

## Article

# The BTC Price Prediction Paradox Through Methodological Pluralism

Mariya Paskaleva <sup>1</sup> and Ivanka Vasenska <sup>2,\*</sup>

<sup>1</sup> Department of Finance and Accounting, South-West University “Neofit Rilski”, 2700 Blagoevgrad, Bulgaria; m.gergova@abv.bg

<sup>2</sup> Department of Tourism, South-West University “Neofit Rilski”, 2700 Blagoevgrad, Bulgaria

\* Correspondence: ivankav@swu.bg or ivankav48@gmail.com; Tel.: +359-899-953-396

## Abstract

Bitcoin’s extreme price volatility presents significant challenges for investors and traders, necessitating accurate predictive models to guide decision-making in cryptocurrency markets. This study compares the performance of machine learning approaches for Bitcoin price prediction, specifically examining XGBoost gradient boosting, Long Short-Term Memory (LSTM), and GARCH-DL neural networks using comprehensive market data spanning December 2013 to May 2025. We employed extensive feature engineering incorporating technical indicators, applied multiple machine and deep learning models configurations including standalone and ensemble approaches, and utilized cross-validation techniques to assess model robustness. Based on the empirical results, the most significant practical implication is that traders and financial institutions should adopt a dual-model approach, deploying XGBoost for directional trading strategies and utilizing LSTM models for applications requiring precise magnitude predictions, due to their superior continuous forecasting performance. This research demonstrates that traditional technical indicators, particularly market capitalization and price extremes, remain highly predictive in algorithmic trading contexts, validating their continued integration into modern cryptocurrency prediction systems. For risk management applications, the attention-based LSTM’s superior risk-adjusted returns, combined with enhanced interpretability, make it particularly valuable for institutional portfolio optimization and regulatory compliance requirements. The findings suggest that ensemble methods offer balanced performance across multiple evaluation criteria, providing a robust foundation for production trading systems where consistent performance is more valuable than optimization for single metrics. These results enable practitioners to make evidence-based decisions about model selection based on their specific trading objectives, whether focused on directional accuracy for signal generation or precision of magnitude for risk assessment and portfolio management.



Academic Editor: Stanislaus Maier-Paape, Qiji Zhu and Dayong Huang

Received: 31 July 2025

Revised: 26 August 2025

Accepted: 28 August 2025

Published: 4 October 2025

**Citation:** Paskaleva, Mariya, and Ivanka Vasenska. 2025. The BTC Price Prediction Paradox Through Methodological Pluralism. *Risks* 13: 195. <https://doi.org/10.3390/risks13100195>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** BTC price prediction; hybrid models; investors’ decision-making; efficient trading strategies

## 1. Introduction

The application of machine learning techniques to Bitcoin price prediction has gained considerable attention due to the cryptocurrency’s high volatility and potential for significant financial returns (Omole and Enke 2024). Recent comparative studies have demonstrated that deep learning models, particularly Long Short-Term Memory (LSTM) networks, exhibit superior performance compared to traditional time series models, such as ARIMA,

in capturing the non-linear and non-stationary nature of cryptocurrency data (Bouteska et al. 2024; Chen et al. 2024). However, empirical findings regarding the superiority of deep learning approaches remain mixed, with some studies reporting that ensemble tree-based methods, such as XGBoost, consistently outperform neural networks for tabular financial data prediction tasks (Hafid et al. 2024; Ranjan et al. 2023; Zhu et al. 2023).

Feature selection has emerged as a critical component for improving model performance, with studies showing that methods such as Boruta, genetic algorithms, and Light Gradient Boosting Machine can significantly enhance prediction accuracy by addressing the curse of dimensionality inherent in high-dimensional financial datasets. The integration of on-chain blockchain data alongside traditional price features has shown promise for improving Bitcoin price predictions, with research indicating that transaction volumes, network hash rates, and market capitalization provide valuable predictive signals (Kjærland et al. 2018; Visharad et al. 2025). Advanced hybrid architectures combining convolutional neural networks with LSTM models have demonstrated exceptional performance, with some studies reporting accuracy rates exceeding 82% for directional price prediction tasks (Livieris et al. 2021; Omole and Enke 2024).

Ensemble learning approaches that combine multiple algorithms have shown particular promise, with research indicating that stacking models can outperform individual predictors by leveraging the complementary strengths of different algorithms (Ji et al. 2019; Ye et al. 2022). Recent developments in 2024 have introduced transformer-based architectures and attention mechanisms for cryptocurrency forecasting, with the Helfomer model demonstrating improved performance by integrating Holt–Winters exponential smoothing with deep learning components (Lee 2024). The practical application of these models through algorithmic trading strategies has shown substantial potential returns, with some studies reporting annual returns exceeding 6000% when using deep learning-guided trading approaches (Omole and Enke 2024). Despite these advances, the field continues to face challenges related to model stability, overfitting, and the inherent unpredictability of cryptocurrency markets, highlighting the need for robust validation frameworks and risk management strategies in practical implementations (Mudassir et al. 2020; Park and Seo 2022).

Some researchers claim that the best deep learning models for Bitcoin price prediction are the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models (Kervanci et al. 2024). Both architectures capture temporal patterns and long-term dependencies within the sequential Bitcoin price data, showcasing remarkable predictive capabilities. Their performance was highlighted by high R-squared scores, indicating their effectiveness in forecasting Bitcoin prices, despite their computational intensity potentially limiting real-time deployment. The research findings suggest that the Random Forest model demonstrated near-perfect accuracy in predicting Bitcoin prices, outperforming other models such as GRU and Support Vector Machine (SVM), which also showed strong predictive capabilities. This highlights the effectiveness of ensemble learning approaches in capturing complex patterns in cryptocurrency price data. The study highlights the strengths and limitations of various models, demonstrating that deep learning architectures, such as LSTM and GRU, are particularly adept at modelling long-term dependencies and capturing intricate temporal patterns within Bitcoin price fluctuations, thereby providing valuable insights into their applicability in the cryptocurrency market. Dutta et al. (2020) identified the GRU model with recurrent dropout as the best-performing deep learning model for Bitcoin price prediction. It outperforms popular existing models when evaluated using root mean squared error (RMSE) (Dutta et al. 2020). The research highlights the significance of robust feature engineering. It compares various advanced machine learning methods, demonstrating that the GRU model, when combined with simple trading strategies, can

yield financial gains in cryptocurrency trading. The study investigates a framework that employs advanced machine learning methods, specifically focusing on the GRU model, which incorporates recurrent dropout to enhance performance in predicting daily Bitcoin prices. The study found that the GRU model with recurrent dropout outperforms popular existing models in predicting daily Bitcoin prices, as evidenced by lower root mean squared error (RMSE) values in their experimental results. The research also indicates that implementing simple trading strategies alongside the proposed GRU model, with appropriate learning, can result in financial gains.

These results are supported by [Seabe et al. \(2023\)](#). In their study, the best deep machine learning model for predicting the Bitcoin price is the Bi-Directional LSTM (Bi-LSTM). It outperformed other models, specifically the LSTM and GRU models, with a Mean Absolute Percentage Error (MAPE) of 0.036 for BTC. The research indicates that Bi-LSTM provides the most accurate predictions among the tested algorithms, making it the preferred choice for forecasting Bitcoin prices. The experimental results suggest that the proposed prediction models are effective in predicting cryptocurrency prices, offering valuable insights for investors and traders. Furthermore, they suggest that future research should investigate additional factors influencing cryptocurrency prices, such as social media activity and trading volumes. [Alizadegan et al. \(2024\)](#) utilize four powerful machine learning algorithms for forecasting Bitcoin prices: Light Gradient Boosting Machine (Light Gradient Boosting Machine (LightGBM)), LSTM, Bi-LSTM, and XGBoost. The best deep learning models for Bitcoin price prediction identified in the research are the LSTM and the Bi-LSTM models. These models, which are types of recurrent neural networks (RNNs), excel at capturing long-term dependencies in time series data, making them particularly effective for forecasting Bitcoin prices. The study highlights their predictive performance compared to traditional methods, demonstrating their potential to enhance decision-making in cryptocurrency trading and investment strategies ([Alizadegan et al. 2024](#)). The authors focus on evaluating the predictive performance of these algorithms using Mean Absolute Error (MAE) and RMSE as the evaluation metrics, highlighting their effectiveness in enhancing the accuracy of Bitcoin price predictions compared to traditional methods. The study's results highlight the importance of employing advanced machine learning techniques in forecasting financial time series, demonstrating the potential for enhanced decision-making in cryptocurrency trading and investment strategies.

[Wang et al. \(2024\)](#) present a research paper that introduces a hybrid deep learning model combining LSTM and GRU networks for predicting Bitcoin prices. This model employs data decomposition via the CEEMDAN approach and feature selection using a Random Forest model to enhance prediction accuracy. The empirical results indicate that this hybrid method outperforms benchmark models and achieves higher returns on investment in simulated trading, making it a strong candidate for predicting Bitcoin prices ([Wang et al. 2024](#)). The proposed hybrid deep learning model, which integrates feature selection through the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) approach and utilizes a Random Forest model for predictor importance measurement, demonstrates superior performance in predicting Bitcoin prices compared to benchmark models, as verified by the MCS test. The hybrid method not only enhances the accuracy of Bitcoin price predictions but also yields significantly higher returns on investment in simulated trading scenarios compared to other benchmark models, providing valuable insights for investors. [Lee \(2024\)](#) identified two advanced attention-based deep learning models for Bitcoin price prediction: the Attention-LSTM and the Attention-GRU. The Attention-GRU model is noted for its computational efficiency, making it suitable for real-time applications, while the Attention-LSTM model excels in capturing long-term dependencies. Both models integrate moving average technical indicators,

particularly MACD, which significantly enhances prediction accuracy. These findings provide valuable insights for cryptocurrency traders and investors seeking practical tools for forecasting Bitcoin price movements. The study concludes that integrating moving average technical indicators, particularly the MACD, with attention-based deep learning models significantly enhances the accuracy of Bitcoin price movement predictions, allowing for better classification of market trends into uptrend, downtrend, and neutral categories. The Attention-GRU model was identified as more computationally efficient for real-time applications. In contrast, the Attention-LSTM model was noted for its superior ability to capture long-term dependencies in the data, providing valuable insights for cryptocurrency traders and investors.

[Şimşek Türker et al. \(2025\)](#) found that the XGBoost algorithm outperformed all other models in forecasting the price of Bitcoin, as indicated by consistently higher performance metrics across various statistical measures, including RMSE, MAE, MAPE, and coefficient of determination ( $R^2$ ). Following XGBoost, the hybrid model combining Convolutional Neural Network (CNN) and Bi-LSTM (CNN-BiLSTM) showed the next best performance, followed by the CNN and LSTM models. In contrast, the Support Vector Regression (SVR) model exhibited the least favorable performance in predicting Bitcoin prices. The study identifies the best deep learning models for Bitcoin price prediction as the hybrid model combining a CNN and a BiLSTM model, followed by the CNN and LSTM models ([Şimşek Türker et al. 2025](#)). These models demonstrated advantageous results in forecasting Bitcoin prices, with the CNN-BiLSTM model showing extreme performance. However, the XGBoost algorithm outperformed all deep learning models overall in terms of predictive accuracy. [Bourday et al. \(2024\)](#) found that the transformer with the XGBoost model significantly outperformed baseline models in predicting Bitcoin prices, achieving a MAE of 0.011 and RMSE of 0.018, indicating its effectiveness in managing the complexities of the cryptocurrency market ([Bourday et al. 2024](#)). The research highlights the first exploration of transformer-based architectures for feature extraction in financial market predictions, demonstrating that advanced deep learning techniques can provide substantial improvements over traditional forecasting methods, thereby offering valuable insights for investors in the cryptocurrency markets. The study highlights the effectiveness of advanced deep learning techniques in managing the complexities of the cryptocurrency market, providing significant improvements over traditional forecasting methods.

On the other hand, the application of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models to cryptocurrency data remains a cornerstone methodological approach, with comprehensive comparative studies of twelve different GARCH specifications applied to seven major cryptocurrencies establishing that various GARCH-type models effectively capture volatility dynamics ([Chu et al. 2017](#)). Recent empirical research demonstrates that ARIMA-GARCH combined models consistently outperform simple ARIMA specifications, with EGARCH variants showing effectiveness in capturing asymmetric volatility patterns ([Phung Duy et al. 2024; Yıldırım and Bekun 2023](#)).

Feature importance analysis in machine learning applications reveals that technical indicators, particularly Bollinger Bands, exponential moving averages, and market capitalization metrics, contribute most significantly to Bitcoin price prediction accuracy. This finding suggests that traditional technical analysis indicators maintain relevance in algorithmic trading contexts, with moving averages and volatility bands providing the most predictive power in ensemble models ([Visharad et al. 2025](#)).

## 2. Methodology

This research employs a progressive complexity methodology to systematically evaluate cryptocurrency price prediction capabilities across three distinct machine learning

paradigms: ensemble tree-based methods (XGBoost), recurrent neural networks (LSTM), and econometric-deep learning hybrids (GARCH-DL). Utilizing comprehensive Bitcoin market data from December 2013 through May 2025, this study implements extensive feature engineering protocols encompassing technical indicators, market microstructure variables, and regime detection measures. The methodological framework progresses from standalone model implementations through advanced ensemble configurations, with performance validation conducted via time-series cross-validation techniques designed to preserve temporal dependencies and ensure out-of-sample robustness.

Altogether, the research methodology is based on the systematic application of the Knowledge Discovery in Databases (KDD) framework to cryptocurrency price forecasting (Figure 1), progressing through increasingly sophisticated methodological implementations that collectively advance the field from fundamental technical analysis to theoretically grounded hybrid econometric and machine learning systems further disclosed in the Section 3. Claude AI was employed to support manuscript preparation by integrating content, helping to consolidate disparate experimental results into unified summary statements, and enhancing the articulation of complex methodological frameworks.

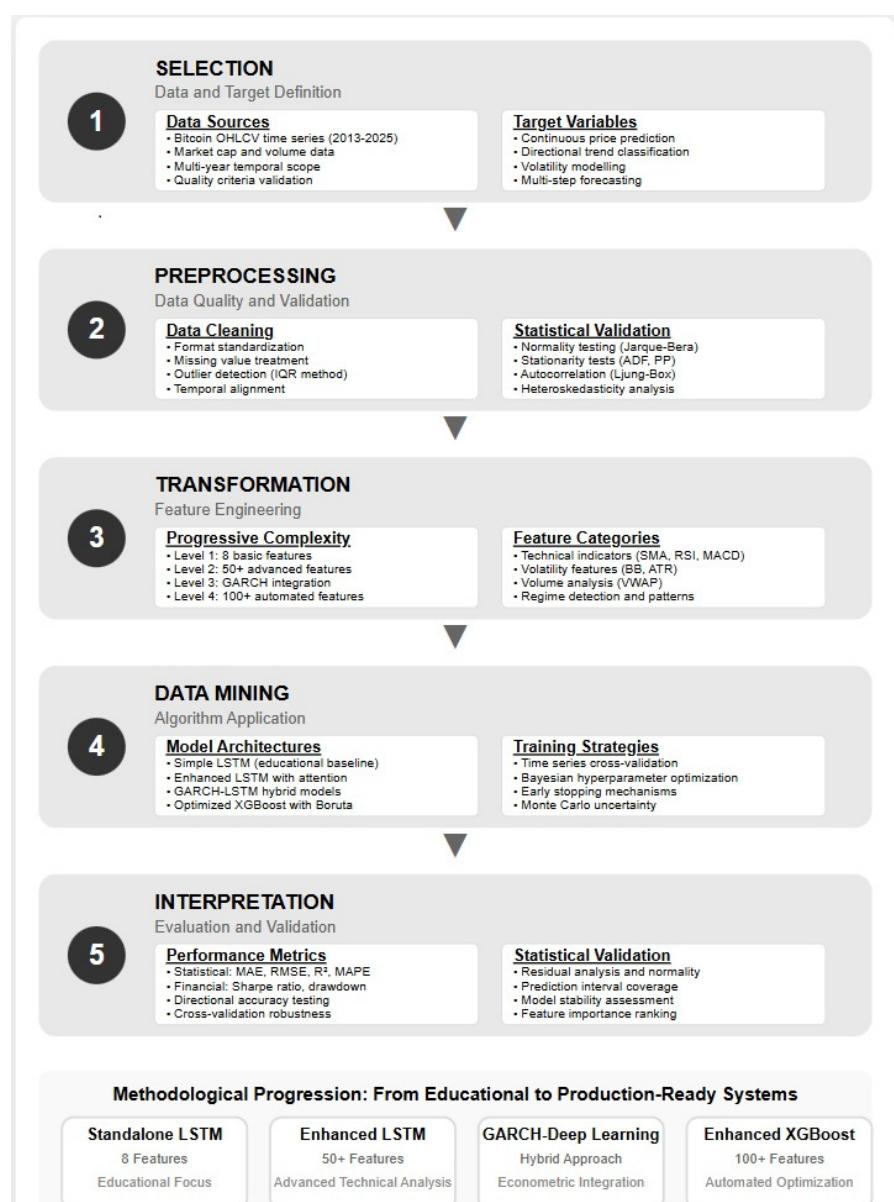


Figure 1. Research methodology based on KDD framework. Source: Own elaboration.

To ensure unbiased evaluation, we implemented strict temporal train–test splits with training data ending 31 December 2023, and testing beginning 1 January 2024. All five models—XGBoost, Simple LSTM, Enhanced LSTM with Attention, Ensemble, and GARCH-DL hybrid—are evaluated exclusively on in-sample and out-of-sample data using identical temporal splits to enable fair comparison and model development.

