "number of epochs"—with values of 70, 90, 100, and 150—on the performance of the algorithms. This study concludes that both models have high predictive ability and obtained the best metrics with the highest number of epochs proposed, i.e., 150 epochs. Furthermore, it is emphasized that GRU networks had higher performance than LSTM networks, reaching RMSE, MAE, and R^2 values of 8.209, 6.323, and 0.935, respectively, while LSTM networks obtained values of 8.964, 7.016, and 0.922. Kaur et al. (2025) [19] also tested LSTM and GRU networks, expanding the scope to three cryptocurrencies: LTC, ETH, and BTC. Consistent with other studies, both models exhibited good performance; however, the authors highlight GRU networks, as they achieved better results in two of the three proposed cryptocurrencies (ETH and BTC), obtaining MAPE values of 8.037%, 4.415%, and 3.54% for LTC, ETH, and BTC, respectively. For LSTM, the MAPE values were 7.653%, 5.075%, and 9.162%.

Other research contrasts traditional models with Deep Learning (DL) models in predicting complex time series. For example, the study by Yu (2024) [5] explores three statistical models—moving average (MA), logistic regression (LR), and ARIMA—and two DL models—LSTM and a hybrid convolutional network with LSTM (CNN-LSTM). The five models were used to predict BTC, LTC, and ETH. Their research indicates that DL models showed better performance in contrast to traditional models, which usually work better with linear patterns or short-term periods. This is reflected in the metrics: DL models achieved MAPE values between 4% and 8%, while traditional models ranged from 60% to over 1000%. In addition, this study highlights the performance of the CNN-LSTM hybrid model in obtaining the best results for all three cryptocurrencies, demonstrating that hybrid models have the potential to combine the strengths of individual models to improve predictive performance.

Kabo et al. (2025) [20] contrast ARIMA models and LSTM networks for the prediction of Bitcoin. This research also includes, in addition to the historical data of the cryptocurrency, economic indicators such as GDP growth rates and sentiment data extracted from X via Tweepy for Twitter. The study concluded that LSTM networks performed better than ARIMA models. Furthermore, both approaches performed better when using historical prices combined with GDP and sentiment data. ARIMA achieved RMSE, MAE, and R^2 values of 2518.35, 2081.66, and 91.44%, respectively, using the combined data, while LSTM achieved values of 1717.65, 1253.24, and 96.02%.

Within the field of predictive models, hybrid models that combine statistical techniques with deep learning methods have also been proposed. Research such as Xu et al. (2022) [8] addresses hybrid ARIMA-LSTM and ARIMA-SVR models and contrasts them with individual models for drought prediction, observing that the hybrid approach improved the accuracy in predicting complex events. Other authors, such as Pierre et al. (2023) [9] and Fan et al. (2021) [21], address the ARIMA-LSTM hybrid model for the prediction of peak power consumption and well production, respectively. The main idea behind combining the ARIMA model with a DL model is to leverage ARIMA's ability to capture linear patterns, while LSTM captures the nonlinear components of complex time series. Both authors propose a similar hybrid model structure in which ARIMA makes the initial predictions and an LSTM network is trained to predict the residuals generated by ARIMA, so that the final predictions are the sum of the ARIMA and LSTM outputs. In both investigations, the authors report that the ARIMA-LSTM hybrid outperforms the individual models.

Dave et al. (2021) [22] also address the ARIMA-LSTM hybrid, focusing on forecasting Indonesian exports. The structure of the hybrid model in this research consists of separating the components of the time series into trend, seasonality, and residuals, and training ARIMA to predict the trend and LSTM to predict seasonality and residuals. This study also concludes that the hybrid model performs better than the independent models, obtaining a MAPE of 7.38% versus 8.56% and 9.38% for LSTM and ARIMA, respectively.

3. Research Methodology

The methodology of this research is structured as follows: in the first stage, data collection was performed, in which the historical data of the cryptocurrencies under study—BTC, LTC, and ETH—were collected from the Investing platform. On the extracted data, an exploratory data analysis was conducted to identify missing data and outliers (anomalous or inconsistent values). Then, data cleaning was carried out to eliminate missing values and outliers. For this purpose, KNN imputation was applied. After cleaning, statistical analyses were used to verify that the imputed variables had not been significantly altered and that the data were ready for model design.

In all models, the data were normalized using the Min-Max method to facilitate training, and then split into a training set and a test set. For ARIMA modeling, manual experimentation was performed with ranges for the hyperparameters: "p" from 0 to 5, "d" from 0 to 2, and "q" from 0 to 5. The selection criteria for all models explored in this research were the metrics obtained on the test set: root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). In the case of ARIMA, the Akaike Information Criterion (AIC) was also considered. To determine whether the metrics obtained were satisfactory, thresholds commonly used in the literature were referenced. According to Gutiérrez & De la Vara (2008) [23], R^2 values close to 100% are desirable, with 70% being the minimum recommended. Based on Klimberg et al. (2010) [24], a model with a MAPE of less than 10% is considered highly accurate.

For modeling with LSTM networks, four approaches were tested: a univariate-unistep approach using the closing price as the sole predictor; a multivariate-unistep approach using the six variables provided by Investing as predictors; a multivariate-unistep approach applying the PFI method of Fisher et al. (2019) [25] to select the three most relevant features; and a univariate-multistep approach. Various combinations of hyperparameters (number of LSTM layers and Dropout layers, Dropout rate, LSTM units, learning rate, batch size, number of epochs, number of steps in the input, and percentage of data dedicated to training) were tested to predict the closing price of the cryptocurrencies under study.

For the hybrid ARIMA-LSTM model, the ARIMA configurations with the best performance—corresponding to each training percentage—were selected, and the same LSTM configuration as in the individual model was used. Two approaches were addressed: univariate-unistep and univariate-multistep. In the proposed hybrid model, both the ARIMA model and the LSTM networks were trained with the same training data. The ARIMA model generated the initial predictions of the time series, and the LSTM networks were trained to predict the residuals generated by ARIMA. Additionally, for both the LSTM networks and the hybrid model, an analysis was performed to forecast values outside the historical data used—i.e., data beyond the test set.

For the discussion of the results, the findings obtained in this research were contrasted with those from other authors. Finally, conclusions were drawn, summarizing the results obtained. The procedure is shown in Figure 1.

