**IAES International Journal of Artificial Intelligence (IJ-AI)**
Vol. 14, No. 3, June 2025, pp. 1910~1918
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i3.pp1910-1918 

# Averaged bars for cryptocurrency price forecasting across different horizons

**Ahmed El Youssefi[1], Abdelaaziz Hessane[1,2], Imad Zeroual[1], Yousef Farhaoui[1]**
[1]IMIA Laboratory, T-IDMS, Faculty of Sciences and Techniques of Errachidia, Moulay Ismail University of Meknès, Meknes, Morocco
[2]Department of Computer Science, Faculty of Science of Meknès, Moulay Ismail University of Meknès, Meknes, Morocco

**Article Info**

***Article history:***

Received Apr 1, 2024
Revised Nov 18, 2024
Accepted Nov 24, 2024

***Keywords:***

Averaged bars
Boosting algorithms
Charting techniques
Cryptocurrency price forecasting
Heikin-ashi candlesticks

**ABSTRACT**

Technical analysis uses past price movements and patterns to predict future trends and help traders make informed decisions about their cryptocurrency portfolios. This study investigates the effectiveness of different forecasting algorithms and features in predicting the future log return of cryptocurrency close price across various horizons. Specifically, we compare the performance of AdaBoost, light gradient boosting machine (LightGBM), random forest (RF), and k-nearest neighbor (KNN) regressors using Kline open, high, low, close (OHLC) prices data and averaged bars (Heikin-Ashi) features. Our analysis covers ten of the most capitalized cryptocurrencies: Cardano, Avalanche, Binance Coin, Bitcoin, Dogecoin, Polkadot, Ethereum, Solana, Tron, and Ripple. We have observed nuanced patterns in predictive performance across different cryptocurrencies, forecasting horizons and features. Then we have found that AdaBoost and RF models consistently exhibit a competitive performance, with LightGBM showing promising results for specific cryptocurrencies. The impact of forecast horizons on forecasting performance underscores the need for tailored forecasting models. In summary, the use of Kline OHLC data as features outperforms averaged bars in forecasting the first and second horizons, while averaged bars outperform Kline OHLC data for mid- to relatively long-term horizons (starting from the third horizon). Our findings suggest that averaged bars merit more attention from researchers instead of relying solely on Kline OHLC data.



***Corresponding Author:***

Ahmed El Youssefi
IMIA Laboratory, T-IDMS, Faculty of Sciences and Techniques of Errachidia
Moulay Ismail University of Meknès
Meknes, Morocco
Email: ah.elyoussefi@edu.umi.ac.ma

## 1. INTRODUCTION

Cryptocurrencies, known for their high volatility, offer traders lucrative returns and high risks based on their decisions' accuracy [1]. Traders commonly decide whether to go long or short on a cryptocurrency based on their forecasting of the future trend of the cryptocurrency they hold or to which they want to take a position. Forecasting cryptocurrency trends poses a significant challenge that prompts traders to employ various techniques, including technical analysis, fundamental analysis, statistical methods, and artificial intelligence-based approaches.

Technical analysis seeks to predict future market behavior through pattern recognition and algorithmic trading strategies. This approach leverages technical indicators and chart formations levels [2], [3], derived from historical price and volume data (e.g., Kline open, high, low, close (OHLC) values) within a defined timeframe. These insights inform cryptocurrency portfolio management decisions.

Alternatively, fundamental analysis constitutes an asset valuation methodology [4] centered on determining the intrinsic value of a cryptocurrency. This approach involves a comprehensive evaluation of the foundational components driving a particular coin project, encompassing factors such as its underlying blockchain technology and upcoming project-related events (such as partnerships, halving events, or introducing a new consensus algorithm) [5]. Additionally, fundamental analysis considers the adoption rate of the project's services, updates from the project team, and other pertinent project-related information. Using these insights, fundamentalists forecast the trend of a cryptocurrency's price. However, applying fundamental analysis to cryptocurrency price forecasting remains a complex task due to the significant influence of speculative behavior, market sentiment [6], and other factors such as the high market volatility [7].

Machine learning algorithms have also garnered significant attention as viable approaches for developing cryptocurrency forecasting models. By leveraging these algorithms, researchers and traders aim to capitalize on their ability to process large datasets and uncover trends and patterns that could signal forthcoming price fluctuations. Typically, these algorithms utilize historical Kline OHLC data, sourced directly from exchange archives or aggregated from diverse platforms, to train models for predicting future cryptocurrency prices over single or multiple time steps (horizons). There are two main approaches to cryptocurrency forecasting: classification and regression). Classification approaches predict price trends by learning from labeled data (e.g., trend direction based on Kline OHLC) and incorporating additional features [8]. Meanwhile, regression models utilize historical price data to directly forecast future cryptocurrency values within a specified timeframe.