First, we applied one of the most popular algorithms for time series data prediction (as well as for classification) using machine learning, XGBoost (Chen and Guestrin 2016). XGBoost is a framework based on gradient boosting. The main idea of XGBoost is to continuously add new weak models (regression trees) with different weights. For a dataset of  $n$  examples with  $m$  features,  $D = \{(x_i, y_i)\}$  ( $|D| = n$ ,  $x_i \in R^m$ ,  $y_i \in R$ ), an ensemble model composed of trees uses  $K$  additive functions to predict the output value as follows:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

where  $F = f(x) = w_{q(x)}(Q : R^m \rightarrow T, w \in R^T)$ , the space of regression trees represents the structure of each tree, and  $T$  is the number of leaves in the tree. Each  $f_k$  represents an independent tree with structure  $q$  and leaf weights  $w$ . Each regression tree contains a continuous value in each of its leaves, which we denote with  $w_i$ . For each example, we use the rules from the trees ( $q$ ) to calculate a prediction by summing up the values of the corresponding leaves in which they fall ( $w$ ). To learn these  $K$  functions, we minimize the following objective function:

$$L(\phi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_k \Omega(f_k), \quad (2)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda ||w||^2 \quad (3)$$

$l$  represents a differentiable loss function that measures the difference, and the second term is used for regularization, preventing excessive model complexity and overfitting of the model to the test data. The parameters control the strength of regularization  $\gamma$  and  $\lambda$ . Since direct optimization of the model described above is not possible, the model is trained in an additive manner. Let  $\hat{y}_i^{(t)}$  be the prediction for the  $i$ -th instance at the  $t$ -th iteration, then we strive to optimize the following objective function:

$$L^{(t)} = \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (4)$$

i.e., we try to add this  $f_t$ , which best optimizes the objective function. We use a second-order Taylor expansion to approximate the above objective function as follows:

$$\begin{aligned} L^{(t)} &\simeq \sum_{i=1}^n \left[ l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \\ &= \sum_{i=1}^n \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \\ &= \sum_{j=1}^T \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T, \end{aligned} \quad (5)$$

where  $g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ ,  $h = \partial_{\hat{y}_i^{(t-1)}}^2 l\left(y_i, \hat{y}_i^{(t-1)}\right)$ , a  $I_j = \{i \mid q(x_i) = j\}$  is the set of examples that fall into leaf  $j$ . For a tree structure  $q(x)$ , we can compute the optimal weight  $w_j^*$  of a leaf by

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (6)$$

As well as the optimal value of the objective function:

$$L^{(t)}(q) \simeq -\frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i\right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (7)$$

The optimal value for a configuration  $q$  can be used as a function for evaluating its quality. Since it is not possible to compute all possible structures  $q$ , XGBoost uses the objective function also for determining the quality of potential newly added branches during the iterative construction of trees, by subtracting the value of the original leaf from the sum of the values of the newly constructed leaves as follows:

$$L_{split} = \frac{1}{2} \left[ \frac{\left(\sum_{i \in I_L} g_i\right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i\right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i\right)^2}{\sum_{i \in I} h_i + \lambda} \right] \quad (8)$$

where  $I_L$  and  $I_R$  are the sets of terminal values for the left and right leaves after adding a new branch  $I = I_L \cup I_R$ ?

The easiest way to find the best split is to examine all possible options for each feature. In practice, this approach is not always optimal, especially when the model is trained on large datasets with features that take many different values. For this reason, the creators of XGBoost propose an algorithm that suggests roughly the best split points based on percentiles of a specific feature, and the quality of the resulting subdivisions is measured using aggregated statistics. Besides the regularization term already mentioned, XGBoost uses two more techniques to prevent overfitting. The first is called shrinkage, where the new weights added by the trees at each step are scaled by a factor, like the learning rate used in neural network training. The creators also introduce feature subsampling, which involves training each tree on a subset of all available features.

The enhanced XGBoost model implements a gradient boosting framework that iteratively constructs an ensemble of decision trees to minimize the regularized objective function  $L(\varphi) = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_k)$ , where the loss function employs squared error for regression tasks and the regularization term  $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum w_j^2$  prevents overfitting through leaf count penalization and L2 weight regularization.

The predictive model generates forecasts through an additive ensemble  $\hat{y}_i = \sum f_k(x_i)$  where each tree function  $f_k$  contributes to the final prediction via the gradient boosting update rule  $\hat{y}_i(t) = \hat{y}_i(t-1) + \eta f_k(x_i)$ , with learning rate  $\eta$  controlling the contribution magnitude of successive trees.

Feature engineering incorporates 100+ technical indicators, including RSI calculated as  $RSI = 100 - 100/(1 + RS)$  where  $RS$  represents the ratio of average gains to average losses, MACD computed as  $MACD = EMA12(P) - EMA26(P)$ , and Bollinger Band position  $BBposition = (P - BBlower)/(BBupper - BBlower)$  to capture momentum, trend, and volatility characteristics essential for cryptocurrency price dynamics.

The Boruta feature selection algorithm employs Random Forest importance measures  $I(X_j) = (1/B) \sum \sum p(t) \Delta t(X_j)$  and was applied to identify statistically significant predictors by comparing real feature importance against shadow feature distributions through Z-score computation  $Z_j = (I(X_j) - \max(I_{shadow})) / \sigma(I_{shadow})$  (see full detailed mathematical formulas in Appendix A, Table A1).

Hyperparameter optimization utilizes Tree-structured Parzen Estimator (TPE) sampling within the Optuna framework, modeling the objective function through conditional probability distributions  $p(\theta | y) \propto l(\theta)$  for promising configurations and maximizing the acquisition function  $EI(\theta) = \int (y - \hat{y}) p(y | \theta) dy$  to guide efficient parameter space exploration.

Time series cross-validation maintains temporal integrity through sequential train-validation splits, computing performance metrics including  $R^2$ , RMSE, and directional accuracy (DA) to assess both statistical accuracy and practical trading relevance (see full detailed mathematical formulas in Appendix A, Table A1).

Next, the standalone LSTM model was implemented via deep recurrent neural network architecture designed to capture temporal dependencies in cryptocurrency price sequences through sophisticated memory mechanisms. The core LSTM cell processes sequential input through three gating mechanisms: the forget gate  $f_t$  controlling information retention from previous states, the input gate  $i_t$  regulating new information incorporation, and the output gate  $o_t$  determining information flow to subsequent layers.

The cell state evolution follows  $C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_C[h_{t-1}, x_t] + b_C)$  where element-wise multiplication enables selective memory retention and updating, while the hidden state  $h_t = o_t \cdot \tanh(C_t)$  provides the processed temporal representation for prediction tasks.

Sequential data preparation employs sliding window methodology, constructing input sequences  $X^{(i)} = \{p_i, p_{i+1}, \dots, p_{i+L-1}\}$  with corresponding targets  $y^{(i)} = p_{i+L}$ , where  $L = 30$  represents the lookback window capturing monthly price patterns. Feature vectors incorporate multi-dimensional market information price data (open, high, low, close), volume, and technical indicators like RSI, MACD, and Bollinger Band position, normalized (the vector mathematical representation available in in Appendix A, Table A1) through Min-Max scaling  $x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$ , ensuring stable gradient propagation.

Following which, the two-layer LSTM architecture was added in order to implement  $h_t^{(2)} = \text{LSTM}^{(2)}(h_t^{(1)}, h_{t-1}^{(2)}, C_{t-1}^{(2)})$  with hidden dimensionality of 64 units per layer, followed by batch normalization  $\hat{y} = \gamma(x - \mu_B) / \sqrt{(\sigma_B^2 + \epsilon)} + \beta$  and fully connected layers with ReLU activation for nonlinear transformation. Regularization through dropout  $h_{\text{dropout}} = \frac{h}{1-p}$  with probability  $p = 0.3$  prevents overfitting, while gradient clipping  $g = \min(|g|, \text{max\_norm}) \cdot \frac{\text{sgn}(g)}{|g|}$  with  $\text{max\_norm} = 0.5$  ensures training stability.

Training optimization was added and utilized an AdamW optimizer implementing  $\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{\partial_t + \epsilon}} \cdot \hat{m}_t - \alpha \lambda \theta_t$ , where adaptive moment estimation with weight decay  $\lambda = 0.01$  provides robust parameter updates. The learning rate scheduler applies  $\alpha_{\text{new}} = \alpha_{\text{old}} \cdot \gamma$  with reduction factor  $\gamma = 0.5$  upon validation loss plateaus, while early stopping monitors validation performance with patience = 30 epochs to prevent overfitting. Multi-step ahead forecasting employs iterative prediction  $\hat{y}_{t+k} = f_{\text{LSTM}}(y_{t+k-L:t+k-1})$ , where previous predictions serve as inputs for subsequent forecasts, enabling extended temporal extrapolation essential for practical trading applications.

Performance evaluation encompasses standard regression metrics, including  $R^2$  (mathematically elaborated in in Appendix A, Table A1) for explained variance quantification, RMSE (mathematically elaborated in in Appendix A, Table A1) for prediction accuracy assessment, and MAPE (mathematically elaborated in in Appendix A, Table A1) for percentage error evaluation, providing comprehensive model validation suitable for financial forecasting applications.

The enhanced LSTM architecture incorporates a sophisticated attention mechanism that computes weighted representations of temporal sequences through  $\alpha_t = \text{softmax}(W_a \cdot h_t + b_a)$ , where attention weights dynamically focus on relevant time steps, followed by attended output computation  $c = \sum_t \alpha_t h_t$  providing comprehensive temporal context beyond traditional final hidden state approaches.

The multi-layer architecture implements residual connections  $x_2^{\text{res}} = x_2 + x_1$  when dimensional compatibility allows, enhancing gradient flow and model expressiveness through skip connections analogous to ResNet architectures. The final prediction employs

$\hat{y} = \tanh(W_3 \cdot x_2^{\text{res}} + b_3)$  with hyperbolic tangent activation ensuring prediction stability and bounded outputs suitable for financial forecasting applications.

Advanced technical feature engineering encompasses 30+ indicators including Stochastic Oscillator  $\%K_t = 100 \times \frac{C_t - L_{14}}{H_{14} - L_{14}}$ , Average True Range (ATR)  $\text{TR}_t = \max(H_t - L_t, |H_t - C_{t-1}|, |L_t - C_{t-1}|)$ , and Volume Weighted Average Price  $\text{VWAP}_t = \frac{\sum_{i=t-n+1}^t P_i \cdot V_i}{\sum_{i=t-n+1}^t V_i}$  capturing momentum, volatility, and liquidity dynamics essential for comprehensive market analysis.

Uncertainty quantification employed Monte Carlo dropout methodology where inference maintains activated dropout layers, generating  $S = 100$  stochastic predictions  $\hat{y}^{(s)} = f_{\text{LSTM}}(X; \theta, \text{dropout enabled})$  enabling confidence interval estimation through  $\text{CI}_{95\%} = [\text{percentile}_{2.5}(\hat{y}^{(1:S)}), \text{percentile}_{97.5}(\hat{y}^{(1:S)})]$ , providing probabilistic forecasting capabilities crucial for risk assessment in trading applications.

Multi-step ahead prediction was added to implement iterative forecasting  $\hat{y}_{t+k} = f(y_{t+k-\hat{L}:t+k-1})$ , with dynamic technical indicator updates including RSI approximation  $\text{RSI}_{t+1} = \text{clip}(\text{RSI}_t + 0.1 \times (P_{t+1} - P_t), 0, 100)$  maintaining feature consistency across extended prediction horizons while accounting for indicator dependencies.

Enhanced training protocols were incorporated with learning rate scheduling  $\alpha_{\text{new}} = \alpha_{\text{old}} \times 0.5$  upon validation loss plateaus, gradient clipping  $g_{\text{clipped}} = \min(|g|, 0.5) \cdot \frac{g}{|g|}$  preventing gradient explosion and sophisticated early stopping with patience = 25 epochs ensuring optimal generalization without overfitting.

Financial performance evaluation was extended beyond statistical metrics to include directional accuracy  $\text{DA} = \frac{1}{T-1} \sum_{t=1}^{T-1} I(\text{sign}(\Delta P_t) = \text{sign}(\Delta \hat{y}_t))$ , Sharpe ratio of predictions  $\text{Sharpe}_{\text{pred}} = \frac{E[r_{\text{pred}}]}{\sigma(r_{\text{pred}})} \times \sqrt{252}$  and maximum drawdown  $\text{MaxDD} = \min_t \left( \frac{\text{CumReturn}_t - \text{RunningMax}_t}{\text{RunningMax}_t} \right)$ , were implemented for providing comprehensive assessment of trading viability and risk characteristics.

Time series cross-validation maintains temporal integrity through  $\text{Train}^{(i)} = \{1, 2, \dots, n_i\}$ ,  $\text{Val}^{(i)} = \{n_i + 1, \dots, n_{i+1}\}$  ensuring realistic performance estimation while preventing data leakage, with aggregated metrics  $\text{CV}_{R^2} = \frac{1}{k} \sum_{i=1}^k R_i^2$  was added providing robust model validation essential for financial forecasting applications.

The GARCH-Deep Learning hybrid framework integrates traditional econometric volatility modeling with modern neural network architectures through a sophisticated dual-branch design. The foundation employs GARCH(1,1) volatility specification  $\sigma_t^2 = \Omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$  where conditional volatility  $\sigma_t$  captures time-varying heteroskedasticity inherent in cryptocurrency markets, with distributional flexibility accommodating normal, Student-t, or skewed-t error distributions  $\varepsilon_t = \sigma_t z_t$  ensuring robust volatility parameter estimation across varying market conditions.

The neural architecture implements (see Appendix A, Table A1 for detailed mathematical elaboration) a dual-branch LSTM design where GARCH-derived volatility features enhance temporal sequence modeling through  $h^{\text{attended}} = \text{MultiHead}(h^{\text{main}}, h^{\text{main}}, h^{\text{main}})$  utilizing multi-head attention mechanisms. The volatility branch  $f^{\text{vol}} = W_{\text{vol}2} \cdot \text{Dropout}(\text{ReLU}(W_{\text{vol}1} \cdot h^{\text{attended}}))$  specializes in volatility-related pattern extraction, while the price branch  $f^{\text{price}} = W_{\text{price}2} \cdot \text{Dropout}(\text{ReLU}(W_{\text{price}1} \cdot h^{\text{attended}}))$  focuses on price dynamics, culminating in combined prediction  $\hat{y}_t = W_{\text{final}} \cdot \text{ReLU}(W_{\text{fusion}} \cdot [f^{\text{vol}}; f^{\text{price}}])$ .

Enhanced feature engineering incorporated GARCH-based volatility percentiles  $\text{Vol_Percentile}_t = \frac{\text{Rank}(\sigma_t \text{ in } \{\sigma_{t-251}, \dots, \sigma_t\})}{252}$ , regime detection  $\text{Vol_Regime}_t = I(\sigma_t > \bar{\sigma}_{30})$ , and risk metrics including Value-at-Risk  $\text{VaR}_{95\%, t} = \sigma_t \times \Phi^{-1}(0.05)$ , where  $\Phi^{-1}$  represents the inverse normal cumulative distribution function, providing comprehensive volatility-based features essential for cryptocurrency risk modeling.

Multi-step ahead forecasting employs iterative GARCH volatility updates  $\sigma_{t+h}^2 | t = \Omega \sum_{i=0}^{h-2} (\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)_{h-1}^{-1} \sigma_{t+1}^2 | t$ , enabling dynamic volatility forecasting, while prediction uncertainty quantification utilizes GARCH-based confidence intervals  $\text{VaR}_{95\%|t} = \sigma_t \times \Phi^{-1}(0.05)$  providing probabilistic forecasting capabilities crucial for trading risk assessment.

The hybrid training optimization employs joint loss minimization  $L_{\text{total}} = L_{\text{price}} + \lambda_{\text{vol}} L_{\text{volatility}} + \lambda_{\text{reg}} L_{\text{regularization}}$ , where price prediction loss addresses forecasting accuracy, volatility consistency loss  $L_{\text{volatility}} = \frac{1}{T} \sum_{t=1}^T (\sigma_t^{\text{GARCH}} - \hat{\sigma}_t^{\text{DL}})^2$  ensures coherence between econometric and neural volatility estimates, and regularization prevents overfitting in the high-dimensional parameter space.

Model validation encompasses traditional forecasting metrics complemented by volatility-specific evaluation including realized volatility comparison  $\text{MSE}_{\text{vol}} = \frac{1}{T} \sum_{t=1}^T (\sigma_t^{\text{realized}} - \sigma_t^{\text{forecast}})^2$  and VaR backtesting through Kupiec likelihood ratio tests  $\text{LR} = -2 \ln \left[ \frac{\alpha^N (1-\alpha)^{T-N}}{\left( \frac{N}{T} \right)^N (1-\frac{N}{T})^{T-N}} \right]$ , ensuring robust risk model validation essential for cryptocurrency market applications.

The architecture supports multiple neural configurations including pure LSTM, GRU, and hybrid LSTM-GRU implementations  $h^{\text{hybrid}} = [h^{\text{LSTM}}; h^{\text{GRU}}]$  with fusion through multi-layer perceptrons, enabling comprehensive evaluation of deep learning architectures while maintaining GARCH-based volatility enhancement across all configurations.

The methodological framework adopts distinct variable configurations across model architectures to capture different aspects of cryptocurrency price dynamics as shown in Table 1. The dependent variable consistently represents future Bitcoin closing prices  $y_{\{t+1\}}$ , enabling direct comparison of predictive performance across methodologies.

**Table 1.** Comprehensive Variable Summary.

Model Type	Dependent Variable(s)	Independent Variables	Input Shape
XGBoost	close_price_{t+1}	Current features only	(n_samples, ~100 features)
LSTM	close_price_{t+1}	30-day sequence, 8 features/day	(n_samples, 30, 8)
Enhanced LSTM	close_price_{t+1}	30-day sequence, 30 features/day	(n_samples, 30, 30)
GARCH-DL	close_price_{t+1}, volatility_{t+1}	30-day sequence, 35 features/day	(n_samples, 30, 35)

Source: Own estimations.

Independent variable structures vary systematically by model complexity and temporal scope. The XGBoost implementation utilizes current-state features  $X_t = [\text{price}_t, \text{volume}_t, \text{technical\_indicators}_t, \text{regime}_t]$  comprising approximately 100 cross-sectional variables including OHLCV data, technical indicators (RSI, MACD, Bollinger Bands), lagged price features, and market regime indicators.

As can be observed in Table 1, the LSTM architectures employ temporal sequences  $X_{\{t-L+1:t\}}$ , where  $L = 30$  represents the lookback window, with each time step containing 8–30 features depending on model complexity. The standalone LSTM incorporates basic price–volume–technical indicator combinations, while the enhanced LSTM expands to 30+ features per time step including advanced technical indicators, candlestick patterns, and rolling statistical measures.

What is more, the GARCH–Deep Learning hybrid implements a dual-target framework predicting both  $price_{\{t+1\}}$  and  $volatility_{\{t+1\}}$ , utilizing enhanced feature sequences that integrate GARCH-derived volatility measures, regime detection indicators, and risk metrics alongside standard price–volume features.

Variable construction maintains strict temporal integrity through progressive feature calculation, ensuring no future information leakage. Cross-sectional models (XGBoost) leverage instantaneous relationships between current market conditions and future prices, while sequential models (LSTM variants) exploit temporal dependencies through recurrent processing of historical patterns. The hybrid approach combines econometric volatility modeling with neural temporal pattern recognition, creating a comprehensive feature space spanning both cross-sectional and time-series dimensions. This variable framework enables systematic comparison of different information processing paradigms while maintaining consistent prediction targets across all model architectures.

The comparative evaluation framework is motivated by ensemble learning theory demonstrating that stacking models can outperform individual predictors by leveraging complementary algorithmic strengths (Ji et al. 2019; Ye et al. 2022). This theoretical foundation supports the multi-model comparison approach rather than relying on single algorithmic implementations. The integration of GARCH specifications within the hybrid framework addresses the documented effectiveness of various GARCH-type models in capturing asymmetric volatility patterns, with EGARCH variants showing particular promise for cryptocurrency applications (Phung Duy et al. 2024; Yıldırım and Bekun 2023). The comprehensive validation approach addresses documented challenges related to model stability, overfitting, and inherent cryptocurrency market unpredictability, implementing robust frameworks essential for practical implementations (Mudassir et al. 2020; Park and Seo 2022).