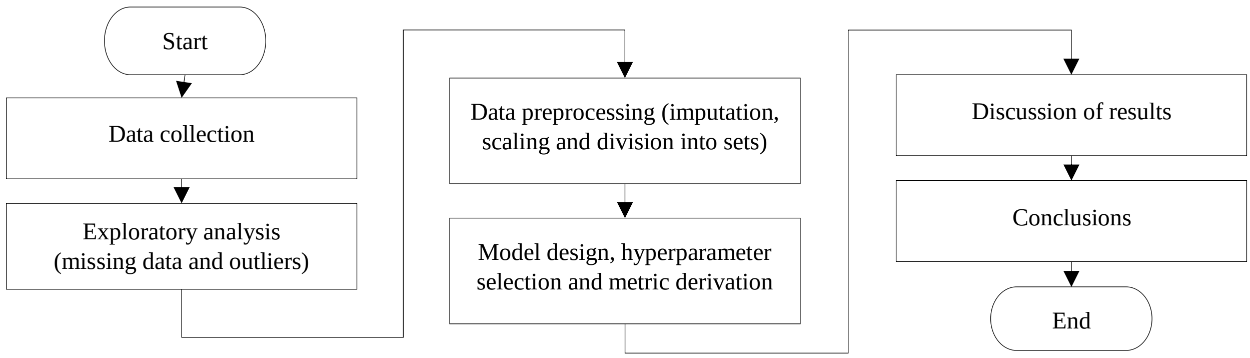

Figure 1. Research procedures

4. Results

4.1. Data Collection, Exploratory Analysis and Data Preprocessing

BTC, LTC, and ETH were selected as they are three of the ten most recognized and capitalizable cryptocurrencies in the market [2]. The historical data were extracted from the Investing platform, recorded in USD, with a daily frequency. The data collection periods are shown in Table 1.

Table 1. Collection periods for BTC, LTC and ETH cryptocurrencies

Cryptocurrency	Data collected		
	Home	End	Quantity
BTC	18/07/2010	30/10/2024	5219
LTC	24/08/2016	30/10/2024	2990
ETH	10/03/2016	30/10/2024	3157

The Investing.com platform provides six variables for the historical data of cryptocurrencies:

- ***Opening price:*** Initial price recorded on a given day
- ***Closing price:*** Final or closing price recorded on a given day
- ***Maximum price:*** Highest price reached during the day
- ***Minimum price:*** Lowest price reached during the day
- ***Volume:*** Represents the flow of transactions measured in monetary units during the day
- ***% Variation:*** Represents the daily variability of the cryptocurrency

According to Regal et al. (2019) [26], in the cryptocurrency market, the variable typically targeted for prediction is the closing price or closing price. In this study, this will also be the variable to be predicted.

The exploratory analysis focused on identifying missing data and outliers. Table 2 shows the percentage of missing data for the three cryptocurrencies. For this, the criteria of Dagnino (2014) [27] were followed, which establish that a percentage below 10% is acceptable for the validity of data usage, as a higher percentage may introduce significant bias.

Table 2. Percentage of missing data for BTC, LTC and ETH

Cryptocurrency	Percentage of missing data					
	Closing	**Opening**	**Max.**	**Min.**	**Vol.**	**% var.**
BTC	0.00%	0.00%	0.00%	0.00%	0.11%	0.00%
LTC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ETH	0.00%	0.00%	0.00%	0.00%	0.25%	0.00%

As shown in Table 2, no variable exceeds 1%, making all variables valid for use. To identify outliers, Boxplot diagrams were used. Figure 2 presents the Boxplots for the closing price of BTC, LTC, and ETH. While several outliers appear in the data, they were not considered anomalies due to the high inherent volatility of cryptocurrencies.

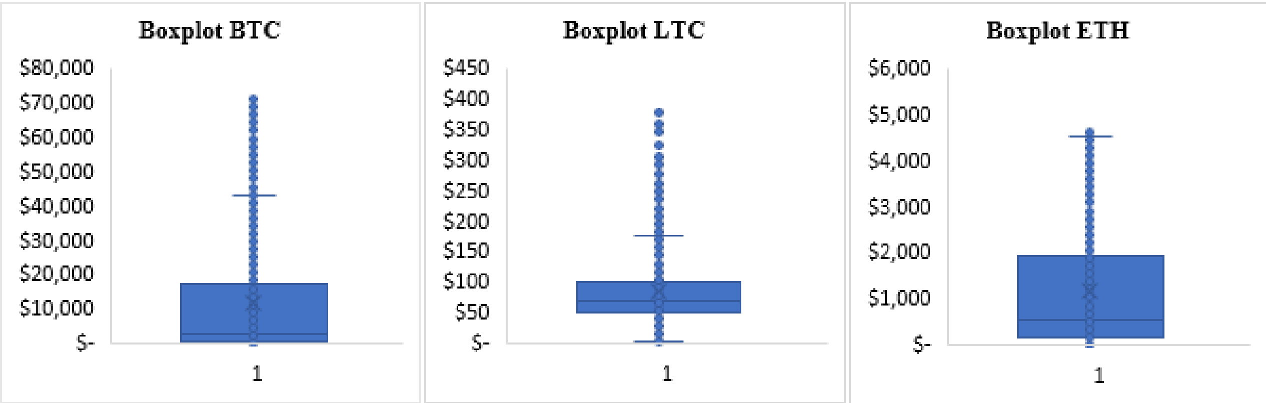

Figure 2. Boxplot Plots for the Closing Price of BTC, LTC and ETH

For the treatment of missing data, since the percentage was null or very low for all variables, the k-nearest neighbor (KNN) imputation method was applied. A value of $k = 3$ was chosen, as Nuñez Sánchez (2023) [28] recommended using an odd k and notes that this value is directly proportional to bias and computation time. After imputation, it was verified that the statistics of the variables with missing data (BTC volume and ETH volume) had not been significantly altered. The results are shown in Table 3, where the variation is observed to be less than 1%. Therefore, the data was considered valid for model design.

Table 3. Percentage change in statistics after imputation

Percentage change in statistics		
Variable	**BTC volume**	**ETH Volume**
Mean	-0.11%	-0.24%
Standard deviation	-0.06%	-0.13%
Minimum	0.00%	0.00%
Maximum	0.00%	0.00%

4.2. Design of ARIMA Models

The data was scaled to a range from -1 to 1 to facilitate training, as normalization helps prevent overfitting and avoid bias. For partitioning the training and test sets, different ratios were tested: 80–20%, 85–15%, and 90–10%. Regarding the ARIMA hyperparameters "p," "d," and "q," values in the ranges of 0 to 5 for "p" and "q," and 0 to 2 for "d" were explored. These ranges were selected because values above those thresholds reduce the model's ability to abstract the linear behavior of the time series.

Likewise, the Akaike Information Criterion (AIC) was used to evaluate performance, as the optimization of ARIMA must balance predictive accuracy with model simplicity and computational cost. Additionally, performance was assessed using the metrics obtained on the test set.