Statistical approaches are also employed to forecast cryptocurrencies. These methods include autoregressive integrated moving average (ARIMA) models for time series analysis [9] and other statistical techniques that capture patterns and trends in historical price data. Statistical models provide a quantitative framework for understanding and predicting cryptocurrency price movements. However, they also need help with challenges, including assumptions about stationarity and the difficulty of capturing the complex dynamics of the cryptocurrency market. Traders and analysts often combine multiple methods, drawing on each other's strengths to enhance the robustness of their forecasts [10].

Our study aims to compare the effectiveness of using averaged bars and Kline OHLC data as features for machine learning algorithms to forecast close price log returns across various cryptocurrencies and horizons. While previous literature primarily focuses on one-step future forecasting with fewer studies utilizing averaged bars, our research extends beyond these limitations. By targeting different horizons and close prices of ten cryptocurrencies, we seek to address several key questions that have not been thoroughly explored:

– Which charting technique exhibits greater efficacy when forecasting deeper horizons into the future and using which machine learning algorithm?
– Is there consistent behavior among the different algorithms considered in this study when applied across different cryptocurrencies?

The novelty of our research lies in its comprehensive approach, as we extend beyond prior studies by examining multiple cryptocurrencies and exploring a more extensive range of forecasting horizons. Additionally, unlike our previous research [5] which focused primarily on Bitcoin, one horizon, and different time sampling windows using Kline OHLC data and Heikin-Ashi, this study broadens the scope to include additional cryptocurrencies and more horizons. Through this comparative analysis, we aim to provide insights into the predictive capabilities of these charting techniques, ultimately contributing to a deeper understanding of cryptocurrency market dynamics.

The subsequent sections of this paper are organized as follows: The second section will introduce core formulas for calculating averaged bars and provide an overview of recent studies about the application of Heikin-Ashi candlesticks in forecasting stock and cryptocurrency prices. Following that, the third section will describe the procedures for data collection, preprocessing, and the methodologies employed. In the fourth section, we will delve into the presentation and discussion of our results. The concluding fifth section will summarize the paper and outline potential directions for future research endeavors.

## 2. BACKGROUND AND RELATED WORKS

Kline candlesticks are extensively used in research related to stocks and cryptocurrency prices [11]–[14]. However, the averaged bars known as Heikin-Ashi candlesticks need more attention from the scientific community. Averaged bars candlesticks exhibit smoother patterns compared to Kline-based candlesticks due to their utilization of an average of the preceding candlesticks' open and close prices to determine the open price. Within this paper, averaged bars and Heikin-Ashi will be used interchangeably. Heikin-Ashi candlesticks are derived directly from Kline OHLC data. The formulas [15] to calculate their OHLC (HA_Open, HA_High, HA_Low, HA_Close) data are as shown in (1) to (7):

$$\text{HA\_Open} = \frac{\text{previous\_HA\_Open} + \text{previous\_HA\_Close}}{2} \quad (1)$$

$$\text{HA\_High} = \max(\text{KLine\_}High\,, \text{HA\_Open}, \text{HA\_Close}) \quad (2)$$

$$\text{HA\_Low} = \min(KLine\_Low, \text{HA\_Open}, \text{HA\_Close}) \quad (3)$$

$$\text{HA\_}Close = \frac{KLine\_Open + KLine\_High + KLine\_Low + KLine\_Close}{4} \quad (4)$$

Each visualized Heikin-Ashi candlestick has a body and two thin shadows (upper and lower).

$$HA\_Body = HA\_Close - HA\_Open \quad (5)$$

$$HA\_UpperShadow = HA\_High - max(HA\_Open,\ HA\_Close) \quad (6)$$

$$\text{HA\_LowerShadow} = \min(HA\_Open,\ HA\_Close) - HA\text{\_Low} \quad (7)$$