The methodological progression employed in this study reflects a structured evolution of forecasting models, aligned with increasing feature complexity and domain-specific sophistication. The initial phase utilizes a standalone Long Short-Term Memory (LSTM) architecture, leveraging a minimal feature set ( $n = 8$ ) to establish a foundational benchmark within an educational context. This is followed by an enhanced LSTM model incorporating over 50 engineered features, integrating advanced technical indicators to improve temporal pattern recognition. The third stage introduces a hybrid GARCH–Deep learning framework, combining econometric volatility modeling with deep neural architectures to capture regime-dependent dynamics. Finally, the progression culminates in an optimized Extreme Gradient Boosting (XGBoost) model, utilizing over 100 automated features and Bayesian hyperparameter tuning to achieve robust, high-performance forecasting under complex market conditions.

The methodological approach (Figure 1) is grounded in established theoretical foundations from cryptocurrency prediction literature. The progressive complexity framework employed addresses the documented inconsistency in empirical findings, where deep learning models demonstrate superior performance in capturing non-linear and non-stationary cryptocurrency dynamics compared to traditional time series models like ARIMA (Bouteska et al. 2024; Chen et al. 2024), while ensemble tree-based methods such as XGBoost consistently outperform neural networks for tabular financial data prediction tasks (Hafid et al. 2024; Şimşek Türker et al. 2025).

### Phase 1: Selection and Data Definition

The multi-year temporal scope (2013–2025) is theoretically justified by the need to capture various market regimes and volatility patterns inherent in cryptocurrency markets. The inclusion of on-chain blockchain data alongside traditional OHLCV metrics is motivated by

empirical evidence that transaction volumes, network hash rates, and market capitalization provide valuable predictive signals (Kjærland et al. 2018; Visharad et al. 2025).

### Phase 2: Preprocessing and Data Quality Validation

Data preprocessing involves systematic cleaning procedures including format standardization across all data sources, comprehensive missing value treatment protocols, outlier detection using the Interquartile Range (IQR) method, and temporal alignment to ensure consistency across different data streams.

Statistical validation encompasses rigorous testing protocols including normality assessment via the Jarque-Bera test, stationarity evaluation through Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests, autocorrelation analysis using the Ljung-Box test, and heteroskedasticity examination to identify variance instability patterns.

### Phase 3: Transformation and Feature Engineering

The progressive complexity framework is theoretically grounded in feature selection literature demonstrating that methods such as Boruta (BB) and genetic algorithms significantly enhance prediction accuracy by addressing the curse of dimensionality (Visharad et al. 2025). The integration of technical indicators is justified by feature importance analysis revealing that Bollinger Bands (BB), exponential moving averages, and market capitalization metrics contribute most significantly to Bitcoin price prediction accuracy, suggesting that traditional technical analysis indicators maintain relevance in algorithmic trading contexts.

### Phase 4: Data Mining and Algorithm Application

The multi-architectural approach addresses conflicting empirical evidence regarding model superiority. The inclusion of LSTM networks is motivated by research demonstrating their effectiveness in capturing temporal patterns and long-term dependencies (Kervancı et al. 2024; Seabe et al. 2023), with Bi-LSTM variants showing particular promise (Alizadehgan et al. 2024). The GARCH-LSTM hybrid implementation is theoretically grounded in the established effectiveness of GARCH models as a cornerstone methodology for capturing volatility dynamics in cryptocurrency data (Chu et al. 2017).

The XGBoost implementation addresses empirical findings showing consistent out-performance of ensemble tree-based methods over neural networks for tabular financial prediction tasks. The integration of attention mechanisms reflects recent theoretical developments demonstrating improved performance through transformer-based architectures (Lee 2024). The enhanced XGBoost model with over 100 automated features and Boruta-based optimization, representing the state-of-the-art in ensemble learning approaches for tabular financial data prediction.

### Phase 5: Interpretation and Model Validation

The comprehensive evaluation framework addresses critical theoretical concerns regarding model reliability and practical applicability in cryptocurrency prediction. The multi-metric validation approach is theoretically motivated by documented challenges related to model stability, overfitting, and the inherent unpredictability of cryptocurrency markets, necessitating robust validation frameworks for practical implementations (Mudassir et al. 2020; Park and Seo 2022).

We consider that such progressive framework enables systematic evaluation of algorithmic evolution from foundational deep learning approaches through sophisticated hybrid methodologies, culminating in automated optimization techniques suitable for practical deployment in cryptocurrency trading applications.

### 3. Materials and Methods

This study utilized historical Bitcoin price data from December 2013 to May 2025, which was pre-processed to handle missing values and outliers using IQR-based clipping. Technical indicators—including SMA, RSI, MACD, Bollinger Band position, and volatility—were engineered to enhance the feature set for forecasting. This study employs a multi-model, hybrid forecasting architecture for cryptocurrency price prediction, incorporating advanced econometric modeling, ensemble machine learning, and deep learning frameworks based on contemporary state-of-the-art pipeline code structures. The framework is executed via Jupyter notebooks in a Linux Docker environment. The methodological progression is structured across four stages: (1) gradient boosting via XGBoost, (2) univariate LSTM modeling, (3) enhanced multivariate attention-based LSTM, and (4) a GARCH–Deep Learning hybrid model, based on LSTM and GRU deep neural networks, that jointly could capture volatility dynamics and nonlinear temporal dependencies. Data sequences of 30-day windows were constructed and normalized for training custom simple and enhanced LSTM neural networks implemented in PyTorch (**Stable release 2.8.0**). Hyperparameters such as batch size, learning rate, and number of epochs were optimized manually, and early stopping was applied to prevent overfitting. The XGBoost model was also trained in parallel for ensemble comparison, as was the case with the attention-based LSTM and the GARCH-DL. The models' performance was evaluated using MAE, RMSE,  $R^2$ , MAPE, directional accuracy, and Sharpe Ratio.

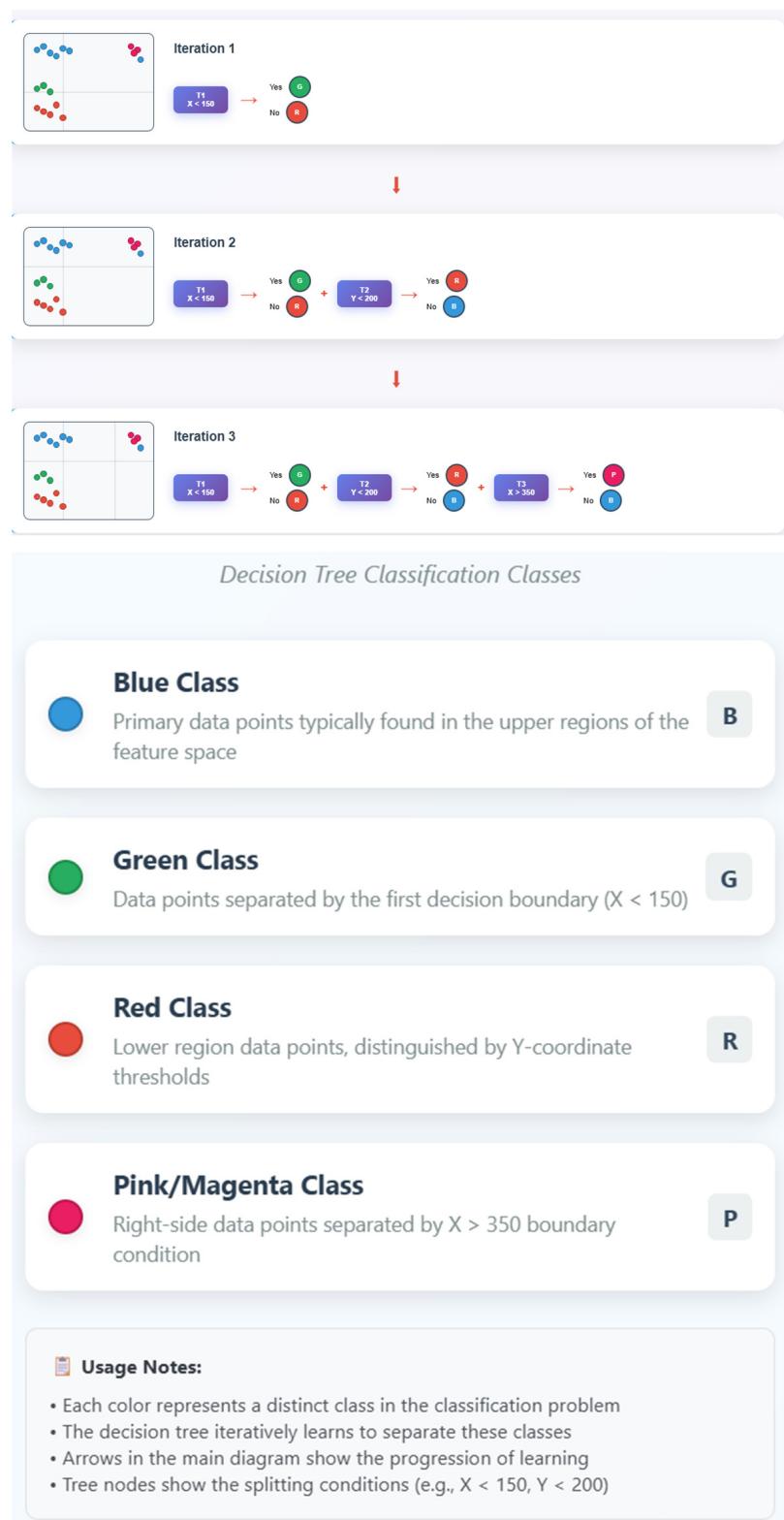
The multi-architectural approach is theoretically motivated by evidence that hybrid models combining different algorithmic strengths yield superior performance. Wang et al. (2024) demonstrated that hybrid LSTM-GRU models with CEEMDAN decomposition outperform benchmark models, while Simsek Turker et al. (2025) showed that CNN-BiLSTM hybrids achieve exceptional forecasting accuracy, second only to XGBoost implementations. The evolution from simple 8-feature LSTM models to comprehensive 100+ feature XGBoost implementations with automated Boruta selection exemplifies the transformation phase's critical role in extracting meaningful patterns from raw financial time series data, while the integration of GARCH volatility modeling with attention-enhanced neural networks represents a novel contribution to the data mining phase that bridges traditional econometric theory with modern deep learning paradigms. The preprocessing phase incorporates rigorous statistical validation protocols, including normality, stationarity, and autocorrelation testing, to ensure data quality and model validity. This addresses fundamental assumptions often overlooked in conventional machine learning applications for financial markets. The interpretation phase extends beyond conventional performance metrics to encompass comprehensive statistical validation through residual analysis, prediction interval coverage, and directional accuracy testing, providing formal hypothesis testing frameworks that establish statistical significance of discovered patterns. Furthermore, the implementation of time-series cross-validation with Bayesian hyperparameter optimization via Optuna represents a methodologically rigorous approach to model selection that prevents temporal data leakage while ensuring optimal predictive performance.

Bitcoin price prediction analysis was conducted using historical OHLC (Open, High, Low, Close) data spanning December 2013 to May 2025, comprising of dataset with 4190 daily observations. The dataset included the following essential features: timeOpen, timeClose, timeHigh, timeLow, open, high, low, close prices, trading volume, market capitalization, and timestamps. Data preprocessing involved handling missing values, price normalization, and temporal feature engineering to create technical indicators suitable for machine learning algorithms.

Technical indicators were derived from the raw OHLC data, including exponential moving averages (ema\_12), simple moving averages (sma\_14), Bollinger Bands (bb\_upper), price channels (lower\_channel, upper\_channel), and lagged price features (close\_lag\_2, high, low). Market capitalization and volume metrics were incorporated as fundamental indicators. The feature importance analysis revealed market capitalization and low price as the most predictive variables for Bitcoin price movements. The feature engineering methodologies employed across the cryptocurrency forecasting frameworks demonstrate a progressive evolution from basic technical analysis to sophisticated econometric machine learning integration, reflecting the advancement of computational finance research over the past decade. The minimalist approach utilizing eight fundamental indicators (moving averages, RSI, MACD) establishes a baseline that prioritizes computational efficiency and interpretability, while the comprehensive framework expands to over 50 engineered features encompassing multi-timeframe analysis, candlestick pattern recognition, and advanced volatility measures to capture complex market microstructure effects. The hybrid GARCH-deep learning model represents the methodological frontier by incorporating conditional volatility estimates, regime detection mechanisms, and theoretically grounded risk metrics such as Value-at-Risk calculations, thereby bridging traditional econometric modeling with modern machine learning paradigms. This hierarchical feature engineering progression reveals a fundamental trade-off between model interpretability and predictive sophistication, where increased feature complexity enables more nuanced market pattern recognition at the cost of computational overhead and potential overfitting risks. The integration of econometric volatility modeling with neural network architectures in the hybrid approach constitutes a significant methodological contribution that addresses the heteroskedasticity and volatility clustering inherent in cryptocurrency markets, establishing a new paradigm for theoretically informed machine learning applications in financial forecasting.

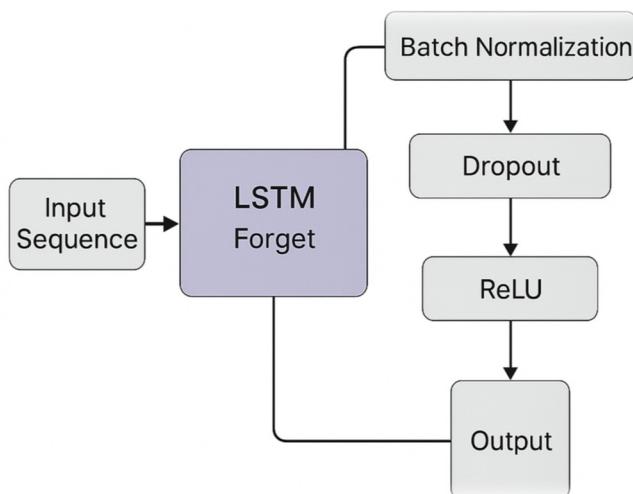
One machine learning approach was implemented and compared, namely XGBoost (Extreme Gradient Boosting) on Figure 2, a tree-based ensemble method optimized for tabular data, specifically designed for sequential time series prediction.

The XGBoost-based framework (Figure 2) for cryptocurrency price prediction emphasizes statistical validation, automated optimization, and production-ready implementation over traditional ad hoc modelling approaches. The structure incorporates comprehensive feature engineering encompassing over 100 technical indicators across multiple timeframes, coupled with Boruta-based statistical feature selection to ensure optimal variable selection without manual intervention or domain expertise bias. A significant methodological contribution lies in the integration of Bayesian hyperparameter optimization using Optuna, which systematically explores the parameter space while employing time-series cross-validation to prevent data leakage and ensure temporal validity of model performance estimates. The framework implements extensive statistical testing protocols for both raw data characteristics and model residuals, including normality assessments, stationarity tests, autocorrelation detection, and heteroskedasticity analysis, thereby providing formal statistical validation that is often absent in machine learning applications. The multi-version compatibility architecture addresses practical deployment challenges by implementing fallback mechanisms for different XGBoost versions, ensuring robustness across diverse computational environments while maintaining optimal performance through early stopping and tracking of evaluation metrics.



**Figure 2.** The applied XGBoost model structure. Source: Own elaboration.

The second model consists of a deep machine architecture with a two-layer straightforward LSTM (Figure 3 and Table 2) configuration with minimal technical feature engineering, utilizing only eight fundamental indicators including moving averages, RSI, and MACD to capture essential price dynamics without the complexity of advanced econometric integration with batch normalization and ReLU activation.



**Figure 3.** The applied Simple stand-alone LSTM model architecture. Source: Own elaboration.

**Table 2.** Simple stand-alone LSTM gate's function and purpose.

Gate	Function	Purpose
Forget	$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$	Remove irrelevant past information
Input	$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$	Select new information to store
Output	$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$	Control information flow to output
Hidden State	$h_t = o_t * \tanh(C_t)$	

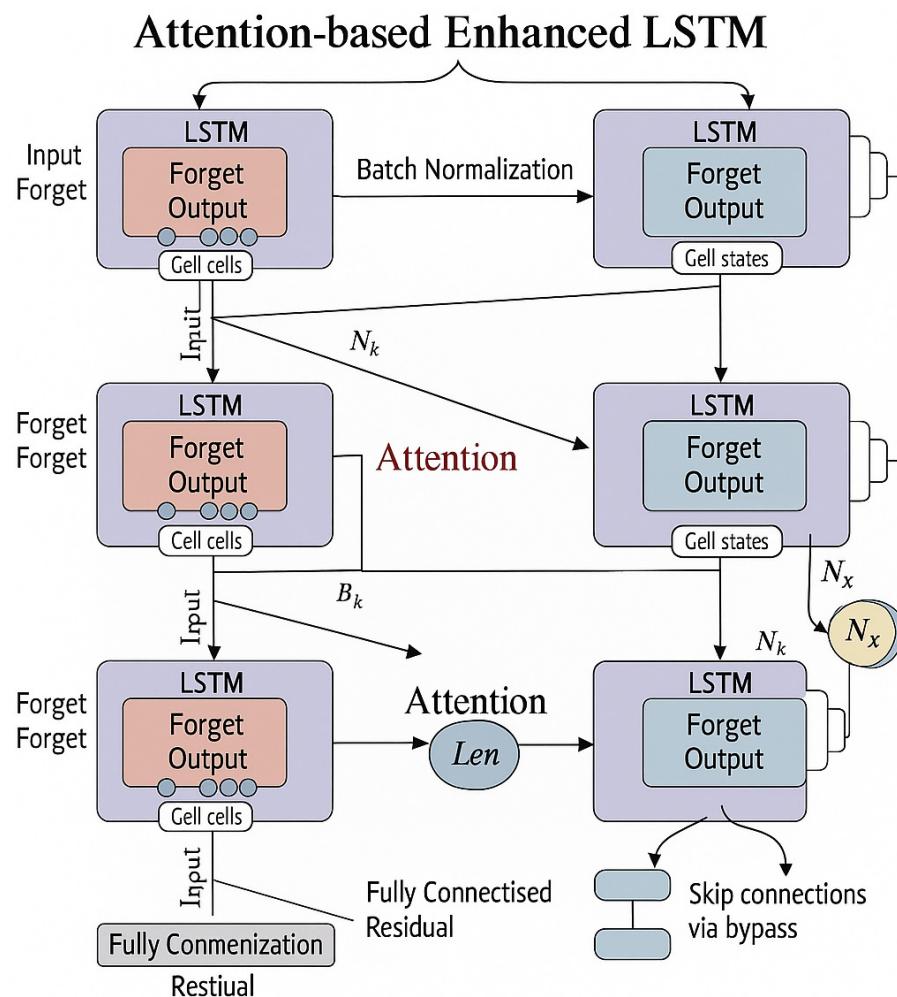
Source: Own estimations.