After experimentation, it was observed that the best results were obtained with 90% of the data used for training and 10% for testing in the case of BTC, and with 85% for training and 15% for testing in the cases of LTC and ETH. Table 4 presents the selected ARIMA configurations along with their corresponding test set performance metrics.

Table 4. Selected ARIMA models and metrics obtained in the test set

Cryptocurrency	Model	AIC	RMSE	MAE	MAPE	R^2
BTC	(3, 2, 0)	-23526.04	$9683.78	$7102.32	13.79%	64.03%
LTC	(4, 0, 2)	-9282.81	$11.20	$8.70	11.71%	-38.80%
ETH	(5, 2, 1)	-11477.21	$555.29	$451.47	16.99%	34.18%

As shown in Table 4, MAPE values between 10% and 20% were obtained, which, according to Klimberg et al. (2010) [24], indicate a good forecast. However, the R^2 values are below 70%—and even negative in the case of LTC—indicating that the ARIMA model explains little or none of the variability in the data. This makes it an inefficient model for predicting cryptocurrency prices, which exhibit volatile nature and highly nonlinear patterns.

4.3. Design of LSTM Networks

For the LSTM networks, three unistep approaches were addressed: a univariate approach using the "Closing Price" as the sole predictor; a multivariate approach using the six variables provided by Investing as predictors; and another multivariate approach in which the Permutation Feature Importance (PFI) method by Fisher et al. (2019) [25] was applied to select the three most relevant features for predicting the closing price. A univariate-multistep approach was also explored, and its results are presented below to compare performance with the hybrid model. During data preprocessing and model training, models with the highest predictive capacity were prioritized, as well as features with high importance according to the PFI method.

Before designing the models, the data was scaled to the range [-1, 1] to improve training stability and ensure that variables in multivariate models were on the same scale. The data was divided into training and test sets, with ratios of 80–20%, 85–15%, and 90–10% tested. The optimizer used was ADAM, and the loss function was Mean Square Error (MSE), both widely used in the literature and supported by studies such as López (2023) [3], and Yu (2024) [5], Arranz Barcenilla (2022) [29].

After experimentation, the best training–testing ratios were determined to be 85–15% for BTC and LTC, and 80–20% for ETH, as these partitions yielded more favorable performance metrics. The best-performing configuration consisted of 1 input step, 2 LSTM layers with a 20% dropout rate, 750 LSTM units, a learning rate of 0.00009, and batch size and epochs set to 70. Additionally, the experimentation revealed that LSTM networks appear to be particularly sensitive to the following hyperparameters: number of input steps, number of LSTM units, learning rate, batch size, and training percentage.

For the PFI method, RMSE was selected as the reference metric. As this is a regression model, RMSE is closely aligned with the objective of minimizing variability relative to actual data. Four iterations were conducted for each cryptocurrency to obtain average results. The three most important variables—"Closing," "Maximum," and "Minimum"—were found to be the same across all three cryptocurrencies. The feature importance results are shown in Table 5, and the test set metrics for the three LSTM approaches are presented in Table 6.

Table 5. Feature Importance for multivariate LSTM models

BTC		LTC		ETH	
Variable	**FI**	**Variable**	**FI**	**Variable**	**FI**
Closing	45.60	Closing	3.45	Maximum	13.98
Maximum	34.30	Maximum	2.41	Closing	11.72
Minimum	24.71	Minimum	1.92	Minimum	9.09
Opening	17.00	Opening	1.87	Opening	8.23
% var.	1.00	% var.	1.26	% var.	1.15
Vol.	1.00	Vol.	1.00	Vol.	1.00

Table 6. LSTM network metrics obtained on test set

Cryptocurrency	Model	RMSE	MAE	MAPE	R^2
BTC	Univariate	$1159.99	$729.35	1.76%	99.60%
	Multivariate	$1149.80	$741.40	1.83%	99.61%
	Multivariate with PFI	$1161.73	$730.06	1.77%	99.60%
LTC	Univariate	$2.54	$1.66	2.23%	92.86%
	Multivariate	$2.55	$1.70	2.29%	92.81%
	Multivariate with PFI	$2.59	$1.68	2.26%	92.58%
ETH	Univariate	$75.30	$49.21	1.98%	98.80%
	Multivariate	$75.16	$49.04	1.99%	98.80%
	Multivariate with PFI	$75.69	$48.87	1.97%	98.78%

In general, it was observed that LSTM networks demonstrated higher predictive performance than ARIMA models, with MAPE values around 2%. This indicates that the error relative to the mean has a maximum representative variation of 2%, which, according to Klimberg et al. (2010) [24], qualifies as highly accurate forecasting. In terms of R^2, values above 92% were obtained—exceeding 98% and even 99% for ETH and BTC, respectively. This reflects a high proportion of explained variance in the predicted variable based on its input data.

Similarly, the RMSE and MAE metrics also showed superior performance. For instance, in the univariate model for BTC, LSTM achieved an RMSE of $1,159.99 and an MAE of $729.35, compared to ARIMA's $9,683.78 and $7,102.32. For LTC, LSTM achieved RMSE and MAE values of $2.54 and $1.66, versus $11.20 and $8.70 with ARIMA. For ETH, LSTM obtained RMSE and MAE values of $75.30 and $49.21, compared to ARIMA's $555.29 and $451.47. These results highlight the strong predictive ability of LSTM networks to capture complex and highly nonlinear patterns, which is illustrated in Figure 3, where the actual and predicted values from the univariate-unistep approach appear to overlap, indicating a highly accurate fit.

Regarding the unistep approaches, no significant performance differences were observed among the cryptocurrencies studied. However, this research favors univariate models due to their less complex structure, which still delivers high performance. Slightly better metrics were observed for BTC and LTC using univariate models.

As for the PFI method, similar metrics were obtained when compared with the multivariate models without PFI. In some cases, slightly lower MAE and MAPE values were achieved, suggesting that PFI is a useful approach in this context, as it maintains strong performance while reducing model complexity. Additionally, PFI can be applied in other scenarios involving sentiment data, economic indicators, or multiple predictors, helping to select the most relevant features while reducing both model complexity and processing time.