Applying Heikin-Ashi candlesticks in analyzing cryptocurrency and stock markets has received relatively limited attention in the existing literature. Shalini *et al*. [16] employs the average directional index (ADX) on Heikin-Ashi with other technical indicators. The backtesting process focuses on providing entry and exit points for stock market participants. Their analysis reveals that the generated technical indicator using Heikin-Ashi was one of the effective indicators for most studied stocks. El Youssefi *et al.* [5] compared Heikin-Ashi and Japanese candlesticks, single-step future log return of the close price of Bitcoin over different time windows ranging from 1 day to 5 minutes, and employing various regression algorithms including k-nearest neighbor (KNN) regressor, linear regression, light gradient boosting machine (LightGBM), the Huber regressor, and random forest (RF) regressor. Their key findings suggest that using OHLC candlesticks consistently outperforms Heikin-Ashi candlesticks across all considered periods. Piasecki and Hanćkowiak [17] used fuzzy numbers to represent the Heikin-Ashi transformation, accounting for the inherent uncertainty in historical price data. They reported that despite introducing additional imprecision through their averaging approach, Heikin-Ashi candlesticks efficiently detect trends in volatile price data, resulting in notable forecasting accuracy. Madbouly *et al.* [18] integrated cloud models, fuzzy time series, and Heikin-Ashi candlesticks to predict and confirm stock trends to address nonlinearity and noise in stock market data, the model leveraged a model to cover the randomness gap in fuzzy logic, bridging qualitative and quantitative concepts. The model handled ambiguity and uncertainty in Japanese candlestick definitions and actual stock prices, constructing dynamic weighted fuzzy logical relationships for OHLC prices forecasting. El Youssefi *et al.* [19] used kMens clustering to categorize averaged bars candlesticks and logarithmic returns of prices, and to determin optimal class numbers for cryptocurrency logarithmic returns. The study establishes the importance of clustering in feature preprocessing for effective classification in cryptocurrency forecasting.

## 3. METHODS AND MATERIALS

Our study began by constructing our dataset using historical data from Binance, focusing on ten selected cryptocurrencies. As depicted in Figure 1, we obtained Kline OHLC data from Binance. Next, we generated various features including Kline OHLC candlesticks and averaged bars candlesticks. Additionally, we calculated different horizon targets, specifically the logarithmic returns from one to ten horizons. In the data preprocessing stage, we addressed missing values using simple mean-based imputation. We then applied a time series split strategy with ten folds to divide each dataset into a training split (70% of the data) and a test split (30% of the data). The data were normalized using the robust scaler. Finally, we applied our selected regressors to the resulting datasets. The performance of these regressors was evaluated using the coefficient of determination, R-squared ($R^2$), to assess the accuracy of our forecasts.

### 3.1. Data collection and preprocessing

Based on market capitalization available at [20], we have used the historical data provided by Binance of the ten most capitalized cryptocurrencies as of January 7, 2024, at 16:23 GMT. Another criterion is that the cryptocurrencies should be tradable in the USD Tether (USDT) market. The chosen cryptocurrencies alphabetically ordered are Cardano (ADA), Avalanche (AVAX), Binance coin (BNB), Bitcoin (BTC), Doge (DOGE), Polkadot (DOT), Ethereum (ETH), Solana (SOL), Tron (TRX), and Ripple (XRP).

We downloaded all the Kline OHLC data from their start date of trading on Binance until November 30, 2023, for each cryptocurrency. These data include the OHLC prices, the volume, and the number of trades. Heikin-Ashi OHLC and the candlesticks' characteristics (body, upper shadow, and lower shadow) for Kline and Heikin-Ashi have been then calculated and appended to the data. Then, we calculated logarithmic returns from one to ten horizons for each datapoint of data. Logarithmic return is a transformation widely used in stock and cryptocurrency regression-based forecasting using machine learning algorithms. To calculate the logarithmic return for the n$^{th}$ horizon, using the close price of a cryptocurrency, the (8) is used:

$$\text{Log\_return}_n = \ln(KLine\_Close(n)) - \ln(KLine\_Close(0)) \quad (8)$$

Where $KLine\_Close(0)$ is the current candlestick close price and $KLine\_Close(n)$ is the close price of the n$^{th}$ future Kline candlestick.

Leveraging the PyCaret automated machine learning library [21], we have applied a time series cross-validation strategy with ten folds. This approach prioritizes temporal fidelity by exclusively training models on past data and keeping future unseen data for unbiased evaluation. This aligns with the inherent structure of cryptocurrency time series, where predictions based on unavailable data are more applicable. Additionally, we have addressed missing values through mean imputation. The dataset was subsequently split into a 70/30 training-testing ratio for model training and performance assessment. Finally, robust scaling has been applied to mitigate the influence of outliers on the employed algorithms.