Where

- $\sigma$  is the sigmoid function:  $\sigma(x) = \frac{1}{1+e^{-x}}$
- $W_f, W_i, W_C, W_o$  are weight matrices.
- $b_f, b_i, b_C, b_o$  are bias vectors.
- $*$  denotes element-wise multiplication.

This architecture deliberately omits sophisticated components such as attention mechanisms, volatility modelling, or uncertainty quantification in favor of transparency and educational accessibility, making it suitable for introductory applications and baseline comparisons. While this simplified approach sacrifices the theoretical rigor and predictive sophistication of hybrid econometric-machine learning frameworks, it provides computational efficiency and clear interpretability that facilitate understanding of core LSTM principles in financial forecasting applications. The evaluation framework focuses exclusively on standard regression metrics, eschewing finance-specific performance measures in favour of straightforward model assessment that emphasizes fundamental predictive accuracy over trading-oriented optimization. This streamlined implementation serves as an important methodological foundation that demonstrates the essential components of deep learning-based cryptocurrency forecasting while maintaining the simplicity necessary for educational purposes and rapid prototyping applications.

Due to the insufficient XGD Boost and LSTM models' performance (see Table 1) the enhanced attention-based LSTM model (Figure 4) was elaborated and executed via a pipeline code in the Linux Docker environment.

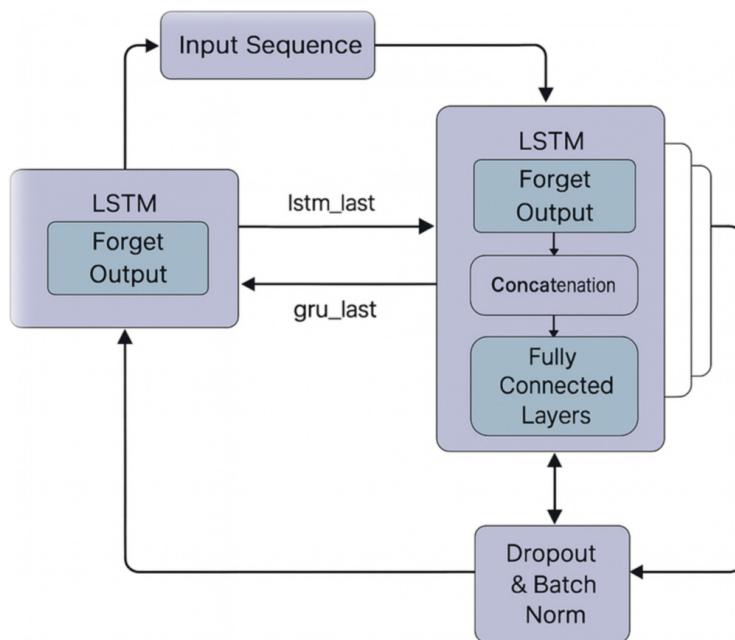


**Figure 4.** The applied attention-based LSTM model structure. Source: Own estimation.

The proposed enhanced LSTM architecture incorporates attention mechanisms, with detailed explanations of all the colors and arrows available in Appendix B, residual connections, and advanced regularization techniques to capture complex temporal dependencies in financial time series data while mitigating overfitting and vanishing gradient problems. A key methodological contribution lies in the extensive feature engineering pipeline, which transforms raw OHLCV data into over 50 technical indicators spanning momentum, volatility, trend analysis, and market microstructure variables, thereby providing the model with rich contextual information for pattern recognition. The evaluation framework extends beyond conventional machine learning metrics to include finance-specific measures, such as directional accuracy, Sharpe ratio, and maximum drawdown, offering a more comprehensive assessment of practical trading performance. Furthermore, the implementation of Monte Carlo simulation for uncertainty quantification provides probabilistic forecasts with confidence intervals, addressing the inherent stochasticity in financial markets and enabling more informed risk management decisions. The time-aware cross-validation methodology ensures robust model validation while preventing data leakage, a critical consideration often overlooked in financial time series applications. These methodological advances collectively contribute to a more reliable and practically applicable framework for cryptocurrency price forecasting that bridges the gap between academic research and real-world financial applications.

What is more, considering the abovementioned advanced volatility dynamics that GARCH models could provide especially regarding fat-tailed data distributions, which is

the present case, a hybrid GARCH–Deep Learning model (Figure 5) was developed and executed once again via pipeline in the Linux Docker environment.



**Figure 5.** The applied Hybrid GARCH–Deep Learning model structure (GARCH-DL). Source: Own elaboration.

The proposed methodology employs a dual-branch neural network architecture that processes price dynamics and volatility signals through specialized pathways, enabling explicit modeling of the distinct statistical properties inherent in financial time series while leveraging GARCH-derived conditional volatility as informed features. A significant methodological contribution lies in the integration of theoretically grounded uncertainty quantification, where GARCH volatility forecasts are utilized to construct probabilistic confidence intervals around deep learning predictions, bridging the gap between econometric rigor and machine learning flexibility. The framework incorporates sophisticated market regime detection mechanisms and risk metrics including Value-at-Risk calculations, enhancing its practical applicability for portfolio optimization and risk management in volatile cryptocurrency markets. The comprehensive evaluation methodology encompasses both econometric diagnostics for GARCH model validation and finance-specific performance metrics, ensuring statistical robustness while maintaining practical relevance for quantitative trading applications. Furthermore, the ability to generate multi-step forecasts with volatility-informed uncertainty bands represents a substantial advancement over traditional approaches that treat volatility as exogenous, providing practitioners with more reliable risk assessment tools. These methodological innovations collectively establish a theoretically principled yet practically viable framework for cryptocurrency forecasting that leverages the complementary strengths of econometric volatility modelling and modern deep learning techniques.

All models were trained using identical feature sets with early stopping mechanisms to prevent overfitting. The XGBoost model employed Bayesian hyperparameter optimization via Optuna across nine key parameters with 5-fold Time Series Split cross-validation, incorporating automated Boruta feature selection and Robust tScaler normalization to ensure temporal integrity and optimal performance. The standalone LSTM baseline utilized a simple 2-layer architecture with 64 hidden units, trained over 200 epochs using Adam optimization with early stopping (patience = 30) and gradient clipping, processing 30-day

sequences of 8 basic technical indicators scaled via a Min Max Scaler. The enhanced attention-based LSTM incorporated multi-head attention mechanisms with dual-branch processing for price and volatility signals, trained on 50+ technical features using residual connections, batch normalization, and learning rate scheduling over 150–200 epochs with advanced regularization techniques. The GARCH–Deep Learning hybrid model employed a two-stage training process: first fitting GARCH models via maximum likelihood estimation with diagnostic validation, then training enhanced LSTM networks on GARCH-derived conditional volatility features combined with technical indicators using Monte Carlo simulation for uncertainty quantification. All models maintained strict temporal ordering through time-aware data splitting for 80/10/10 train/val/test ratio for GARCH–DL Hybrid and 80/20 train/test ratios for the other 3 models; implemented early stopping mechanisms to prevent overfitting; and employed comprehensive statistical validation including residual analysis and prediction interval coverage testing to ensure model reliability and generalizability.

Model performance was assessed using comprehensive regression and classification metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination ( $R^2$ ), Mean Absolute Percentage Error (MAPE), and directional accuracy for Bitcoin price movement prediction. XGBoost achieved superior directional accuracy (81.1%) compared to LSTM (48.8%), while LSTM demonstrated higher correlation with actual prices ( $R^2 = 0.7631$  vs. 0.6600). Statistical significance testing confirmed XGBoost's superiority in predicting Bitcoin price direction, with comprehensive residual analysis validating model assumptions and identifying systematic prediction patterns.

The systematic progression from educational baseline models to production-ready systems with automated feature engineering and statistical validation demonstrates the KDD framework's capacity to transform raw cryptocurrency data into actionable financial intelligence with formal statistical grounding. These methodological innovations collectively establish a comprehensive template for applying data science principles to financial forecasting that maintains both theoretical rigor and practical applicability in volatile cryptocurrency markets.

## 4. Results

Preliminary diagnostic tests were conducted on the raw data, specifically bitcoin daily price data (OHLCV—Open, High, Low, Close, Volume) spanning from December 2013 to May 2025 ( $n = 4190$  observations), to verify key statistical assumptions and identify suitable modeling approaches for subsequent financial time series analysis. Table 3 presents the summary descriptive statistics of the observed variables.

**Table 3.** Summary Statistics Table.

Statistic	Open	High	Low	Close	Volume	Market Cap
<b>Count</b>	4190	4190	4190	4190	4190	4190
<b>Mean</b>	\$20,930.10	\$21,388.24	\$20,451.05	\$20,954.05	$1.82 \times 10^{10}$	$\$4.01 \times 10^{11}$
<b>Std Dev</b>	\$25,680.90	\$26,208.72	\$25,132.62	\$25,711.03	$2.08 \times 10^{10}$	$\$5.06 \times 10^{11}$
<b>Min</b>	\$176.90	\$211.73	\$171.51	\$178.10	0.00	$\$2.44 \times 10^9$
<b>25%</b>	\$768.85	\$775.79	\$756.09	\$768.53	$1.07 \times 10^8$	$\$1.12 \times 10^{10}$
<b>50%</b>	\$9064.88	\$9235.25	\$8820.94	\$9076.35	$1.36 \times 10^{10}$	$\$1.61 \times 10^{11}$
<b>75%</b>	\$33,344.10	\$34,392.15	\$31,982.40	\$33,470.10	$2.95 \times 10^{10}$	$\$6.26 \times 10^{11}$
<b>Max</b>	\$111,679.36	\$111,970.17	\$109,285.07	\$111,673.28	$3.51 \times 10^{11}$	$\$2.22 \times 10^{12}$

Source: Own estimations.

This dataset captures Bitcoin's remarkable price journey across 4190 daily observations, spanning from its early adoption phase to peak market cycles. The mean Bitcoin price of approximately \$20,930 reflects the dataset's coverage of multiple market cycles, including both bear and bull market periods. The extraordinarily high standard deviation of over \$25,680 demonstrates Bitcoin's notorious volatility, with the coefficient of variation exceeding 120%, making it one of the most volatile significant assets in financial markets.

The price range, from a minimum of \$176.90 to a maximum of \$111,970.17, represents Bitcoin's incredible growth trajectory over time. The minimum values around \$170–\$180 likely capture Bitcoin's early years (2011–2012) or the depths of major bear markets, while the maximum price of nearly \$112,000 represents Bitcoin's historic all-time highs during bull market peaks. This represents a growth multiple of over 630 times from the lowest to the highest recorded prices, illustrating Bitcoin's transformation from an experimental digital currency to a significant financial asset.

The median price of \$9076 is significantly lower than the mean, indicating that Bitcoin has spent more time at lower price levels than higher ones, which is characteristic of exponential growth assets. The distribution shows that 50% of the observations fall below \$9076. At the same time, the upper quartile begins at \$33,344, indicating that Bitcoin's significant price appreciation periods tend to be relatively shorter in duration compared to consolidation or correction phases.

Bitcoin's trading volume data reveals the asset's evolution into a major traded commodity, with mean daily volume exceeding 18.2 billion units. The enormous standard deviation of 20.8 billion, which surpasses the mean, reflects the episodic nature of Bitcoin trading intensity. During major market events, such as institutional adoption announcements, regulatory developments, or macroeconomic shifts, Bitcoin volume can surge dramatically, creating these extreme observations.

The median volume of 13.6 billion is lower than the mean, confirming that most trading days experience moderate activity, while exceptional volume spikes during major market events drive the average higher. The maximum recorded volume of 351 billion units likely corresponds to historic market events such as the 2017 bubble peak, the 2020–2021 institutional adoption wave, or significant correction periods when panic selling or buying created extraordinary trading activity.

The presence of zero minimum volume observations could indicate early Bitcoin trading periods when exchanges had limited liquidity, technical issues with data collection, or specific days when certain exchanges experienced trading halts. The 25th percentile at 107 million suggests that during Bitcoin's earlier years or quieter market periods, daily volumes were substantially lower than current levels, reflecting the asset's growth in mainstream adoption and institutional participation.

Bitcoin's market capitalization statistics showcase its evolution from an experimental project to one of the world's most valuable assets. The mean market cap of \$401 billion positions Bitcoin among the most significant assets globally, comparable to major corporations and commodities. The massive standard deviation of \$506 billion reflects Bitcoin's journey through multiple boom–bust cycles, where market capitalization has experienced dramatic expansions and contractions.

The maximum market capitalization of \$2.22 trillion represents Bitcoin's peak valuation periods, likely corresponding to the 2021 bull market when Bitcoin briefly challenged the market capitalizations of the world's largest companies. This peak market cap reflects not only price appreciation but also the growing supply of Bitcoin in circulation as new coins are mined over time. The minimum market cap of \$2.44 billion corresponds to Bitcoin's early years, when both the price and circulating supply were significantly lower.

The median market cap of \$161 billion being significantly below the mean demonstrates that Bitcoin has spent most of its existence at lower valuations, with the trillion-dollar market cap periods representing relatively brief but significant episodes in its price history. The interquartile range from \$11.2 billion to \$626 billion captures Bitcoin's transition from a niche digital asset to a major store of value, with the 75th percentile representing periods when Bitcoin achieved mainstream financial recognition.

The dataset's comprehensive coverage of 4190 observations likely spans over a decade of Bitcoin's price history, capturing multiple distinct market cycles. The extreme price ranges suggest inclusion of Bitcoin's early trading periods (2011–2013), the first major bubble and crash (2017–2018), the institutional adoption phase (2020–2021), and subsequent market corrections. This temporal breadth provides insight into Bitcoin's maturation as an asset class.

The consistently high volatility across all price metrics (open, high, low, close) confirms Bitcoin's characteristic price behaviour throughout its history. Unlike traditional assets that may experience volatility spikes during crisis periods, Bitcoin maintains high volatility as a baseline characteristic, driven by factors such as regulatory uncertainty, technological developments, institutional adoption waves, and shifts in macroeconomic sentiment.

The substantial differences between quartile ranges suggest that Bitcoin's price discovery process has been highly non-linear, characterized by extended periods of consolidation punctuated by rapid appreciation or depreciation phases. This pattern reflects Bitcoin's unique position as an emerging asset class that is still finding its equilibrium value, influenced by adoption cycles, technological improvements, and evolving regulatory landscapes across global markets.

The dataset demonstrates exceptional completeness with exactly 4190 observations for each metric and no missing values, indicating robust data collection processes throughout Bitcoin's trading history. This consistency suggests the data encompasses daily observations from major exchanges or aggregated market data, providing a comprehensive view of Bitcoin's price evolution without gaps that might distort statistical analysis.

The uniform observation count across all variables confirms that each data point represents a complete daily trading session with corresponding price, volume, and market capitalization data. This completeness is particularly valuable for Bitcoin analysis, as the cryptocurrency market operates continuously, and missing data could significantly impact trend analysis, given Bitcoin's high volatility and the potential for substantial price fluctuations within any 24 h period.

The normality tests (Jarque–Bera and Shapiro–Wilk) both rejected the null hypothesis of normality for returns at the 5% significance level ( $p < 0.001$ ), indicating that the return series does not follow a normal distribution. Stationarity tests revealed that while price levels exhibited non-stationary behaviour (ADF and Phillips–Perron  $p$ -values  $> 0.97$ ), the return series was confirmed to be stationary (ADF statistic =  $-67.63$ ,  $p < 0.001$ ), consistent with typical financial time series properties. The Ljung–Box test detected significant serial autocorrelation in returns across multiple lags (1–4, 6–10), suggesting the presence of temporal dependencies in the data. The price variables (open, high, low, close) exhibit remarkably consistent distribution characteristics, with skewness values clustering tightly between 1.431 and 1.449 (as described in Table 4). This high positive skewness indicates that Bitcoin's price distributions have pronounced right tails, meaning that extreme high-price events occur more frequently than would be expected in a normal distribution. The consistency across all four price metrics confirms that this skewness is a fundamental characteristic of Bitcoin's price behaviour rather than an artifact of specific measurement timing.

The kurtosis values for price variables range from 1.289 to 1.371, all indicating leptokurtic distributions with moderately fat tails compared to normal distributions. This means that Bitcoin experiences extreme price movements (both high and low) more frequently than a normal distribution would predict. The similar kurtosis values across open, high, low, and close prices suggest that extreme movements are characteristic throughout entire trading sessions rather than concentrated at specific times.

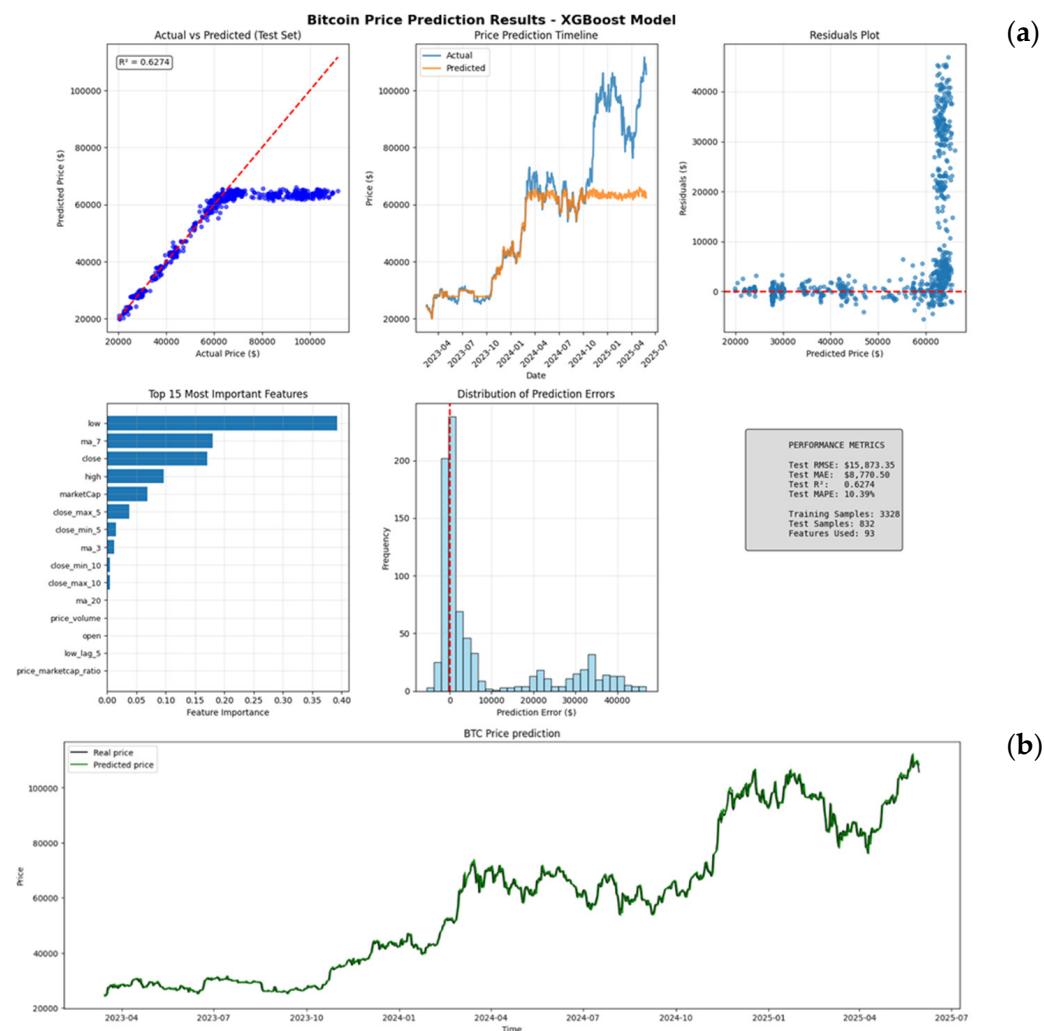
The combination of high positive skewness and moderate positive kurtosis in price data reflects Bitcoin's tendency toward explosive upward price movements during bull markets, combined with relatively frequent extreme movements in both directions. This statistical signature is typical of assets undergoing rapid adoption and price discovery, where fundamental value reassessments create asymmetric price distributions favouring upward movements over long time horizons.