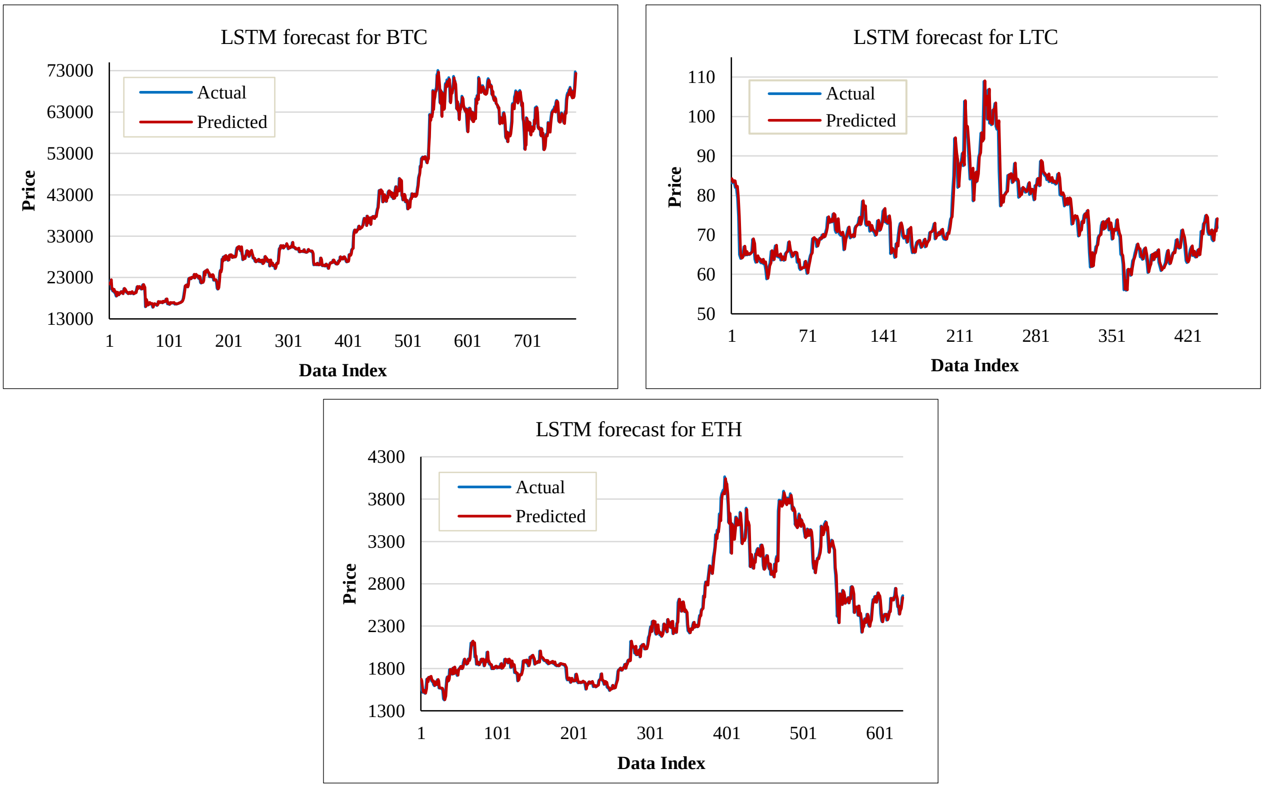

Figure 3. Behavior of the LSTM model on the test set for BTC, LCT and ETH

4.4. Design of Hybrid Models

Hybrid models are proposed with the objective of leveraging the strengths of individual models. In this research, a hybrid approach combining the ARIMA model with LSTM networks is proposed. As noted by Bouteska et al. (2024) [10], the central idea is to combine ARIMA's ability to extract linear patterns with LSTM's ability to capture complex or highly nonlinear patterns.

The proposed approach is described as follows: with the preprocessed data, the ARIMA model is fitted and generates predictions that capture the linear component of the series L(t). By comparing the ARIMA predictions with the actual values, the residuals—which represent the nonlinear component of the series N(t)—are obtained. These residuals are then predicted using LSTM networks. Finally, both the linear component predicted by ARIMA and the nonlinear component predicted by LSTM are added together to obtain the hybrid predictions.

This approach has already been applied in previous research such as Fan et al. (2021) [21], Pierre et al. (2023) [9], and Xu et al. (2022) [8], where it was observed that the hybrid ARIMA + LSTM model outperformed the individual models. The general representation is shown in Equation (1) and illustrated in the diagram in Figure 4.

$$\mathrm{HYBRID} = \mathrm{residuals}[(\mathrm{ARIMA}(p, d, q)] + \mathrm{LSTM}(\mathrm{mShortTerm}, \mathrm{mLongTerm}) \tag{1}$$

Regarding the architectures, the training-testing percentages of the LSTM networks were used, i.e. 85-15% for BTC and LTC and 80-20% for ETH. As mentioned earlier, these partitions showed better performance in the initial training phases, as reflected in the error metrics. The ARIMA configurations used were those that showed the best performance during experimentation for each training percentage: (0, 2, 0) for BTC, (4, 0, 2) for LTC, and (2, 2, 4) for ETH. The LSTM configuration matched the one used in the individual univariate model, as it performed well with lower complexity.

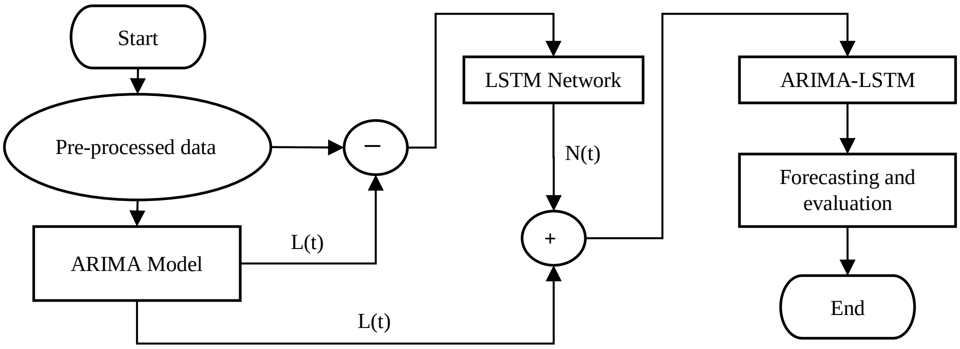

Figure 4. Hybrid model procedure

The performance metrics obtained on the test set for the hybrid model are shown in Table 7. Figure 5 compares the actual and predicted values, where the hybrid model's results were observed to be graphically similar to those of the LSTM network.