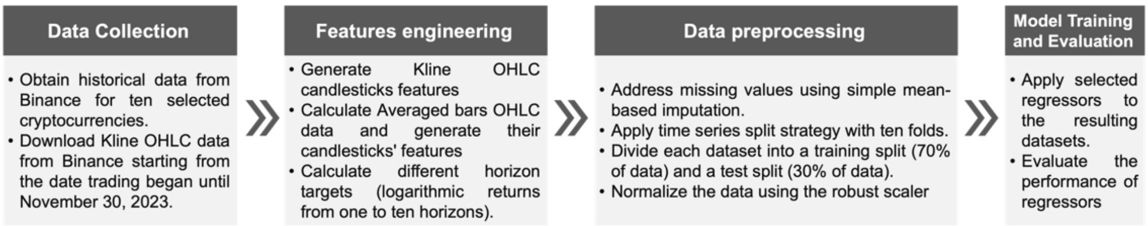


Figure 1. Workflow for data collection, feature engineering, data preprocessing, and model training and evaluation

### 3.2. Learning regression algorithms

We employed four regression models—AdaBoost, LightGBM, RF, and KNN regressors—to compare log returns across different forecasting horizons for various cryptocurrencies. Each of these models was implemented using the Scikit-learn library [22]. By utilizing diverse ensemble and non-parametric techniques, we aimed to capture different aspects of the underlying data patterns and assess their predictive capabilities across multiple time horizons.

#### 3.2.1. AdaBoost regressor

AdaBoost employs a boosting approach. It iteratively trains a sequence of weak learners (models with slightly better accuracy than random guessing, like simple decision trees). These weak learners are trained on modified versions of the dataset. Their predictions are then combined using a weighted vote (sum for regression) to create the final prediction. At each boosting iteration, the data undergoes modifications where weights $w_1, w_2, \ldots, w_n$ are assigned to each training sample. Initially, these weights are uniformly set to $\frac{1}{n}$, allowing the first step to train a weak learner on the original data. AdaBoost adjusts sample weights dynamically. Samples misclassified by the previous boosted model gain higher weights, while correctly classified samples have their weights decreased. This iterative process focuses the learning algorithm on previously challenging samples in subsequent iterations [23].

#### 3.2.2. Light gradient boosting machine

LightGBM is a gradient boosting decision tree algorithm which use a forward stage-wise approach, initially building a base model to predict the target variable's mean. Subsequently, it iteratively refines this model by constructing decision trees that focus on the residuals (errors) between the actual values and the current predictions. These residuals are calculated as the negative gradient of the chosen loss function. The model is then updated by incorporating a portion of each new tree's predictions. This approach enhances model performance while maintaining computational efficiency [24].

### 3.2.3. Random forest regressor

RF is a machine learning algorithm that utilizes a multitude of decision trees for prediction. During training, it constructs a forest of unpruned decision trees, where each tree is built on a random subset of features and a random subset of training data. Predictions are made by averaging the predictions from all individual trees within the ensemble. This approach enhances model robustness and reduces overfitting compared to single decision trees.

### 3.2.4. K-nearest neighbor regressor

KNN is a non-parametric, instance-based learning method prevalent in both statistics and machine learning. Unlike parametric approaches, KNN avoids assumptions about data distribution and relies on the training data for predictions. It predicts continuous variables by considering the 'K' closest neighbors in the training set. The algorithm identifies these neighbors using a distance metric, e.g., Euclidean distance. The final prediction is then calculated as the average of the dependent variable values associated with these 'K' neighbors [5] as shown in (9).

$$\hat{y} = \frac{1}{K}\sum_{i=1}^{K} y_i \tag{9}$$

While the KNN regressor offers flexibility and simplicity and can effectively capture complex variable relationships in specific datasets, its performance relies heavily on factors such as the data's dimensionality and scaling, the chosen value of 'K' and distance metric.

### 3.3. Evaluation metrics

The coefficient of determination, or $R^2$, is a statistical metric that indicates how well a regression model fits the data. It measures the proportion of variation in the dependent variable (Y) that is explained by the independent variables (X) included in the model. Essentially, $R^2$ is the ratio of the explained sum of squares (ESS) to the total sum of squares (TSS) [25]. ESS reflects the variation captured by the model, while TSS represents the total variation in the dependent variable. The formula for calculating $R^2$ as shown in (10):

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_{i=1}^{n}\left(y_{p_i} - y_{a_i}\right)^2}{\sum_{i=1}^{n}\left(y_{p_i} - \bar{y}\right)^2} \tag{10}$$

A higher $R^2$ value indicates that a more significant proportion of the total variance in the dependent variable is accounted for by the independent variables in the regression model, suggesting a better fit of the model to the observed data. This study used $R^2$ as an evaluation metric for different cryptocurrency forecasting models. $R^2$ quantifies the proportion of variance in the dependent variable explained by the model, thereby offering a straightforward assessment of model performance. Notably, $R^2$ is scale-independent, facilitating comparisons across diverse datasets or scales. As $R^2$ was recommended in [26], we will report a graphical representation of $R^2$ regression values only.