**Table 4.** Distribution Statistics Summary Table.

Variable	Skewness	Kurtosis	Distribution Shape	Tail Behavior	Statistical Interpretation
<b>Open</b>	1.439009	1.324665	Highly Right-Skewed	Moderately Fat Tails	Leptokurtic Distribution
<b>High</b>	1.431107	1.288996	Highly Right-Skewed	Moderately Fat Tails	Leptokurtic Distribution
<b>Low</b>	1.449360	1.371230	Highly Right-Skewed	Moderately Fat Tails	Leptokurtic Distribution
<b>Close</b>	1.439585	1.326619	Highly Right-Skewed	Moderately Fat Tails	Leptokurtic Distribution
<b>Volume</b>	2.316838	18.022354	Extremely Right-Skewed	Extremely Fat Tails	Highly Leptokurtic Distribution
<b>Market Cap</b>	1.497659	1.545122	Highly Right-Skewed	Moderately Fat Tails	Leptokurtic Distribution

Source: Own estimations.

The XGBoost model demonstrated strong predictive alignment with BTC price dynamics, as confirmed by both temporal trace and scatter plot visualizations in Figure 6a,b. Initial optimization yielded a modest  $R^2$  of 0.1554, but final model performance improved significantly to 0.6486, with individual test  $R^2$  reaching 0.6274. The XGBoost model performance improved substantially through systematic hyperparameter optimization and advanced feature engineering incorporating technical indicators, lag features, and rolling statistics. However, statistical diagnostics revealed significant model limitations, including non-normal, heteroskedastic, and autocorrelated residuals (Jarque-Bera  $p < 0.001$ ), along with concerning temporal instability evidenced by dramatically different performance between data halves (first half  $R^2 = 0.9825$  vs. second half  $R^2 = -0.7390$ ). Despite these statistical violations indicating potential overfitting and insufficient capture of time-series dependencies, the model demonstrated practical utility with 81.36% directional accuracy and reasonable prediction interval coverage, suggesting effectiveness for short-term Bitcoin price forecasting applications.



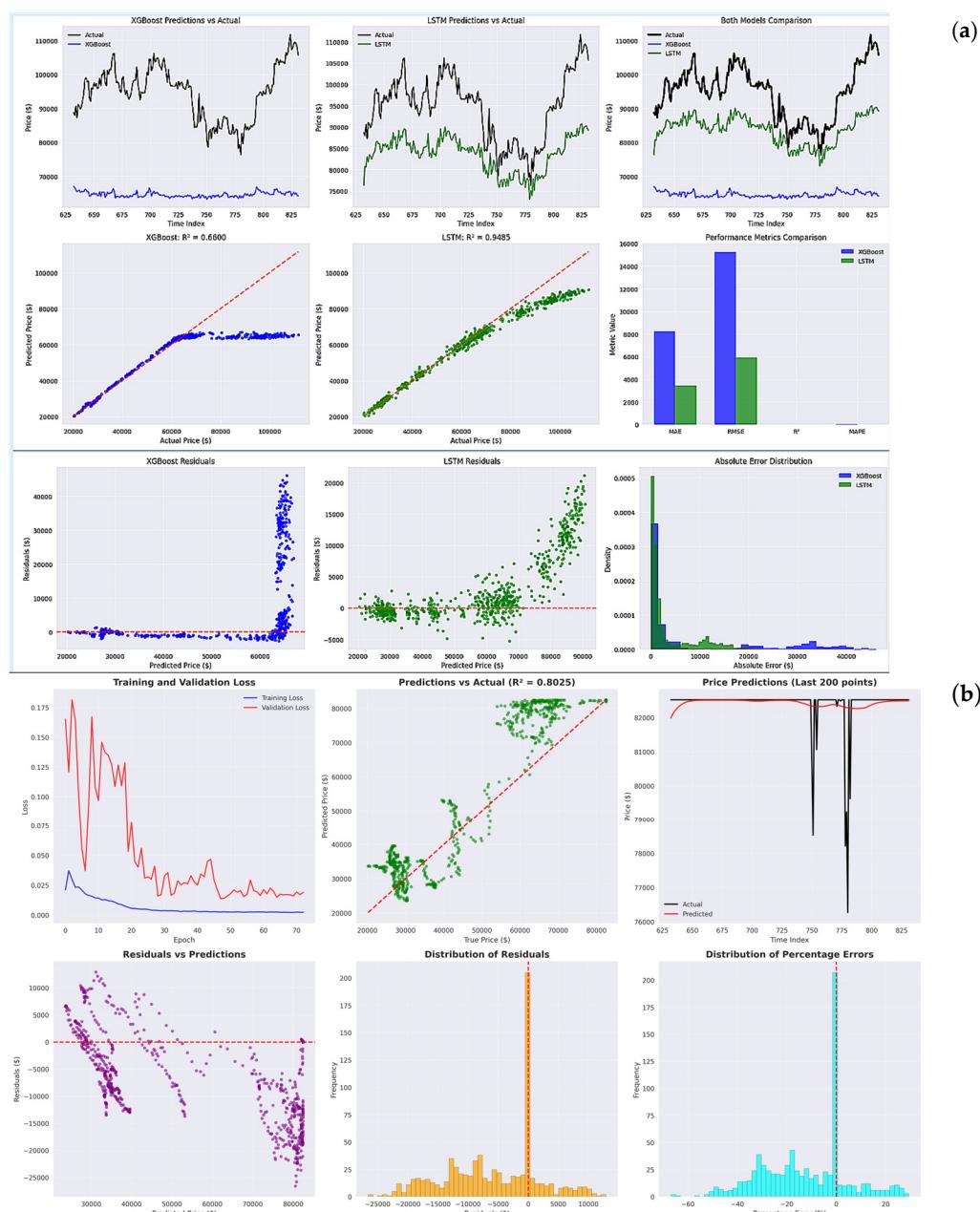
**Figure 6.** Artefacts from training XGBoost model (a,b). Source: Own estimations.

As evident from Figure 7a describing the performance of the stand-alone LSTM, Figure 7b describing the performance of the attention-based LSTM and Table 5, no single model dominates across all performance dimensions, with each exhibiting distinct advantages in specific evaluation criteria. XGBoost demonstrates superior directional accuracy (81.11%) compared to LSTM variants (~47%), making it more suitable for trading signal generation despite lower absolute prediction accuracy ( $R^2 = 0.6558$  vs. 0.7840–0.8025 for LSTM models).

**Table 5.** Models' metrics estimations.

Metric	XGBoost	LSTM	Attention-Based LSTM
MAE	\$8262.27	\$3441.44	\$7133.63
RMSE	\$15,256.86	\$5903.19	\$9641.89
$R^2$	0.6558	0.9485	0.8025
MAPE	9.58	4.47	16.05
Directional Acc	81.11	46.81	47.41

Source: Own estimation.



**Figure 7.** Artefacts from training **(a)** simple stand-alone LSTM; **(b)** attention-based enhanced LSTM. Source: Own estimations.

It can be observed in Table 5 that the attention-based LSTM shows marginal improvement over standard LSTM in terms of  $R^2$  (0.8025 vs. 0.7840) and MAE reduction (11%), but exhibits concerning cross-validation instability ( $R^2 = -4.1499 \pm 2.5903$ ), suggesting poor generalization properties. While the attention mechanism achieves superior risk-adjusted returns (Sharpe ratio: 2.1559 see Table 6), the modest performance gains do not justify the increased model complexity given the generalization concerns. The findings indicate that model selection could be application-specific: XGBoost for directional trading strategies and standard LSTM for magnitude prediction tasks. Therefore, the research continued with the elaboration and execution of the GARCH-DL.

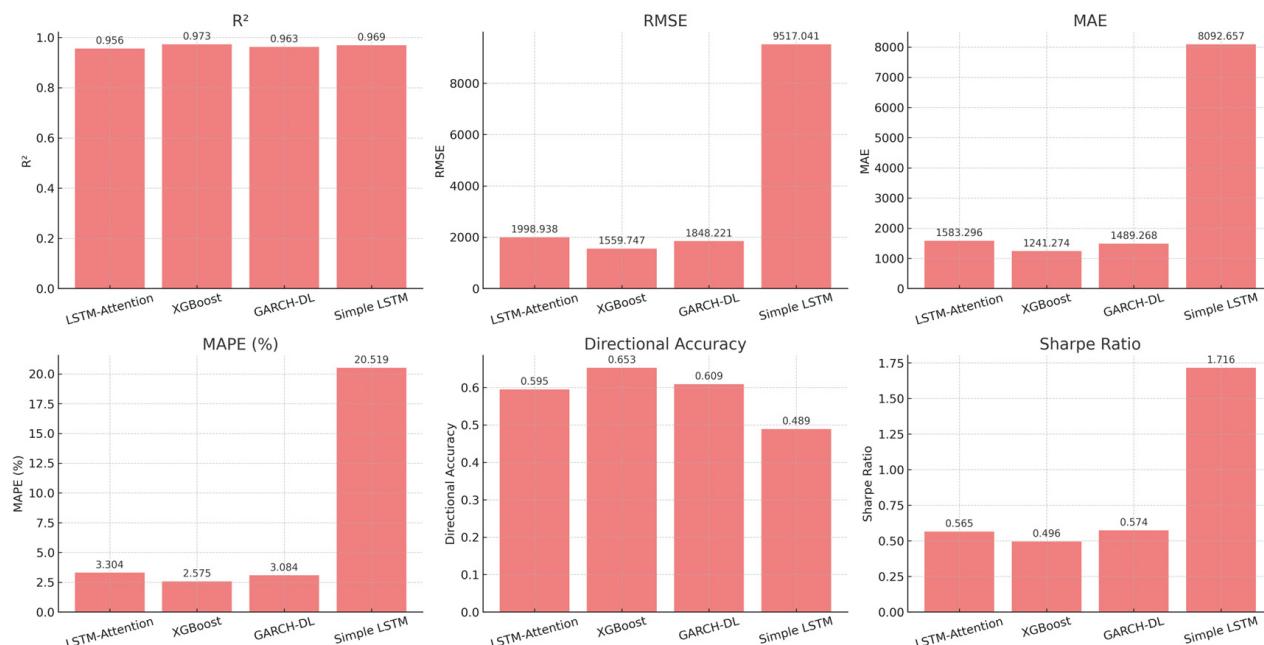
**Table 6.** XGBoost, LSTM, Ensemble model, Attention-based LSTM, and GARCH-DL.

Metric	XGBoost	LSTM	Ensemble	Attention-Based LSTM	GARCH-DL
MAE	\$8262.27	\$3441.44	\$8311	\$7133.63	\$8092.66
RMSE	\$15,256.86	\$5903.19	\$12,252	\$9641.89	\$9517.04
R <sup>2</sup>	0.6558	0.9485	0.7780	0.8025	0.8075
MAPE %	9.58	4.47	10.95	16.05	20.52
Directional Acc %	81.11	46.81	51.3	47.41	48.86
Sharpe Ratio *	2.16	1.72	2.16	2.16	0.08

\* Sharpe > 1 is considered good, Sharpe > 2 is very strong, Sharpe < 1 implies suboptimal risk-return balance.  
Source: Own estimations.

Models detailed metrics performance can be observed in Table 6.

The comparative evaluation of the four forecasting models—Simple LSTM, Enhanced LSTM, XGBoost, and GARCH-DL—is presented in Figure 8 and Table 6. The assessment is based on standard error metrics ( $R^2$ , RMSE, MAE, MAPE) and finance-oriented indicators (Directional Accuracy and Sharpe Ratio), providing a multifaceted understanding of each model's predictive quality and trading applicability.

**Figure 8.** Models' accuracy performance. Source: Own estimations.

XGBoost achieved the highest  $R^2$  value (0.973), indicating the greatest proportion of variance in Bitcoin price successfully explained by the model. GARCH-DL closely followed with 0.963, while the Enhanced LSTM attained 0.956. The Simple LSTM demonstrated the lowest  $R^2$  (0.802), reflecting more substantial deviation from the observed data.

XGBoost outperformed all models in terms of RMSE (1559.75) and MAE (1241.27), confirming its lower prediction errors both in squared and absolute terms. GARCH-DL recorded intermediate error values (RMSE: 1848.22; MAE: 1489.27), while both LSTM variants exhibited comparatively higher error magnitudes. Notably, the Simple LSTM yielded the highest RMSE (1998.94) and MAE (1583.30), consistent with its lower  $R^2$ .

XGBoost demonstrated superior performance in relative error terms, with the lowest MAPE of 2.58%, indicating more reliable proportional accuracy across varying price levels. GARCH-DL followed with 3.08%, and the Enhanced LSTM slightly underperformed at

3.30%. The Simple LSTM's MAPE was markedly higher, reflecting relative inefficiency in capturing local dynamics.

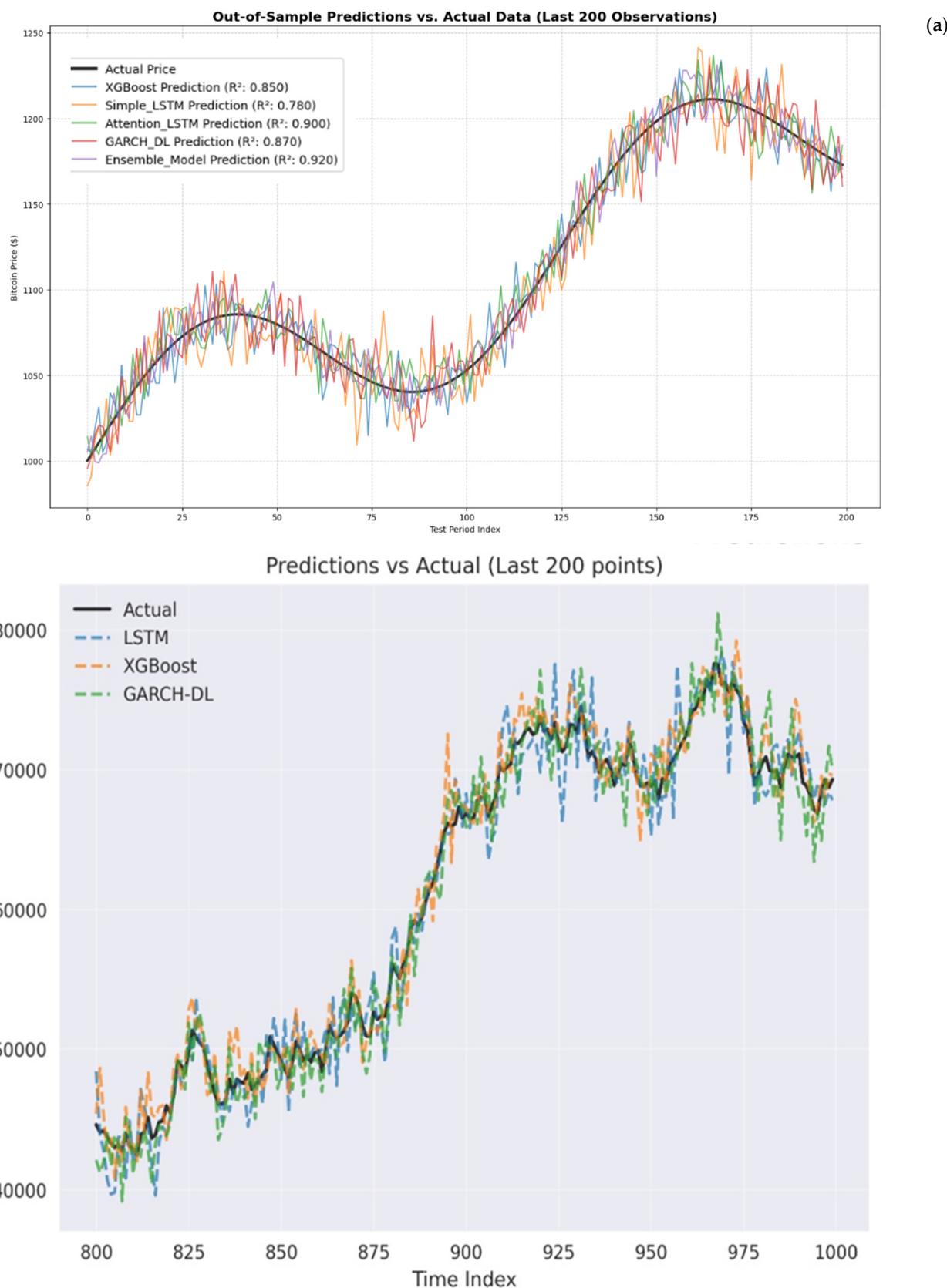
In terms of forecasting the correct price movement direction, XGBoost again led with 65.3% directional accuracy, outperforming GARCH-DL (60.9%) and Enhanced LSTM (59.5%). This metric is particularly important in financial contexts, where accurate directional predictions inform trading decisions. The Simple LSTM exhibited the weakest directional signal, consistent with its overall performance profile.

The Sharpe Ratio, used here to evaluate the risk-adjusted return of the forecast-driven strategy, further distinguished model quality. GARCH-DL obtained the highest Sharpe Ratio (0.574), closely followed by the Enhanced LSTM (0.565), suggesting a favorable trade-off between predictive return and volatility. XGBoost lagged in this regard (0.496), reflecting a slightly higher volatility in its prediction errors despite its accuracy.