Table 7. Hybrid model metrics obtained on test set

Cryptocurrency	RMSE	MAE	MAPE	R^2
BTC	\$1157.47	\$726.21	1.75%	99.60%
LTC	\$2.53	\$1.67	2.23%	92.92%
ETH	\$75.88	\$49.71	1.99%	98.78%

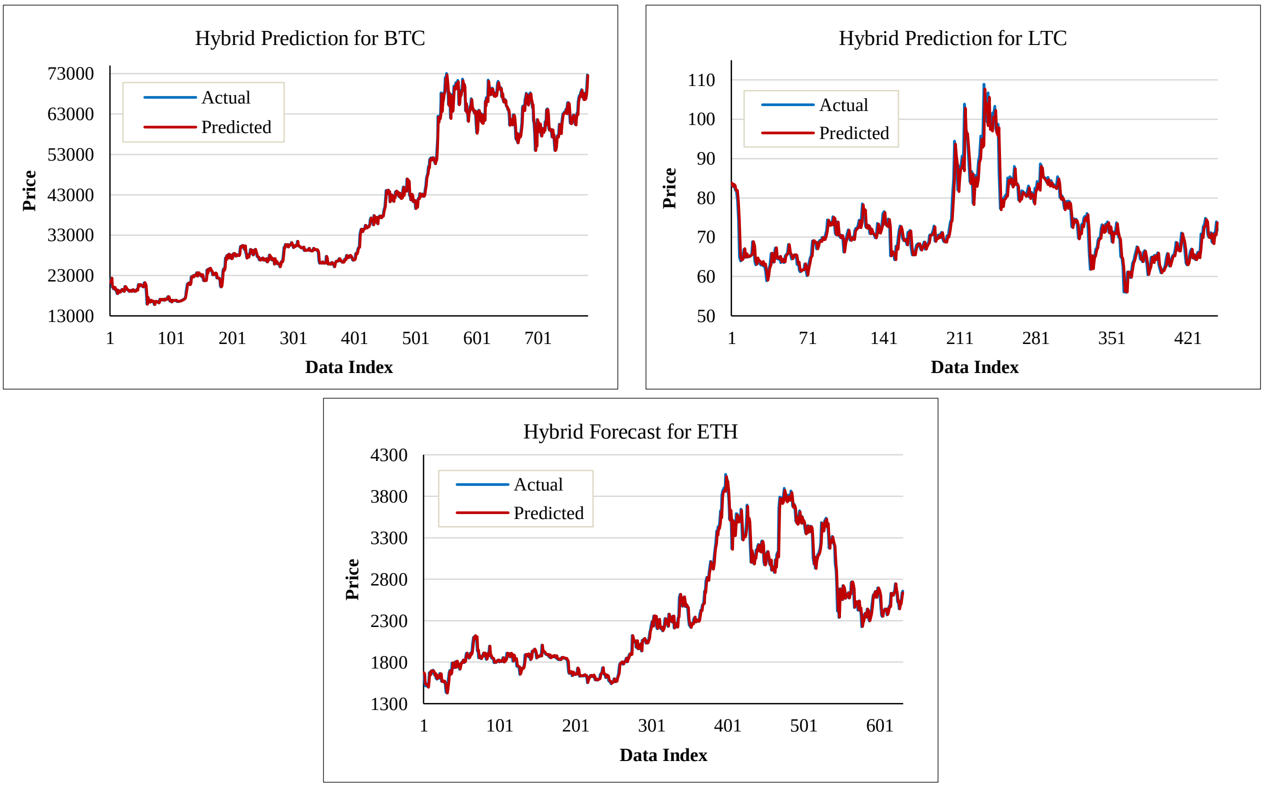

Figure 5. Behavior of the Hybrid model on the test set for BTC, LCT and ETH

The metrics of the hybrid model are also similar to those of the LSTM networks, with MAPE values around 2% and R^2 values exceeding 92%. However, the Bootstrap method—widely used in robust statistics and algorithm comparison—was applied, revealing statistically significant differences in some metrics between the hybrid model and the LSTM network. At a 95% confidence level, the mean difference in MAE was 4.4190, with a confidence interval ranging from 0.8555 to 8.1133, indicating a statistically significant difference.

Similarly, for the MAPE metric, the mean difference was 0.0002, with a 95% confidence interval from 0.0000 to 0.0003—also statistically significant. In contrast, no significant difference was found for the RMSE metric, where the mean difference was 1.6604 and the confidence interval ranged from -1.6831 to 4.7555 (as the interval includes 0). For R^2, the mean difference was -0.0000 with a confidence interval from -0.0000 to 0.0000, suggesting no statistically significant difference. Nonetheless, the hybrid model showed slightly better numerical values, which can be observed graphically in Figure 6.

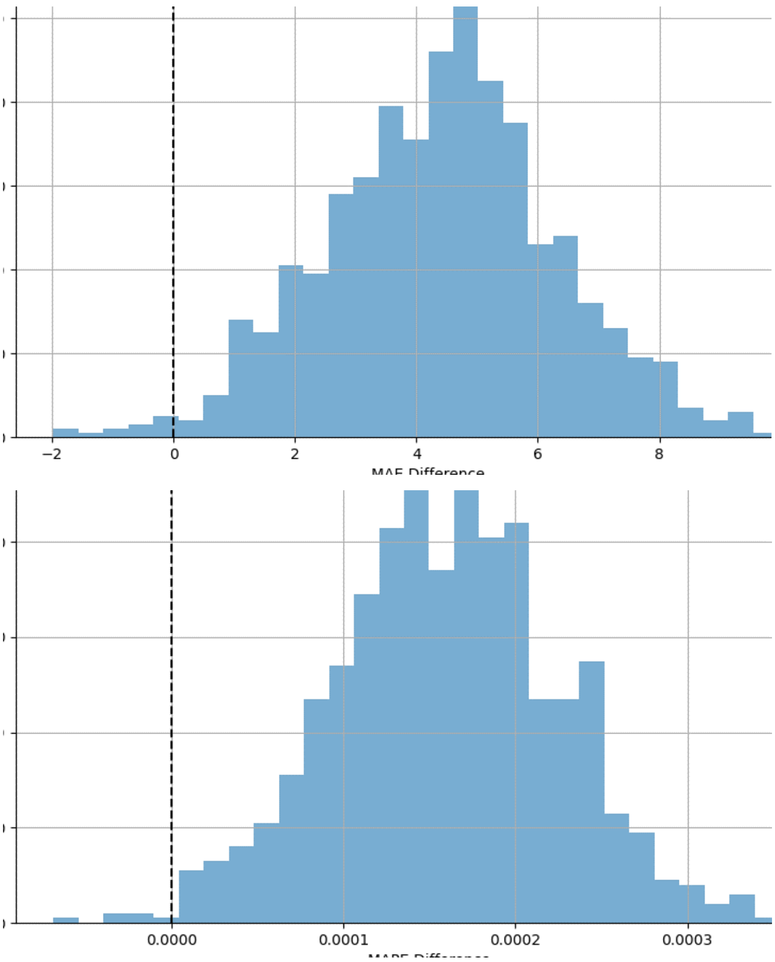

Figure 6. Bootstrap distribution of the MAE and MAPE metrics of the difference (LSTM - Hybrid), for BTC

4.5. LSTM and Hybrid Model Performance in Multi-step Predictions

To further compare the predictive ability of both models, multi-step predictions were performed over 5 days. Table 8 shows the overall performance metrics on the test set for both models, while Table 9 breaks down the results by prediction step.

Table 8. General metrics obtained on the test set for multi-step predictions (5 steps)

Cryptocurrency	Model	RMSE	MAE	MAPE	R^2
BTC	LSTM	$1920.46	$1262.26	3.14%	98.89%
	Hybrid	$1925.56	$1242.36	3.03%	98.89%
LTC	LSTM	$4.28	$2.83	3.86%	79.82%
	Hybrid	$4.53	$3.21	4.30%	77.41%
ETH	LSTM	$128.69	$84.18	3.39%	96.50%
	Hybrid	$129.36	$84.87	3.40%	96.45%

From Table 8, it can be seen that both models generally achieve MAPE values around 4% and R^2 values above 77%, reaching a maximum of 98% for BTC. These results are considered good performance levels in the literature, with efficient predictability demonstrated by MAPE values below 10% and determination coefficients exceeding 77%, 96%, and 98% for LTC, ETH, and BTC, respectively. For LTC, the LSTM model showed slightly better performance than the hybrid model, while for BTC and ETH no considerable differences were observed. This is further reflected in Table 9, where metrics are broken down for each of the 5 steps.

Table 9. Per-step metrics obtained on the test set for multi-step predictions

Cryptocurrency	Step	LSTM				Hybrid			
		RMSE	MAE	MAPE	R^2	RMSE	MAE	MAPE	R^2
BTC	1	$1151.86	$761.43	1.93%	99.60%	$1150.21	$727.60	1.77%	99.60%
	2	$1545.49	$1024.14	2.57%	99.28%	$1554.46	$1009.44	2.47%	99.27%
	3	$1908.94	$1287.78	3.20%	98.91%	$1916.04	$1273.95	3.11%	98.90%
	4	$2211.76	$1520.12	3.77%	98.54%	$2219.22	$1499.73	3.66%	98.53%
	5	$2487.90	$1717.82	4.23%	98.15%	$2490.67	$1701.10	4.15%	98.15%
LTC	1	$2.62	$1.72	2.32%	92.48%	$2.57	$1.73	2.31%	92.77%
	2	$3.47	$2.36	3.20%	86.76%	$3.56	$2.55	3.44%	86.05%
	3	$4.29	$2.96	4.04%	79.71%	$4.42	$3.27	4.40%	78.40%
	4	$4.87	$3.35	4.57%	73.85%	$5.21	$3.86	5.19%	70.01%
	5	$5.53	$3.78	5.17%	66.09%	$6.04	$4.62	6.18%	59.58%
ETH	1	$75.33	$49.22	1.99%	98.80%	$75.48	$49.39	1.99%	98.80%
	2	$104.18	$68.88	2.79%	97.71%	$104.60	$69.44	2.79%	97.69%
	3	$127.20	$86.82	3.50%	96.58%	$127.91	$87.78	3.51%	96.53%
	4	$148.84	$101.27	4.08%	95.31%	$149.72	$102.18	4.08%	95.24%
	5	$167.17	$114.72	4.62%	94.07%	$168.08	$115.55	4.61%	93.99%

Likewise, performance was tested on data outside the historical data used—specifically, for 5 future days from 10/31/2024 to 11/04/2024. The MAPE values for the predicted data from both models are shown in Table 10. For BTC, the LSTM model had the lowest error with a MAPE of 2.06% and an RMSE of $1,519.64. For LTC, the hybrid model performed better, with a MAPE of 1.80% and RMSE of $1.50. For ETH, both models performed similarly, with the hybrid model slightly outperforming LSTM with a MAPE of 7.52% and RMSE of $189.86.

In general, MAPE values are below 10%, described in the literature as highly accurate forecasts, corroborating the good performance of the Hybrid and LSTM models for predicting the closing price of cryptocurrencies.

Table 10. RMSE and MAPE of predicted values by LSTM and Hybrid

Cryptocurrency	Model	RMSE	MAPE
BTC	LSTM	$1519.64	2.06%
	Hybrid	$2181.61	2.98%
LTC	LSTM	$4.35	5.38%
	Hybrid	$1.50	1.80%
ETH	LSTM	$192.19	7.57%
	Hybrid	$189.86	7.52%

5. Discussion of Results

The approach adopted in this research centers on a hybrid LSTM+ARIMA model. The core objective of this work is to evaluate the capacity and performance of predictive tools—specifically, LSTM networks, which are known for their ability to model time series and capture highly complex relationships, and the ARIMA model, which excels at identifying and expressing linear relationships through regression-based methods. While the general processes of data handling, LSTM training, and ARIMA parameterization are framed within the CRISP-DM methodology, the elements of innovation and novelty contributed by this study lie in the construction of the hybrid model itself.

This model, developed by the authors, operates by generating a reference line with a slope defined by ARIMA, which captures the core linear behavior of the time series. The residuals—representing what ARIMA cannot explain—are then predicted by the LSTM network and added to the baseline. This combination produces a final prediction that leverages the strengths of both components, achieving a more comprehensive and accurate forecast by integrating the linear structure modeled by ARIMA with the nonlinear dynamics captured by LSTM.

The results of the research show that the ARIMA model is deficient for the prediction of complex time series such as cryptocurrency prices, obtaining MAPE values above 10% and R^2 values below 70%, even reaching negative values in the case of LTC. These results fall outside the thresholds recommended by Klimberg et al. (2010) [24] and Gutiérrez & De la Vara (2008) [23]. This finding also aligns with the results obtained by Azari (2019) [4], who noted that ARIMA models perform better in sub-periods with a single trend or more linear patterns but are not suitable for time horizons characterized by numerous fluctuations or complex patterns, such as those inherent to the volatile nature of cryptocurrencies. Specifically for ETH, ARIMA obtained an RMSE of $555.29 and a MAPE of 16.99%, results similar to those reported by Mangiwa et al. (2025) [15], who obtained an RMSE of $649.702 and a MAPE of 15.01%, and emphasized that ARIMA models perform better for short-term predictions. The limitations of ARIMA in cryptocurrency forecasting are also evident in Yu's (2024) [5] research, which compared deep learning (DL) models with traditional models, concluding that LSTM achieves MAPE values between 6% and 8%, whereas ARIMA reaches MAPE values between 62% and 210%.