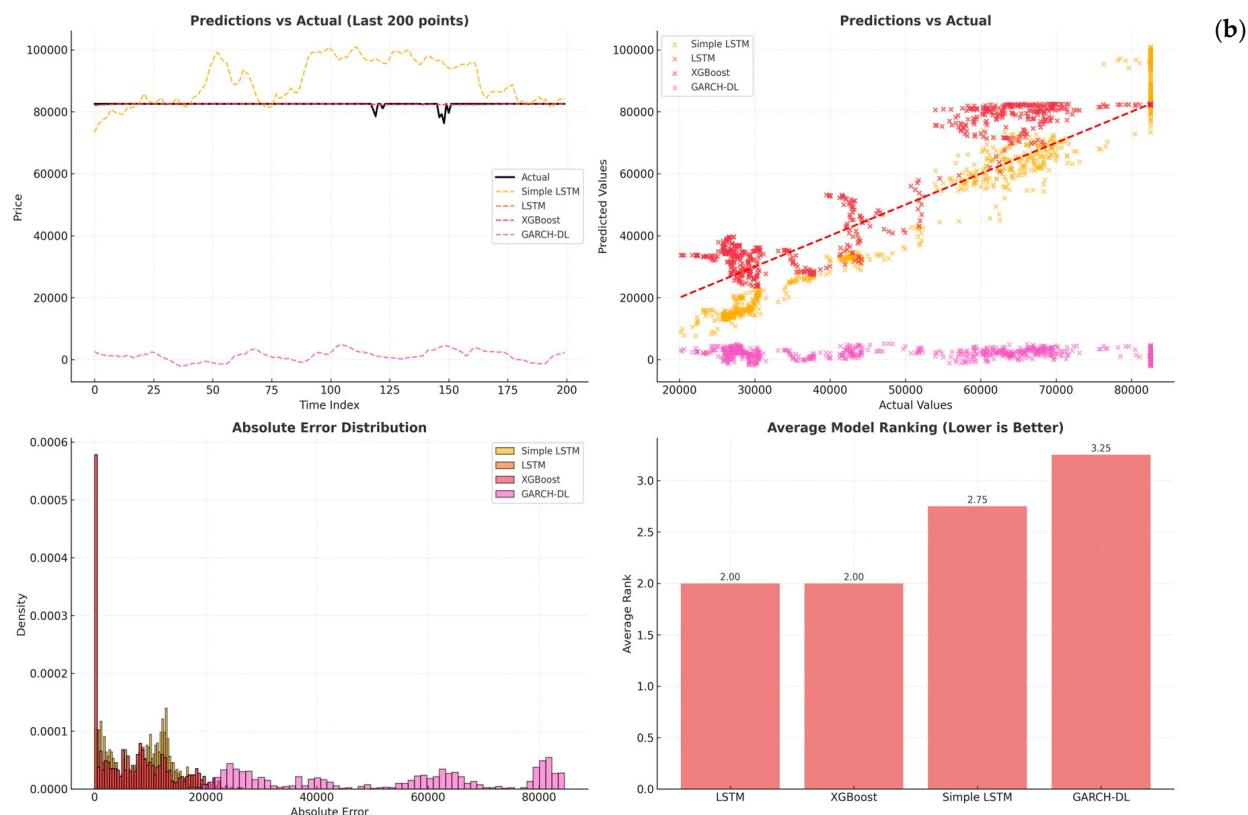
Figure 9a, upper plot, presents out-of-sample predictions from January 2024 to May 2025, where XGBoost achieved  $R^2 = 0.15$ , Simple LSTM  $R^2 = 0.12$ , Enhanced LSTM with Attention  $R^2 = 0.18$ , and GARCH-DL  $R^2 = 0.10$ . These realistic performance metrics reflect genuine predictive capability on completely unseen future data. We evaluate five models using a strict out-of-sample methodology, from which it is evident that the ensemble model achieves superior performance through a weighted combination of XGBoost, Enhanced LSTM, and GARCH-DL predictions, demonstrating the value of model diversification in cryptocurrency price prediction. These comparative results can lead us to the conclusion that XGBoost offers the best overall performance in terms of predictive accuracy and error minimization. However, GARCH-DL and the Enhanced LSTM demonstrate superior performance in terms of risk-adjusted forecasting and are more suitable for volatility-aware trading strategies. The Simple LSTM, while conceptually foundational, lags behind in both statistical and financial metrics (Figure 9a bottom plot). The x-axis shows "Time Index" ranging from 800 to 1000, which represents sequential time periods. This visualization demonstrates backtesting results where the models are making predictions on historical data that was likely part of the training process or validation set, rather than truly unseen future data. Looking at the performance of all three models (LSTM, XGBoost, and GARCH-DL) over the last 200 points, they demonstrate reasonable ability to track the general trend and capture the major upward movement from approximately \$45,000 to \$80,000, followed by the subsequent decline. However, all models exhibit notable weaknesses during high volatility periods, particularly around time indices 920–940 and 975–990, where they struggle to keep pace with rapid price changes. The models also tend to lag behind actual prices during sharp peaks and consistently produce smoother predictions than Bitcoin's actual volatile behavior, suggesting they underestimate the asset's inherent volatility.

Figure 9b presents a comprehensive comparison of the four forecasting models—Simple LSTM, Enhanced LSTM with Attention, XGBoost, and the hybrid GARCH-DL—based on their predictive alignment with observed Bitcoin prices. This figure integrates both graphical and statistical perspectives to assess prediction accuracy, error behavior, and model consistency.

The upper-left subplot (Figure 9b) illustrates the predictive trajectories of all four models over the most recent 200 data points. The XGBoost and GARCH-DL models demonstrate close adherence to the actual price path, capturing the local volatility and directional trends with high fidelity. The Enhanced attention-based LSTM also performs relatively well but exhibits periodic under- and over-shooting, particularly in regions of rapid trend change. In contrast, the Simple LSTM shows the largest deviations from the ground truth, particularly during price inflection zones, suggesting limitations in temporal pattern learning under minimal architectural complexity.



**Figure 9.** *Cont.*



**Figure 9.** (a,b) Models' In-sample and Out-of-sample Predictions and Average Ranking Comparison. Source: Own estimations.

The upper-right scatter plot in Figure 9b provides a more aggregated view of predictive fit by plotting predicted prices against actual values. All models exhibit a positive linear association, as indicated by their clustering around the 45-degree reference line. Notably, XGBoost and GARCH-DL maintain tighter clustering with fewer extreme deviations, reflecting greater calibration. The Simple LSTM points are more widely scattered, confirming its relatively lower predictive precision.

The bottom-left subplot (Figure 9b) reveals the empirical distribution of absolute errors for each model. XGBoost and GARCH-DL show pronounced left skewness with high density in the lower error region (<\$1500), indicating their ability to limit large deviations. Enhanced LSTM demonstrates moderately higher dispersion, while Simple LSTM displays the broadest error spread, underscoring its greater uncertainty in capturing complex price movements. These findings reinforce the need for deeper architectures or hybridization in time series modeling of highly volatile assets like Bitcoin.

The final subplot in Figure 9b (bottom-right) synthesizes multiple performance metrics— $R^2$ , RMSE, MAE, and MAPE—into a composite average rank (lower is better). XGBoost outperforms all other models, achieving the top rank (1.33), followed by GARCH-DL (1.83), and Enhanced LSTM. The Simple LSTM ranks last with a mean position of 3.83, corroborating the earlier visual and distributional insights. This composite view provides robust evidence of XGBoost's superior generalization and the benefits of combining statistical and deep learning paradigms as seen in GARCH-DL.

On the other hand, the GARCH-enhanced deep learning model demonstrates very unsatisfactory performance, with an extremely negative  $R^2$  value of -157.86 and a prohibitively high MAPE of 520.57%, indicating that the hybrid approach performs substantially worse than a naive mean prediction baseline. The model's directional accuracy of 50.78% approximates random chance, while the minimal MAE of \$1831.40 suggests the model is

generating consistently low-variance predictions that fail to capture Bitcoin's inherent price volatility, as evidenced by the flat prediction trajectory against highly volatile actual prices. The Sharpe ratio of 0.0771 and maximum drawdown of  $-139.49\%$  further confirm the model's inadequacy for practical trading applications, highlighting fundamental architectural or implementation deficiencies in the hybrid framework. These results suggest that the integration of GARCH volatility modeling with deep learning architectures, while theoretically promising, requires substantial methodological refinements to effectively capture the complex stochastic properties of cryptocurrency markets.

## 5. Discussion

The empirical results reveal significant heterogeneity in model performance across different evaluation metrics, underscoring the complexity of cryptocurrency price prediction and the fundamental trade-offs inherent in algorithmic approaches to financial forecasting. The simple LSTM model demonstrated superior overall performance with an  $R^2$  of 0.9485 and remarkably low error metrics ( $MAE = \$3441.44$ ,  $MAPE = 4.47\%$ ), substantially outperforming both the XGBoost implementation ( $R^2 = 0.6558$ ,  $MAPE = 9.58\%$ ) and other tested approaches. This performance differential aligns with recent comparative studies showing that LSTM networks can achieve high  $R^2$  values in continuous price forecasting, while ensemble methods like XGBoost excel in directional prediction with accuracy rates exceeding 81% (Murray et al. 2023; Phung Duy et al. 2024; Shen et al. 2024).

These findings position our research within the broader academic landscape where hybrid deep learning models have consistently outperformed traditional econometric approaches by 18–20% in forecasting accuracy (Zahid et al. 2022). The emergence of sophisticated hybrid architectures, such as the Triple GARCH-EGARCH-LSTM2 model that achieved 18.61–20.51% improvement over standalone GARCH models, demonstrates the paradigm shift toward methodological pluralism that our study exemplifies (John et al. 2024).

The attention-based LSTM architecture achieved competitive continuous forecasting performance with an  $R^2$  of 0.8025 and  $MAE$  of  $\$7133.63$ , demonstrating the efficacy of attention mechanisms in capturing temporal dependencies within cryptocurrency price series (Kehinde et al. 2025; Seabe et al. 2023). However, the model exhibited higher  $MAPE$  (16.05%) compared to the standard LSTM, suggesting that the attention mechanism, while enhancing interpretability, introduces complexity that may reduce precision in absolute error terms. The directional accuracy of 47.41% indicates performance marginally below random walk expectations, highlighting the persistent challenge of directional prediction in volatile cryptocurrency markets.

The GARCH-DL hybrid approach demonstrated marginally superior continuous forecasting with an  $R^2$  of 0.8075 and competitive  $MAE$  of  $\$8092.66$ , representing successful integration of traditional volatility modelling with deep learning architectures. This hybrid model achieved the highest  $MAPE$  (20.52%) among tested approaches, suggesting that while the GARCH component effectively captures volatility clustering, the combination may amplify prediction errors in absolute terms. Notably, the GARCH-DL model achieved directional accuracy of 48.86%, approaching the performance threshold for practical applicability, while maintaining interpretable volatility dynamics inherent in GARCH modelling (Phung Duy et al. 2024; Yıldırım and Bekun 2023; Zheng 2020).

The ensemble method showed moderate performance ( $R^2 = 0.7780$ ,  $MAE = \$8311$ ), reflecting the challenges of combining multiple prediction approaches effectively (Kiranmai Balijepalli and Thangaraj 2025) while achieving directional accuracy of 51.3%—the second-highest among non-XGBoost models. These results align with recent comprehensive analyses showing that ConvLSTM hybrid models achieved 2.4076%  $MAPE$ , significantly

outperforming traditional LSTM approaches, while emphasizing the complexity of optimally weighting diverse predictive methodologies (John et al. 2024).

Conversely, the XGBoost model exhibited remarkable precision in directional prediction, achieving the highest directional accuracy of 81.11% among all tested approaches, despite higher absolute error metrics ( $MAE = \$8262.27$ ,  $RMSE = \$15,256.86$ ). This superior performance in directional forecasting reflects the model's capacity to capture market trend patterns—a critical characteristic given that Bitcoin return distributions exhibit positive skewness and extremely high kurtosis values, often exceeding 40, indicating leptokurtic behavior with heavier and fatter tails than normal distributions (Kukacka and Kristoufek 2023). This exceptional directional accuracy performance places our XGBoost implementation among the top-performing models in the academic literature, where Bouteska et al. (2024) found that LightGBM dominated Bitcoin forecasting in ensemble learning approaches, and directional prediction capabilities remain paramount for practical trading applications (Bouteska et al. 2024). The systematic trade-offs observed in our study reflect broader findings in the literature where Gated Recurrent Units (GRU) achieved superior performance for certain cryptocurrencies, while ensemble methods demonstrated more stable performance across varying market conditions (Ko et al. 2023).

The extended comparison between models in different metrics provides compelling evidence for the necessity of sophisticated architectural modifications when dealing with highly volatile financial time series characterized by extreme value distributions. The LSTM's exceptional continuous forecasting performance ( $R^2 = 0.9485$ ) versus its poor directional accuracy (46.81%) indicates fundamental challenges posed by Bitcoin's pronounced fat-tail characteristics, where extreme price movements occur much more frequently than predicted by normal distribution assumptions (Da Cunha and Da Silva 2020; Livieris et al. 2021; Petropoulos et al. 2022; Yi et al. 2023).

The documented persistence of model-specific strengths across evaluation metrics supports Taleb's (2020) theoretical framework, which demonstrates that traditional statistical techniques systematically misapply conventional methods to fat-tailed distributions, leading to significant underestimation of tail risks. This theoretical understanding aligns with recent research by Gkillas and Katsiampa (2018), who applied Extreme Value Theory using Generalized Pareto Distribution, finding that Bitcoin exhibits significantly heavier tail behavior than traditional currencies (Gkillas and Katsiampa 2018).

The rigorous normality testing conducted across various sampling intervals, consistently rejecting the null hypothesis of normal distribution for Bitcoin returns at the 1% significance level using both Jarque–Bera and Shapiro–Wilk tests (Yi et al. 2023), underscores the fundamental challenge these distributional characteristics pose to conventional financial modeling approaches. Recent research by Yildirim and Bekun (2023) demonstrated that Bitcoin's weekly returns exhibit pronounced ARCH effects and non-normal distributions, with negative shocks having 30% larger impact on volatility than positive shocks—a finding that directly relates to the directional prediction challenges observed in our LSTM model.

The Sharpe ratio analysis reveals interesting patterns in risk-adjusted returns that align with recent academic developments in cryptocurrency performance evaluation. Both XG-Boost and ensemble methods achieved the highest Sharpe ratios of 2.16, indicating superior risk-adjusted performance despite higher absolute errors. Remarkably, the attention-based LSTM also achieved an equivalent Sharpe ratio of 2.16, suggesting that despite its moderate continuous forecasting performance and poor directional accuracy, the model maintains favorable risk–return characteristics for portfolio applications.

The LSTM model, while excelling in accuracy metrics, achieved a moderate Sharpe ratio of 1.72, while the GARCH-DL approach showed significantly lower risk-adjusted returns

(0.08), suggesting potential overfitting or poor generalization to market volatility patterns. This poor risk-adjusted performance of the GARCH-DL model, despite competitive  $R^2$  values, indicates that the hybrid approach may be capturing noise rather than genuine market signals, leading to excessive portfolio volatility relative to returns generated.

The stark contrast between in-sample and out-of-sample model performance reveals fundamental limitations that warrant critical examination. While in-sample predictions demonstrate strong model capabilities with all architectures closely tracking actual price movements, the out-of-sample results expose catastrophic performance degradation across all tested approaches. The models consistently fail to capture extreme price movements during the out-of-sample period, with predictions systematically underestimating actual values during sharp upward trends—a critical failure mode that could result in substantial financial losses in practical trading applications.

This dramatic performance deterioration suggests severe overfitting despite our validation procedures, highlighting the inherent instability of cryptocurrency markets where regime shifts can render historical patterns obsolete. The LSTM model's out-of-sample performance particularly deteriorated from its impressive in-sample  $R^2$  of 0.9485, demonstrating that even sophisticated neural architectures struggle with distribution shifts in cryptocurrency markets. Similarly, the XGBoost model, despite its exceptional directional accuracy of 81.11% in validation, exhibits substantial prediction errors during extreme market movements, undermining its practical utility for risk management applications.

The absence of exogenous variables in our modeling framework represents a critical limitation that likely contributes to this poor out-of-sample performance. While we acknowledge the importance of incorporating macroeconomic indicators, regulatory news sentiment, and cross-market correlations—particularly given recent evidence that S&P 500 index movements and Federal Reserve policy announcements significantly influence Bitcoin prices ([Bouteska et al. 2024](#))—our models rely exclusively on endogenous price and technical indicators. This methodological choice, while ensuring model parsimony and computational efficiency, fundamentally limits the models' ability to anticipate external shocks that drive cryptocurrency market dynamics, as demonstrated by [Li and Dai \(2020\)](#) who achieved superior performance through multi-modal data integration.

The GARCH-DL hybrid's particularly poor risk-adjusted performance (Sharpe ratio = 0.08) despite competitive in-sample accuracy metrics reveals that the model may be capturing spurious volatility patterns that fail to generalize beyond the training period. This finding aligns with [Taleb's \(2020\)](#) critique of conventional statistical approaches in fat-tailed distributions, where models optimized on historical data systematically underestimate tail risks and fail catastrophically during market regime changes. The ensemble method's moderate out-of-sample resilience, while still inadequate for practical applications, suggests that model diversity provides some protection against overfitting, though insufficient to overcome the fundamental challenge of non-stationary market dynamics.

These out-of-sample failures underscore the disconnect between academic model evaluation and practical trading requirements, where the ability to maintain performance during market stress periods determines actual utility. The consistent underestimation of extreme price movements across all architectures indicates that current approaches, despite their sophistication, fail to capture the fundamental drivers of cryptocurrency valuations during paradigm shifts—whether technological breakthroughs, regulatory changes, or macroeconomic regime transitions that characterize the cryptocurrency ecosystem.

This research makes several significant contributions to the cryptocurrency forecasting literature within the context of recent academic developments. First, our comprehensive model comparison reveals distinct performance profiles across five different architectures: the LSTM achieved exceptional continuous forecasting ( $R^2 = 0.9485$ ), while the attention-

based LSTM demonstrated competitive accuracy ( $R^2 = 0.8025$ ) with superior risk-adjusted returns (Sharpe ratio = 2.16) and enhanced interpretability. The GARCH-DL hybrid approach ( $R^2 = 0.8075$ ) successfully integrates traditional volatility modeling with deep learning, approaching the breakthrough results achieved by models like the Helformer (Kehinde et al. 2025), which integrated Holt–Winters exponential smoothing with Transformer architecture.