LSTM networks, on the other hand, are designed to efficiently handle long-term temporal dependencies, such as price changes that may be influenced not only by recent days but also by events from weeks or months earlier. This characteristic inspired the development of such recurrent networks, which can retain relevant information over time through their input, forget, and output gates. These mechanisms make LSTM networks more robust than traditional recurrent neural networks (RNNs). Additionally, LSTM networks can learn to filter out irrelevant information during training, proving more resistant to noise than other models. This is particularly advantageous for modeling the behavior of cryptocurrency prices, which are highly nonlinear and chaotic, responding to news, speculation, social media (e.g., tweets), regulation, and more.

The LSTM networks demonstrated strong predictive capability, achieving MAPE values around 2% and R^2 values above 92% in the case of LTC—and even higher than 98% and 99% for ETH and BTC, respectively. These results are consistent with the findings of Lambis et al. (2023) [6], who reported MAPE values of 2.71% and 3.31%, and R^2 values of 98.68% and 98.85% for BTC and ETH, respectively. Regarding the unistep models, no substantial differences were observed among the three approaches. However, considering model complexity, this research recommends the univariate model. The PFI method proposed by Fisher et al. (2019) [25] also showed slightly improved performance in the multivariate models for BTC and LTC and reduced model complexity for ETH while maintaining comparable performance. This supports its usefulness in selecting relevant predictor variables or features.

The metrics obtained by the LSTM networks in this study are similar to—or even better than—those reported in existing literature on DL algorithms. For example, Kabo et al. (2025) [20] reported MAE, RMSE, and R^2 values of 1253.24, 1717.65, and 96.02% for BTC prediction using LSTM, whereas this research achieved values of 729.35, 1159.99, and 99.60%, respectively. Similarly, Tumpa & Maduranga (2024) [17] and Kaur et al. (2025) [19] reported MAPE values of 1.94%, 4.25%, and 1.85% (first study), and 3.54%, 7.65%, and 4.42% (second study) for BTC, LTC, and ETH, respectively—compared to 1.76%, 2.23%, and 1.98% in this research.

The proposed ARIMA-LSTM hybrid model showed good performance, comparable to the LSTM networks, and was slightly better in the case of BTC. These findings are consistent with those of other researchers such as Pierre et al. (2023) [9], Fan et al. (2021) [21], and Xu et al. (2022) [8], who found substantial improvements with the hybrid approach over individual models—though their studies were not applied to cryptocurrency time series. Bouteska et al. (2024) [10], in their research on DL and hybrid models (ARIMA-LSTM and ARIMA-MLP) for cryptocurrency price prediction, observed that individual models performed better, and hybrid approaches are not always superior in every scenario.

In multi-step predictions (five steps), both the LSTM and hybrid models achieved MAPE values between 3% and 4.3%, and R^2 values above 77% for LTC and higher than 96% and 98% for ETH and BTC, respectively. These metrics are considered valid and indicative of good performance according to the literature. The results of both models were similar, with the hybrid model performing slightly better for BTC and the LSTM network performing slightly better for LTC and ETH.

In predictions made on data outside the historical data, the LSTM model had the best performance for BTC, with a MAPE of 2.06% and an RMSE of $1,519.64. For LTC, the hybrid model performed best, with a MAPE of 1.80% and an RMSE of $1.50. For ETH, both models produced similar results, with the hybrid model slightly outperforming LSTM, yielding a MAPE of 7.52% and an RMSE of $189.86.

The degradation in prediction accuracy as the number of forecast steps increases is due to the high volatility of cryptocurrencies. This is corroborated in Table 9, which shows the step-by-step breakdown of metrics for all cryptocurrencies. Nevertheless, the predictions remain within the recommended thresholds by Klimberg et al. (2010) [24], with MAPE values under 10%—considered highly accurate. Likewise, the R^2 values comply with the guidelines of Gutiérrez & De la Vara (2008) [23], who recommend R^2 values close to 100%, with a minimum acceptable value of 70%, except in the case of LTC, where R^2 fell below 70% at the fifth step for both models.

For trading applications, it is recommended to develop a system capable of collecting real-time data, including exogenous variables such as sentiment indicators and macroeconomic data. Furthermore, confidence intervals should be established based on the level of risk aversion the user is willing to tolerate.

6. Conclusion

The hybrid ARIMA-LSTM model demonstrated competitive performance in predicting cryptocurrency prices, excelling in both single-step and multi-step (five-step) forecasts, as well as in extrapolations beyond historical data. In comparative terms, the hybrid approach achieved accuracy metrics similar to those of pure LSTM networks, consistently outperforming the ARIMA model. LSTM networks, due to their ability to capture complex temporal dependencies, delivered outstanding performance—achieving MAPE values close to 2% and coefficients of determination (R^2) exceeding 92% for Litecoin (LTC), 98% for Ethereum (ETH), and 99% for Bitcoin (BTC). These results demonstrate their capacity to adapt to the high volatility inherent in cryptocurrencies, effectively recognizing non-linear patterns and hidden trends.

In contrast, the ARIMA model exhibited significant limitations, as reflected in MAPE values above 10% and R^2 values below 70%—even negative in the case of LTC. This indicates that ARIMA's capacity to model nonlinear time series is insufficient, particularly in high-volatility contexts. The inferior performance of ARIMA compared to both the LSTM and hybrid models confirms that traditional time series modeling techniques are inadequate for complex data such as cryptocurrency prices. In conclusion, the hybrid ARIMA-LSTM model positions itself as a robust solution by combining the strengths of both approaches. However, it is acknowledged that further optimization of hyperparameters and exploration of alternative architectures could enhance its predictive accuracy even more.

7. Declarations

7.1. Author Contributions

Conceptualization, M.J.C. and J.S.R.L.; methodology, M.J.C.; software, J.S.R.L.; validation, J.S.R.L. and M.J.C.; formal analysis, M.J.C.; investigation, M.J.C. and J.S.R.L.; resources, M.J.C. and J.S.R.L.; data curation, J.S.R.L. and M.J.C.; writing—original draft preparation, M.J.C.; writing—review and editing, M.J.C.; visualization, M.J.C. and J.S.R.L.; 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

Data was obtained from the Investing.com platform and are available at https://www.investing.com/.

7.3. Funding and Acknowledgments

This research is being funded by the National University of Piura-Peru, specifically with funds from 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] Acosta Valderrama, J. C. (2018). SABER UCV: Study of Cryptocurrencies Using Time Series and Neural Networks. Universidad Central de Venezuela, Caracas, Venezuela.

[2] E.F.P.A. (2024). Financial advisors, their great challenges. Asesores Financieros, EFPA (Asociación Española de Asesores y Planificadores Financiero), 15, 18–20. Available online: https://www.asesoresfinancierosefpa.es/ejemplares-revista-asesores-financieros-efpa/los-grandes-desafios-del-sector-en-portada-del-numero-15-de-asesores-financieros-efpa/ (accessed on May 2025).

[3] López-García, A. (2023). Comparative analysis of deep neural networks for Bitcoin market price prediction. Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA, 24(1), 1–16. doi:10.24310/recta.24.1.2023.19858.

[4] Azari, A. (2019). Bitcoin price prediction: An ARIMA approach. arXiv preprint arXiv:1904.05315, 1-3. doi:10.48550/arXiv.1904.05315.

[5] Yu, S. (2024). Comparative analysis of machine learning techniques for cryptocurrency price prediction. Applied and Computational Engineering, 32(1), 1–12. doi:10.54254/2755-2721/32/20230175.

[6] Jiménez Gómez, L. M., Lambis-Alandete, E., & Velásquez-Henao, J. D. (2023). Comparison of Deep Learning Algorithms for Cryptocurrency Price Forecasting. Ingeniería y Competitividad, 25(3). doi:10.25100/iyc.v25i3.12845.

[7] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. Proceedings: IEEE International Conference on Big Data, Big Data 2019, 3285–3292. doi:10.1109/BigData47090.2019.9005997.

[8] Xu, D., Zhang, Q., Ding, Y., & Zhang, D. (2022). Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting. Environmental Science and Pollution Research, 29(3), 4128–4144. doi:10.1007/s11356-021-15325-z.

[9] Pierre, A. A., Akim, S. A., Semenyo, A. K., & Babiga, B. (2023). Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches. Energies, 16(12), 4739. doi:10.3390/en16124739.

[10] Bouteska, A., Abedin, M. Z., Hajek, P., & Yuan, K. (2024). Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods. International Review of Financial Analysis, 92, 1–12. doi:10.1016/j.irfa.2023.103055.

[11] Flores-Fernandez, G. A., Jimenez-Carrion, M., Gutierrez, F., & Sanchez-Ancajima, R. A. (2024). Genetic Algorithm and LSTM Artificial Neural Network for Investment Portfolio Optimization. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15(2), 27–46. doi:10.58346/JOWUA.2024.I2.003.

[12] Cuthbertson, A. (2020). Bitcoin price reaches its highest value in history. Independentespanol. Available online: https://www.independentespanol.com/tecnologia/bitcoin-precio-valor-mas-alto-historia-b1764093.html (accessed on May 2025).

[13] Astuti, I. D., Rajab, S., & Setiyouji, D. (2022). Cryptocurrency Blockchain Technology in the Digital Revolution Era. APTISI Transactions on Technopreneurship, 4(1), 9–15. doi:10.34306/att.v4i1.216.

[14] Darda, S. D. (2023). Advantages & Disadvantages of Cryptography & Valuation Analysis. National High School Journal of Science, 1–11.

[15] Mangiwa, R. D. D., Siregar, R., Sari, S. D., & Agustina, N. (2025). Application of the Arima Model in Forecasting Ethereum Prices. Parameter: Jurnal Matematika, Statistika Dan Terapannya, 4(1), 81–94. doi:10.30598/parameterv4i1pp81-94.

[16] Zhang, X. (2024). Analyzing Financial Market Trends in Cryptocurrency and Stock Prices Using CNN-LSTM Models. Preprints, 1-12. doi:10.20944/preprints202407.1119.v1.

[17] Tumpa, S. N., & Maduranga, K. D. G. (2024). Utilizing RNN for real-time cryptocurrency price prediction and trading strategy optimization. arXiv preprint arXiv:2411.05829. doi:10.48550/arXiv.2411.05829.

[18] Pirkhedri, A. (2025). Predicting Solana Cryptocurrency Prices in 2024: A Comparative Study of LSTM and GRU Models. International Journal of Industrial Engineering and Operational Research, 7(2), 11-19..

[19] Kaur, R., Uppal, M., Gupta, D., Juneja, S., Arafat, S. Y., Rashid, J., Kim, J., & Alroobaea, R. (2025). Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum. PeerJ Computer Science, 11, 2675. doi:10.7717/peerj-cs.2675.

[20] Kabo, I. G., Obunadike, G. N., & Samaila, N. A. (2025). Sentiment-Driven and Economic Indicators for Bitcoin Price Forecasting: A Hybrid Time Series Model. Journal of Science Research and Reviews, 2(1), 99–107. doi:10.70882/josrar.2025.v2i1.34.

[21] Fan, D., Sun, H., Yao, J., Zhang, K., Yan, X., & Sun, Z. (2021). Well production forecasting based on ARIMA-LSTM model considering manual operations. Energy, 220, 119708. doi:10.1016/j.energy.2020.119708.

[22] Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. (2021). Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM. Procedia Computer Science, 179, 480–487. doi:10.1016/j.procs.2021.01.031.

[23] Gutiérrez, H., & De la Vara, R. (2008). Análisis y diseño de experimentos. México DF: McGraw-Hill Interamericana. Available online: https://gc.scalahed.com/recursos/files/r161r/w19537w/analisis_y_diseno_experimentos.pdf (accessed on May 2025).

[24] Klimberg, R. K., Sillup, G. P., Boyle, K. J., & Tavva, V. (2010). Forecasting performance measures - What are their practical meaning? Advances in Business and Management Forecasting, 7, 137–147. doi:10.1108/S1477-4070(2010)0000007012.

[25] Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. Journal of Machine Learning Research, 20, 1–81.

[26] 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.

[27] Jorge Dagnino, S. (2014). Datos faltantes (missing values). Revista Chilena de Anestesia, 43(4), 332–334.

[28] Nuñez Sánchez, N. D. (2023). Análisis de algoritmos para la imputación de datos y modelos de predicción: Informe final de pasantía para optar por el título de Matemática. Universidad Distrital Francisco José de Caldas. Available online: https://repository.udistrital.edu.co/server/api/core/bitstreams/d84b225f-429e-4517-9481-0a416f41d0a9/content (accessed on May 2025).

[29] Arranz Barcenilla, V. (2022). Aplicación web para la gestión de carteras de inversión usando técnicas de Inteligencia Artificial (Final Degree Project, Universidad de Valladolid, Escuela de Ingeniería Informática). UVaDOC. Available online: https://uvadoc.uva.es/handle/10324/57220 (accessed on May 2025).