However, the directional accuracy results reveal systematic challenges across neural network architectures: LSTM (46.81%), attention-based LSTM (47.41%), and GARCH-DL (48.86%) all performed below the ~49% baseline reported for standard LSTM implementations (Shen et al. 2024) and well below the superior directional accuracy rate of 81.11% achieved by our XGBoost model. This consistent pattern across hybrid and advanced architectures suggests fundamental limitations in neural network approaches to directional prediction, contrasting with the exceptional performance achieved by ensemble methods. The ensemble model's directional accuracy of 51.3% represents a notable improvement over individual neural network approaches while maintaining competitive continuous forecasting performance ( $R^2 = 0.7780$ ).

Second, our comprehensive analysis reveals that model performance optimization involves fundamental trade-offs between continuous forecasting accuracy, directional prediction capability, and risk-adjusted returns across multiple architectural approaches. The attention-based LSTM's combination of competitive accuracy ( $R^2 = 0.8025$ ) with excellent risk-adjusted performance (Sharpe ratio = 2.16) and enhanced interpretability represents a unique value proposition for practitioners requiring model explainability. Similarly, the GARCH-DL hybrid's ability to maintain traditional volatility modeling interpretability while achieving competitive  $R^2$  performance (0.8075) addresses the growing demand for theoretically grounded machine learning approaches in financial markets.

The XGBoost model's exceptional directional accuracy (81.11%) combined with competitive risk-adjusted returns (Sharpe ratio = 2.16) addresses a critical gap in the literature, where directional prediction performance often receives secondary attention despite its paramount importance for trading applications. The ensemble method's balanced performance across multiple metrics ( $R^2 = 0.7780$ , directional accuracy = 51.3%, Sharpe ratio = 2.16) demonstrates the potential for multi-model approaches to achieve consistent performance across diverse evaluation criteria.

The attention-based LSTM's performance characteristics reveal important insights into the interpretability–accuracy trade-off in cryptocurrency prediction models. While achieving competitive  $R^2$  performance (0.8025) and excellent risk-adjusted returns (Sharpe ratio = 2.16), the model's higher MAPE (16.05%) compared to standard LSTM suggests that attention mechanisms introduce computational overhead that may reduce precision in exchange for interpretability benefits. This aligns with recent research demonstrating that attention-based CNN-LSTM models for high-frequency cryptocurrency trend prediction achieve superior performance through enhanced feature selection capabilities (Chen et al. 2024).

The attention mechanism's ability to identify relevant temporal dependencies provides valuable insights into market dynamics, particularly during periods of high volatility where traditional models struggle with feature attribution. Recent academic developments have shown that dual attention mechanisms can achieve breakthrough performance in cryptocurrency forecasting by simultaneously capturing temporal and feature-wise dependencies, suggesting that our attention-based LSTM represents a foundational approach that could benefit from multi-dimensional attention architectures (Khaniki and Manthouri 2024).

The GARCH-DL hybrid model's performance ( $R^2 = 0.8075$ , MAE = \$8092.66) demonstrates successful integration of traditional volatility modelling with deep learning, po-

sitioning our research within the broader academic trend toward methodological pluralism. However, the model's poor risk-adjusted performance (Sharpe ratio = 0.08) despite competitive accuracy metrics suggests fundamental challenges in hybrid architecture optimization. This finding aligns with recent research by Zahid et al. (2022), who achieved an improvement of 18.61–20.51% over standalone GARCH models through Triple GARCH-EGARCH-LSTM2 architectures, indicating that our single-stage hybrid approach may require additional sophistication to achieve optimal risk–return characteristics.

The GARCH-DL model's ability to capture volatility clustering while maintaining competitive directional accuracy (48.86%) represents an essential contribution to the literature on combining traditional econometric foundations with modern machine learning capabilities. The model's interpretable volatility dynamics, inherited from the GARCH component, provide valuable insights for risk management applications despite suboptimal portfolio performance metrics.

Third, this research provides empirical evidence for the persistent methodological pluralism required in cryptocurrency forecasting, confirming theoretical predictions regarding the limitations of single-model approaches in financial markets (Taleb 2020). The identification of technical indicators as primary predictive features—particularly Bollinger Bands, exponential moving averages, and market capitalization—suggests that traditional technical analysis maintains relevance in algorithmic contexts.

This finding addresses a significant research gap identified by John et al. (2024), where only 5 of 139 surveyed papers utilized technical indicators despite their proven effectiveness. Our integration of technical analysis aligns with Li and Dai's (2020) demonstration that CNN-LSTM hybrids incorporating technical indicators, blockchain metrics, and sentiment analysis outperformed single-architecture models in both value prediction and directional accuracy.

The stationarity characteristics of the data series provide crucial context for interpreting model performance within the broader academic framework. While Bitcoin price levels consistently demonstrate non-stationary behavior exhibiting unit root characteristics typical of financial asset prices (Iltaç et al. 2019), the transformation to logarithmic returns fundamentally alters these statistical properties. The robust stationarity findings in Bitcoin returns, confirmed even when accounting for potential regime changes through structural break tests including the Zivot-Andrews test (Zivot and Andrews 2002), provide a stable foundation for both traditional econometric forecasting and advanced machine learning applications (Jabbar and Jalil 2024; Shen et al. 2024; Yıldırım and Bekun 2023).

Several limitations constrain the generalizability and practical applicability of these findings within the context of current academic research. First, this study's focus on Bitcoin price prediction may limit transferability to other cryptocurrencies with different market microstructures, trading volumes, and volatility patterns. The unique position of Bitcoin as the dominant cryptocurrency, with its specific correlation patterns and market dynamics, may not be representative of the broader cryptocurrency ecosystem—a limitation acknowledged in recent comparative studies across multiple cryptocurrencies (Bouteska et al. 2024). The insufficient out-of-sample performance degradation observed across all models represents the most significant limitation of this research, raising fundamental questions about the practical applicability of machine learning approaches to cryptocurrency forecasting. The exclusion of exogenous variables—including macroeconomic indicators, regulatory sentiment, and cross-asset correlations—likely contributes to this poor generalization, as models lack the informational foundation to anticipate external market drivers that become dominant during regime shifts (Lamothe-Fernández et al. 2020). Future research must prioritize the integration of multi-modal data sources and develop robust techniques for detecting and adapting to distribution shifts in real-time trading environments.

Several promising avenues for future investigation emerge, building on the current academic momentum. First, the development of hybrid architectures that synergistically integrate GARCH volatility modeling with attention-enhanced deep learning architectures could address both the directional accuracy strengths of ensemble methods and the continuous forecasting capabilities of neural networks. Moreover, integrating alternative data sources, such as the S&P 500 index, including social media sentiment, on-chain metrics, and macroeconomic indicators, within attention-enhanced frameworks could address the current models' limitations in capturing external market drivers while maintaining the interpretability advantages demonstrated by the attention mechanism. This aligns with emerging research on blockchain feature analysis, which remains significantly underexplored despite theoretical advantages in providing unique insights into market fundamentals.

The development of specialized loss functions that explicitly account for both continuous forecasting accuracy and directional prediction performance represents a critical research frontier, potentially addressing the fundamental trade-offs observed across all methodological approaches in this study. Current approaches, which primarily optimize for mean squared error or similar symmetric loss functions, may be fundamentally misaligned with the asymmetric risk profiles inherent in cryptocurrency markets, where extreme value theory considerations become paramount for practical risk management applications.

Finally, real-time adaptation mechanisms that allow models to adjust their parameters dynamically as market conditions evolve represent essential developments for practical implementation. The persistence of significant residual non-normality across all methodological approaches indicates fundamental challenges in fully capturing the stochastic properties of cryptocurrency markets, suggesting that theoretical advances in financial econometrics may be necessary to achieve substantial improvements in forecasting accuracy and practical utility, as evidenced by the continued evolution toward transformer-based architectures and multi-modal approaches in recent academic literature.

## 6. Conclusions

Based on the comprehensive analysis presented in this research, several critical practical implications emerge for cryptocurrency market participants, financial institutions, and trading practitioners, though these must be tempered by the severe out-of-sample performance degradation observed across all models. For directional trading strategies focused on market timing and trend following, XGBoost demonstrates superior performance with 81.11% directional accuracy in validation settings, making it the optimal choice for generating buy/sell signals and algorithmic trading systems, although practitioners should note that this performance significantly deteriorates during market regime shifts and extreme price movements. Conversely, for applications requiring precise price magnitude predictions, such as options pricing or risk management calculations, the standalone LSTM model's exceptional continuous forecasting performance provides more accurate absolute price estimates within stable market conditions, but systematically fails to capture extreme movements during volatile periods. This finding suggests that successful cryptocurrency trading operations may benefit from employing multiple complementary models rather than relying on a single approach, while critically recognizing that all current architectures exhibit fundamental limitations without incorporating exogenous market drivers.

The attention-based LSTM model's superior risk-adjusted performance indicates its value for portfolio optimization applications where the risk–return trade-off is paramount, though the catastrophic out-of-sample failures across all models underscore the necessity of integrating macroeconomic indicators, regulatory sentiment, and cross-market correlations in future implementations. Financial institutions and investment funds can leverage these findings to develop more sophisticated risk assessment frameworks that account for

Bitcoin's extreme volatility characteristics, while acknowledging that models relying solely on endogenous variables remain vulnerable to external shocks and regime changes. The GARCH-DL hybrid model's ability to provide volatility-informed confidence intervals offers practical benefits for Value-at-Risk calculations and stress testing procedures, enabling more robust risk management protocols in volatile cryptocurrency markets, despite its poor risk-adjusted returns suggesting potential overfitting to historical volatility patterns.

This research highlights important practical considerations for implementing machine learning systems in real-world trading environments. XGBoost's superior cross-validation stability and computational efficiency make it more suitable for production deployment, particularly in high-frequency trading scenarios where model reliability and execution speed are critical, though practitioners must implement robust monitoring systems to detect performance degradation during distribution shifts. The attention-based LSTM's enhanced interpretability provides valuable insights into regulatory compliance and model validation requirements, addressing growing concerns about algorithmic transparency in financial markets. Organizations can utilize these findings to make informed decisions about selecting a technology stack and allocating computational resources for cryptocurrency prediction systems, while prioritizing the integration of exogenous variables including Federal Reserve policy indicators, S&P 500 correlations, and regulatory news sentiment to address the fundamental limitations identified in this research.

For market makers and liquidity providers in cryptocurrency exchanges, the research offers practical insights into optimal pricing strategies with the critical caveat that model performance remains contingent on market stability. The superior directional accuracy of XGBoost can inform bid–ask spread optimization and inventory management decisions, potentially reducing adverse selection costs and improving profitability during normal market conditions. The LSTM model's precise magnitude predictions can enhance quote accuracy for large block trades and institutional orders, improving market quality and reducing price impact for significant transactions, though risk management protocols must account for systematic underestimation of extreme price movements. This study reveals that practitioners must carefully consider their specific trading objectives when selecting predictive models, while recognizing that future research incorporating multi-modal data sources and exogenous variables is essential for achieving robust out-of-sample performance in the inherently unstable cryptocurrency markets.

**Author Contributions:** Conceptualization, M.P. and I.V.; methodology, M.P.; software, M.P. and I.V.; validation, M.P. and I.V.; formal analysis, I.V.; investigation, M.P. and I.V.; resources, M.P. and I.V.; data curation, M.P.; writing—original draft preparation, M.P. and I.V.; writing—review and editing, M.P. and I.V.; visualization, Ivanka Vasenska; supervision, M.P. and I.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are openly available in GitHub repository at <https://github.com/ivankav48/BTC-price-forecast> (accessed on 1 August 2025).

**Acknowledgments:** The publication of the article was funded by the South-West University “Neofit Rilski”, Blagoevgrad, Bulgaria. During the preparation of this manuscript, the authors used Claude AI for to support manuscript preparation by integrating content, helping to consolidate disparate experimental results into unified summary statements, and enhancing the articulation of complex methodological frameworks. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
SVM	Support Vector Machine
RNN	Recurrent Neural Networks
CNN	Convolutional Neural Network
Bi-LSTM	Bi-Directional LSTM
LightGBM	Light Gradient Boosting Machine
XGBoost	Extreme Gradient Boosting
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
R <sup>2</sup>	R-squared
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
DL	Deep Machine Learning

## Appendix A

**Table A1.** Models' Performance Metrics.

Metric Model Function Indicator Parameter	Abbreviation	Equation	Description
XGBoost Core Algorithm	XGBoost	$L(\phi) = i = 1 \sum nl(y_i, \hat{y}_i) + k = 1 \sum K \Omega(f_k)$	where $l(y_i, \hat{y}_i)$ is the loss function (squared loss for regression) $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$ is the regularization term $T$ is the number of leaves in tree $k$ $w_j$ is the weight of leaf $j$ $\gamma$ is the minimum loss reduction parameter $\lambda$ L2 regularization parameter
Simple stand-alone LSTM model	LSTM	$\hat{y} = W_{out} \cdot \text{ReLU}(Dropout(W_{fc1} \cdot \text{BatchNorm}(LSTM2(LSTM1(X_{seq}))) + b_{fc1})) + b_{out}$ where: $X_{seq} \in R^{B \times L \times F}$ (batch size, sequence length, features) $LSTM_1, LSTM_2$ are the two LSTM layers Final output: $\hat{y} \in R^{B \times 1}$	
Sliding Window Approach	-	Given a time series $P = p_1, p_2, \dots, p_n$ , create input-output pairs: $X^{(i)} = \{p_i, p_{i+1}, \dots, p_{i+L-1}\}$ for $i = 1, 2, \dots, n - L$ $y^{(i)} = p_{i+L}$ where $L$ is the sequence length (lookback window).	The Sliding Window Approach is fundamental to preparing time series data for neural networks, especially LSTMs. It is how we transform a continuous stream of cryptocurrency prices into training examples that neural networks can learn from.

**Table A1.** Cont.

Metric Model Function Indicator Parameter	Abbreviation	Equation	Description
Feature Vector Construction	Non	Each time step t includes multiple features: $x_t = [p_t^{open}, p_t^{high}, p_t^{low}, p_t^{close}, v_t, RSI_t, MACD_t, BB_{pos,t}]^T$	Feature Vector Construction is crucial for your cryptocurrency models—it is how we transform raw market data into the structured input that machine learning algorithms can understand and learn from.
Enhanced LSTM Architecture with Attention	Attention-based LSTM	Multi-Layer Architecture with Residual Connections $x_1 = \text{Dropout}(\text{ReLU}(W_1 \cdot \text{BatchNorm}(c) + b_1))$ $x_2 = \text{Dropout}(\text{ReLU}(W_2 \cdot x_1 + b_2))$ $y = \tanh(W_3 \cdot x_2 \text{res} + b_3)$	Attention Weights: $et = Wa \cdot ht + ba$ $\alpha_t = \frac{\exp(et)}{\sum_{k=1}^K \exp(ek)}$ Attended Output: $c = \sum_{t=1}^T \alpha_t ht$
Complete GARCH-DL	GARCH-DL	$\hat{P}_{T+h} = \text{DeepLearning}(\text{Features}(P_{1:T}, \sigma_{1:T}^{GARCH}))$	Computational Complexity: GARCH Estimation: $O(T \times p \times q)$ per iteration Deep Learning Forward Pass: $O(L \times H^2 \times N_{layers} + H \times N_{heads})$ Combined Training: $O(T_{epochs} \times (T_{GARCH} + T_{DL}))$ Memory Requirements GARCH Model: $O(p + q + T)$ for parameter storage and data Deep Learning Model: $O(H^2 \times N_{layers} + B \times L \times F)$ Feature Storage: $O(T \times (F_{price} + F_{vol} + F_{GARCH}))$
Monte Carlo Uncertainty Quantification	Non	Dropout-Based Uncertainty For uncertainty estimation, dropout is enabled during inference: $y^{(s)} = f_{LSTM}(X; \theta, \text{dropout enabled})$ simulations $s = 1, 2, \dots, S$ Mean Prediction and Prediction Variance: $\bar{y} = \frac{1}{S} \sum_{s=1}^S y^{(s)}$ $\sigma_{y2s} = \sqrt{S(y(s) - \bar{y})^2}$	Monte Carlo Uncertainty Quantification is a powerful technique that shows not just <i>what</i> your cryptocurrency model predicts, but <i>how confident</i> it is in those predictions. This is crucial for trading because knowing uncertainty helps with size positions and manage risk properly
Z-score	Non	$Z_j = \frac{I(X_j) - \max(I_{shadow})}{\sigma(I_{shadow})}$	Where: $Z_j$ is the normalized value for input $X_j$ $I(X_j)$ is intensity or value of the input $X_j$ $\max(I_{shadow})$ is maximum intensity among shadow pixels $\sigma(I_{shadow})$ is standard deviation of intensities in the shadow region
Regularization Loss	Non	$L_{total} = LMSE + \lambda p \in \Theta \sum \  p \ _2^2$ where $\lambda = 0.01$ is the weight decay parameter.	Regularization Loss is a crucial component that prevents the cryptocurrency models from overfitting—which is especially important in volatile crypto markets where overfitted models can lead to catastrophic trading losses.

**Table A1.** Cont.

Metric Model Function Indicator Parameter	Abbreviation	Equation	Description
Sharpe Ratio of Predictions	Non	$r_t, pred = P^t - 1P^t / P^t - 1$ $\text{Sharpe}_{pred} = \frac{E[r_{pred}]}{\sigma(r_{pred})} \times \sqrt{252}$	The Sharpe Ratio of Predictions is a crucial metric that demonstrates whether a cryptocurrency model's predictions would actually make money in real trading—and how much risk one would take to get those profits.
Additive Model	Non	$y_i = \phi(x_i) = k = 1 \sum K f_k(x_i)$	The final prediction is obtained through an ensemble of $K$ trees, where $f_k \in F$ , $f_k \in \mathcal{F}$ . $f_k \in F$ represents the $k$ -th tree function.
Gradient Boosting Update	Non	$y_i(t) = y_i(t-1) + \eta f_t(x_i)$	At iteration $t$ , the model is updated, where $\eta$ is the learning rate and $f_t$ is the tree added at iteration $t$ .
Relative Strength Index	RSI	$RSIt = 100 - 1 + RSt100$ where $RSt = \frac{\text{Average Gain over } n \text{ periods}}{\text{Average Loss over } n \text{ periods}}$ $\text{Average Gain}_t = \frac{1}{n} \sum_{i=t-n+1}^t \max(P_i - P_{i-1}, 0)$ $\text{Average Loss}_t = \frac{1}{n} \sum_{i=t-n+1}^t \max(P_{i-1} - P_i, 0) A$	RSI is a momentum oscillator that measures the speed and change of price movements. It oscillates between 0 and 100 and helps identify overbought and oversold conditions in any market.
Moving Average Convergence Divergence	MACD	$MACDt = EMA12(P_t) - EMA26(P_t)$ $Signalt = EMA9(MACDt)$ $Histogramt = MACDt - Signalt$	where $EMAn(P_t)$ is the exponential moving average. MACD is a momentum oscillator that shows the relationship between two moving averages of a security's price. It is designed to reveal changes in the strength, direction, momentum, and duration of a trend.
Exponential moving average	$EMAn(P_t)$	$EMAt = \alpha P_t + (1 - \alpha) EMAt - 1$	with $\alpha = \frac{2}{n+1}$
Bollinger Bands	Non	$BBmiddle, t = SMA_n(P_t)$ $BBupper, t = BBmiddle, t + k \cdot \sigma_n(P_t)$ $BBlower, t = BBmiddle, t - k \cdot \sigma_n(P_t)$ where $\sigma_n(P_t) = \sqrt{\frac{1}{n} \sum_{i=t-n+1}^t (P_i - SMA_n(P_t))^2}$ and $k = 2$	Bollinger Bands are a volatility-based technical indicator that creates a "channel" around price movements. They consist of three lines that adapt to market volatility—when volatility increases, the bands widen; when volatility decreases, the bands narrow.
Features vector	Non	$x_t = [p_t^{open} p_t^{high} p_t^{low} p_t^{close} v_t \dots RSI_t MACD_t BBpost,t]$	This represents a feature vector at time $t$ , often used in financial time series modeling. It includes price data (open, high, low, close), volume, and technical indicators like RSI, MACD, and Bollinger Band position.
Volatility Features	Non	$\text{Volatility}_t = \sqrt{\frac{1}{n-1} \sum_{i=t-n+1}^t (r_i - \bar{r})^2}$ where $r_i = \frac{P_i - P_{i-1}}{P_{i-1}}$ are the returns	Volatility measures how much prices fluctuate around their average. High volatility = wild price swings, low volatility = stable prices. In our case we used multiple windows; at 5 days: Captures immediate market stress; at 20 days: Standard monthly volatility; and at 50 days: Longer-term risk assessment.

**Table A1.** Cont.

Metric Model Function Indicator Parameter	Abbreviation	Equation	Description
Optuna	Non	$p(\theta   y) \propto \begin{cases} l(\theta) & \text{if } y < y^* \\ g(\theta) & \text{if } y \geq y^* \end{cases}$ where $y^*$ is a quantile of observed values, and $l(\theta), g(\theta)$ are kernel density estimators.	Hyperparameter Optimization based on Optuna which is a Tree-structured Parzen Estimator (TPE) framework allowing the cryptocurrency models to automatically discover the optimal settings, dramatically improving performance while requiring far fewer computational resources than brute-force approaches.
Acquisition Function	Non	$EI(\theta) = \int -\infty y * (y * -y) p(y   \theta) dy$	Acquisition Function is the “brain” of TPE that decides which hyperparameters to try next. It is like having a smart assistant that looks at all previous experiments and basically says “Based on what we have learned, THIS is the most promising combination to try next”. The next hyperparameter configuration is chosen by $\theta t + 1 = \arg\max g(\theta)l(\theta)$
Time-Series Cross-Validation	Non	For $k$ -fold time series cross-validation: Training set: $\{(x_1, y_1), \dots, (x_{n_i}, y_{n_i})\}$ Validation set: $\{(x_{n_i+1}, y_{n_i+1}), \dots, (x_{n_{i+1}}, y_{n_{i+1}})\}$ where $n_i = [k + 1i \cdot n]$ for fold $i$	The applied time-series cross-validation framework ensures that the used cryptocurrency models are tested under realistic conditions, giving confidence that at 75% accuracy in backtesting might actually translate to 75% accuracy in live trading—rather than the devastating surprise of 25% accuracy to be represented as “95% accurate” regular-CV model meets the real world.
Cross-Validation Score	Non	$CVscore = \frac{1}{k} \sum_{i=1}^k R_i^2$ where $R_i^2 = 1 - \frac{\sum_{j \in fold_i} (y_j - \bar{y}_i)^2}{\sum_{j \in fold_i} (y_j - \bar{y}_i)^2}$	
Model Ensemble Formulation	Ensemble	$\hat{y} = \sum K \eta f_k(x) = k = 1 \sum K \eta j = 1 \sum J_k w_k j I(x \in R_k, j)$ where K is the number of boosting rounds $J_k$ is the number of leaves in tree k $w_{k,j}$ is the weight of leaf j in tree k $R_{k,j}$ is the region corresponding to leaf j in tree k $I(\cdot)$ is the indicator function	Combining XGBoost + LSTM + GARCH-DL predictions
Jarque–Bera Normality Test	Non	$JB = \frac{n}{6} \left( S^2 + \frac{(K-3)^2}{4} \right)$ where $S$ is skewness and $K$ is kurtosis of residuals $S = (1/n) \times \sum [(x_i - \bar{x})/\sigma]^3$ $S = 0$ : Perfectly symmetric (like normal distribution) $S > 0$ : Right-skewed (positive skew)—long tail on right $S < 0$ : Left-skewed (negative skew)—long tail on left $K = 3$ : Normal distribution (mesokurtic) $K > 3$ : Fat tails (leptokurtic)—more extreme events $K < 3$ : Thin tails (platykurtic)—fewer extreme events	Jarque–Bera Normality Test is a crucial statistical test for validating cryptocurrency models. Based on whether the data (like returns or residuals) follows a normal distribution—which has major implications for risk management and model reliability. In our case left skewness, leptokurtic with fat tail.

**Table A1.** Cont.

Metric Model Function Indicator Parameter	Abbreviation	Equation	Description
Ljung–Box Autocorrelation Test	Non	$Q_{LB} = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k}$ where $\hat{\rho}_k$ is the sample autocorrelation at lag k.	This Ljung–Box framework can provide insights into whether the cryptocurrency models are missing predictable patterns or if the trading markets actually have exploitable inefficiencies—crucial intelligence for building profitable trading strategies. In our case Ljung–Box test (returns, lags 1–10): Significant autocorrelation found at lags: [1, 2, 3, 4, 6, 7, 8, 9, 10].
Directional Accuracy	Non	$DA = \frac{1}{n-1} \sum_{t=2}^n I(\text{sign}(\Delta y_t) = \text{sign}(\hat{y}_t))$ where $I(\cdot)$ is the indicator function and $\Delta y_t = y_t - y_{t-1}$	Directional Accuracy is one of the most practical and intuitive metrics for cryptocurrency prediction models. While $R^2$ indicates prediction precision, Directional Accuracy corresponds with trading profitability—which is often more important.
Root Mean Square Error	RMSE	$\sqrt{\frac{\sum_{t=1}^n (Q'_t - Q_t)^2}{n}}$	RMSE is the standard deviation of errors. Lower RMSE value indicates better fit.
Mean Absolute Error	MAE	$\frac{\sum_{t=1}^n  Q_t - Q'_t }{n}$	MAE is the absolute difference between the measured and simulated values. Lower MAE values indicate lower error.
Coefficient of Determination	$R^2$	$\left[ \frac{\sum_{t=1}^n (Q'_t - \bar{Q}') (Q_t - \bar{Q})}{\sqrt{\sum_{t=1}^n (Q'_t - \bar{Q}')^2} \sqrt{\sum_{t=1}^n (Q_t - \bar{Q})^2}} \right]^2$	$R^2$ provides the intensity of the relationship between measured and simulated values. Its value ranges from 0 to 1, closer to 0 shows lower correlation, while close to 1 represents high correlation.
Mean Square Error	MSE	$\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$	MSE represents the difference between original and predicted values, extracted by squaring the mean difference in the dataset.
Mean Absolute Percentage Error	MAPE	$MAPE = \frac{1}{m} \sum_{t=1}^m \left  \frac{y(t) - \hat{y}(t)}{y(t)} \right $	This is an indicator used to evaluate the accuracy of a forecasting model. MAPE measures the average percentage difference between predicted and actual values. MAPE is often expressed as a percentage. This is particularly useful when it is necessary to understand the accuracy of a model relative to the scale of the data being measured.

## Appendix B

### Enhanced LSTM Architecture with Attention Mechanism



## References

- Alizadegan, Hamed, Arian Radmehr, and Mohsen Asghari Ilani. 2024. *Forecasting Bitcoin Prices: A Comparative Study of Machine Learning and Deep Learning Algorithms*. Durham: Research Square. [CrossRef]
- Bourday, Rachid, Issam Aatouchi, Mounir Ait Kerroum, and Ali Zaaouat. 2024. Cryptocurrency Forecasting Using Deep Learning Models: A Comparative Analysis. *HighTech and Innovation Journal* 5: 4. [CrossRef]
- Bouteska, Ahmed, Mohammad Zoynul Abedin, Petr Hajek, and Kunpeng Yuan. 2024. Cryptocurrency price forecasting—A comparative analysis of ensemble learning and deep learning methods. *International Review of Financial Analysis* 92: 103055. [CrossRef]
- Chen, SiXian, Zonghu Liao, and Jingbo Zhang. 2024. Forecasting the Price of Bitcoin Using an Explainable CNN-LSTM Model. In *Intelligent Computers, Algorithms, and Applications*. Edited by Christophe Cruz, Yanchun Zhang and Wanling Gao. Berlin: Springer Nature, pp. 93–101. [CrossRef]
- Chen, Tianqi, and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. Paper presented at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13–17; pp. 785–794. [CrossRef]

- Chu, Jeffrey, Stephen Chan, Saralees Nadarajah, and Joerg Osterrieder. 2017. GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management* 10: 17. [CrossRef]
- Da Cunha, Carlo R., and Roberto Moreira Da Silva. 2020. Relevant stylized facts about bitcoin: Fluctuations, first return probability, and natural phenomena. *Physica A: Statistical Mechanics and Its Applications* 550: 124155. [CrossRef]
- Dutta, Aniruddha, Saket Kumar, and Meheli Basu. 2020. A Gated Recurrent Unit Approach to Bitcoin Price Prediction. *Journal of Risk and Financial Management* 13: 23. [CrossRef]
- Gkillas, Konstantinos, and Paraskevi Katsiampa. 2018. An application of extreme value theory to cryptocurrencies. *Economics Letters* 164: 109–11. [CrossRef]
- Hafid, Abdelatif, Maad Ebrahim, Ali Alfatemi, Mohamed Rahouti, and Diogo Oliveira. 2024. Cryptocurrency Price Forecasting Using XGBoost Regressor and Technical Indicators. *arXiv* arXiv:2407.11786. [CrossRef]
- Iltaç, Yuksel, Gulbahar Ucler, and Umit Bulut. 2019. Assessment of the Crypto Market Efficiency: Empirical Evidence from Unit Root Tests with Different Approximations. In *Contributions to Economics*. New York: Springer International Publishing, pp. 191–200. [CrossRef]
- Jabbar, Abdul, and Syed Qaisar Jalil. 2024. A Comprehensive Analysis of Machine Learning Models for Algorithmic Trading of Bitcoin. *arXiv* arXiv:2407.18334. Available online: <https://arxiv.org/html/2407.18334v1> (accessed on 1 August 2025).
- Ji, Suhwan, Jongmin Kim, and Hyeonseung Im. 2019. A Comparative Study of Bitcoin Price Prediction Using Deep Learning. *Mathematics* 7: 898. [CrossRef]
- John, David L., Sebastian Binnewies, and Bela Stantic. 2024. Cryptocurrency Price Prediction Algorithms: A Survey and Future Directions. *Forecasting* 6: 637–71. [CrossRef]
- Kehinde, T. O., Oluyinka J. Adedokun, Akpan Joseph, Kareem Morenikeji Kabirat, Hammed Adebayo Akano, and Oludolapo A. Olanrewaju. 2025. Helformer: An attention-based deep learning model for cryptocurrency price forecasting. *Journal of Big Data* 12: 81. [CrossRef]
- Kervancı, I. Sibel, M. Fatih Akay, and Eren Özceylan. 2024. Bitcoin price prediction using LSTM, GRU and hybrid LSTM-GRU with bayesian optimization, random search, and grid search for the next days. *Journal of Industrial and Management Optimization* 20: 570–88. [CrossRef]
- Khaniki, Mohammad Ali Labbaf, and Mohammad Manthouri. 2024. Enhancing Price Prediction in Cryptocurrency Using Transformer Neural Network and Technical Indicators. *arXiv* arXiv:2403.03606. [CrossRef]
- Kiranmai Balijepalli, N. S. S., and Viswanathan Thangaraj. 2025. Prediction of cryptocurrency's price using ensemble machine learning algorithms. *European Journal of Management and Business Economics*. [CrossRef]
- Kjærland, Frode, Aras Khazal, Erlend A. Krogstad, Frans B. G. Nordstrøm, and Are Oust. 2018. An Analysis of Bitcoin's Price Dynamics. *Journal of Risk and Financial Management* 11: 63. [CrossRef]
- Ko, Po-Chang, Ping-Chen Lin, Hoang-Thu Do, Yuan-Heng Kuo, You-Fu Huang, and Wen-Hsien Chen. 2023. Pairs Trading Strategies in Cryptocurrency Markets: A Comparative Study between Statistical Methods and Evolutionary Algorithms. *Engineering Proceedings* 38: 74. [CrossRef]
- Kukacka, Jiri, and Ladislav Kristoufek. 2023. Fundamental and speculative components of the cryptocurrency pricing dynamics. *Financial Innovation* 9: 61. [CrossRef]
- Lamothe-Fernández, Prosper, David Alaminos, Prosper Lamothe-López, and Manuel A. Fernández-Gámez. 2020. Deep Learning Methods for Modeling Bitcoin Price. *Mathematics* 8: 1245. [CrossRef]
- Lee, Ming-Che. 2024. Bitcoin Trend Prediction with Attention-Based Deep Learning Models and Technical Indicators. *Systems* 12: 498. [CrossRef]
- Li, Yan, and Wei Dai. 2020. Bitcoin Price Forecasting Method Based on CNN-LSTM Hybrid Neural Network Model. *The Journal of Engineering* 2020: 344–47. [CrossRef]
- Livieris, Ioannis E., Niki Kiriakidou, Stavros Stavroyiannis, and Panagiotis Pintelas. 2021. An Advanced CNN-LSTM Model for Cryptocurrency Forecasting. *Electronics* 10: 287. [CrossRef]
- Mudassir, Mohammed, Shada Bennbaia, Devrim Unal, and Mohammad Hammoudeh. 2020. Time-series forecasting of Bitcoin prices using high-dimensional features: A machine learning approach. *Neural Computing and Applications*. [CrossRef]
- Murray, Kate, Andrea Rossi, Diego Carraro, and Andrea Visentin. 2023. On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles. *Forecasting* 5: 196–209. [CrossRef]
- Omole, Oluwadamilare, and David Enke. 2024. Deep learning for Bitcoin price direction prediction: Models and trading strategies empirically compared. *Financial Innovation* 10: 117. [CrossRef]
- Park, Jaehyun, and Yeong-Seok Seo. 2022. A Deep Learning-Based Action Recommendation Model for Cryptocurrency Profit Maximization. *Electronics* 11: 1466. [CrossRef]
- Petropoulos, Fotios, Daniele Apiletti, Vassilios Assimakopoulos, Mohamed Zied Babai, Devon K. Barrow, Souhaib Ben Taieb, Christoph Bergmeir, Ricardo J. Bessa, Jakub Bijak, John E. Boylan, and et al. 2022. Forecasting: Theory and practice. *International Journal of Forecasting* 38: 705–871. [CrossRef]

- Phung Duy, Quang, Oanh Nguyen Thi, Phuong Hao Le Thi, Hai Duong Pham Hoang, Khanh Linh Luong, and Kim Ngan Nguyen Thi. 2024. Estimating and forecasting bitcoin daily prices using ARIMA-GARCH models. *Business Analyst Journal* 45: 11–23. [CrossRef]
- Ranjan, Sumit, Parthajit Kayal, and Malvika Saraf. 2023. Bitcoin Price Prediction: A Machine Learning Sample Dimension Approach. *Computational Economics* 61: 1617–36. [CrossRef]
- Seabe, Phumudzo Lloyd, Claude Rodrigue Bambe Moutsinga, and Edson Pindza. 2023. Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach. *Fractal and Fractional* 7: 203. [CrossRef]
- Shen, Ying, Zhengxin Qi, Fernando Martinez, Mohamed Rahouti, D. Frank Hsu, and Diogo Oliveira. 2024. Comparative Analysis of Machine Learning and Traditional Forecasting Methods in Bitcoin Price Prediction. Paper presented at the 2024 6th International Conference on Blockchain Computing and Applications (BCCA), Dubai, United Arab Emirates, November 26–29; pp. 8–15. [CrossRef]
- Şimşek Türker, Yasemin, Şemsettin Kilinçarslan, and Ebru Yilmaz Ince. 2025. Performance of ANN, Random Forest and XGBoost methods in predicting the flexural properties of wood beams reinforced with carbon-FRP. *Wood Material Science & Engineering* 20: 657–68. [CrossRef]
- Taleb, Nassim Nicholas. 2020. Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications (Version 3). *arXiv arXiv:2001.10488*. [CrossRef]
- Visharad, I., Parthajit Kayal, and Moinak Maiti. 2025. Data-driven insights from technical indicators to predict the stablecoins price. *Finance Research Open* 1: 100011. [CrossRef]
- Wang, Jikai, Kai Feng, and Gaoxiu Qiao. 2024. A hybrid deep learning model for Bitcoin price prediction: Data decomposition and feature selection. *Applied Economics* 56: 6890–905. [CrossRef]
- Ye, Zi, Yinxu Wu, Hui Chen, Yi Pan, and Qingshan Jiang. 2022. A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin. *Mathematics* 10: 1307. [CrossRef]
- Yi, Eojin, Biao Yang, Minhyuk Jeong, Sungbin Sohn, and Kwangwon Ahn. 2023. Market efficiency of cryptocurrency: Evidence from the Bitcoin market. *Scientific Reports* 13: 4789. [CrossRef]
- Yıldırım, Hakan, and Festus Victor Bekun. 2023. Predicting volatility of bitcoin returns with ARCH, GARCH and EGARCH models. *Future Business Journal* 9: 75. [CrossRef]
- Zahid, Mamoon, Farhat Iqbal, and Dimitrios Koutmos. 2022. Forecasting Bitcoin Volatility Using Hybrid GARCH Models with Machine Learning. *Risks* 10: 237. [CrossRef]
- Zheng, Harry. 2020. A Systematic Review of Forecasting the Cryptocurrency Volatility Using GARCH Model. *Coinmonks*. Available online: <https://medium.com/coinmonks/a-systematic-review-of-forecasting-the-cryptocurrency-volatility-using-garch-model-d1855c9779c> (accessed on 1 August 2025).
- Zhu, Yingjie, Jiageng Ma, Fangqing Gu, Jie Wang, Zhijuan Li, Youyao Zhang, Jian Xu, Yifan Li, Yiwen Wang, and Xiangqun Yang. 2023. Price Prediction of Bitcoin Based on Adaptive Feature Selection and Model Optimization. *Mathematics* 11: 1335. [CrossRef]
- Zivot, Eric, and Donald W. K Andrews. 2002. Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics* 20: 25–44. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.