## 4. RESULTS AND DISCUSSION

This study investigated the performance of four forecasting algorithms: AdaBoost, LightGBM, RF, and KNN regressors, using Kline OHLC and Heikin-Ashi (averaged bars) features to forecast the future log return of the close price over different horizons (H1 to H10). While earlier studies have explored various forecasting methods, they have not explicitly addressed the influence of feature types and forecast horizons on predictive performance for cryptocurrencies. Figures 2 to 11 present the $R^2$ values of the considered algorithms across various horizons for both feature types. Notably, all negative $R^2$ values were excluded to focus on meaningful positive data.

Our key findings reveal that the use of averaged bars generally leads to better $R^2$ values, outperforming Kline OHLC features starting from the 3rd horizon. This improvement is evident in cryptocurrencies such as Ethereum, Bitcoin, and Cardano. Specifically, AdaBoost and RF models consistently demonstrate high performance with averaged bars features. Conversely, LightGBM shows promise for specific cryptocurrencies, notably Bitcoin and Ethereum, while KNN models exhibit less consistent performance, with lower $R^2$ values across most cryptocurrencies and horizons except for Avalanche and Binance coin. Our study suggests that averaged bars offer better results for longer horizons in forecasting cryptocurrency prices, whereas Kline OHLC data are more effective for short-term horizons (1st to 3rd horizons), despite their generally lower $R^2$ values.

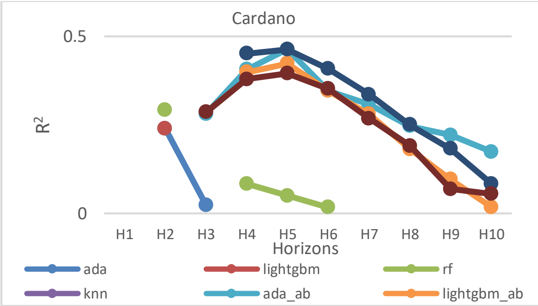


Figure 2. $R^2$ values for ADA/USDT pair close price log return forecasting

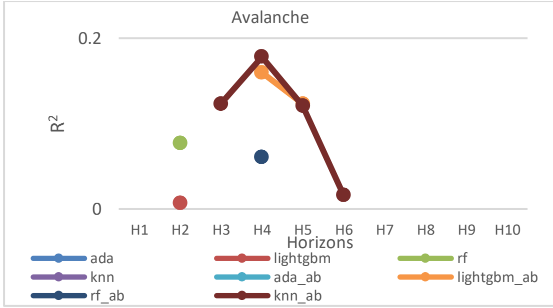


Figure 3. $R^2$ values for AVAX/USDT pair close price log return forecasting

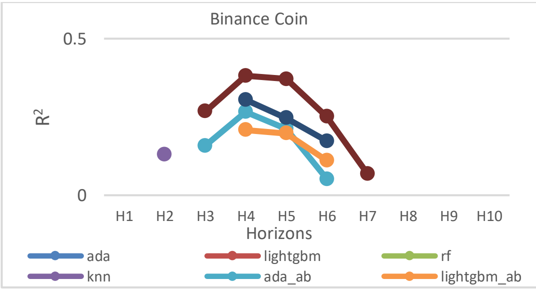


Figure 4. $R^2$ values for BNB/USDT pair close price log return forecasting

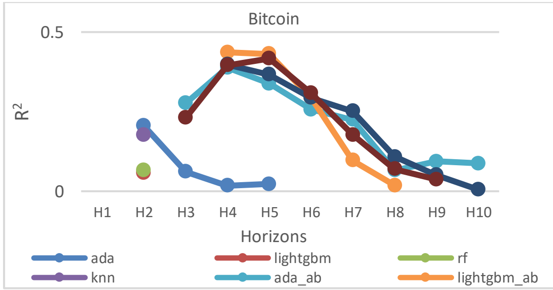


Figure 5. $R^2$ values for BTC/USDT pair close price log return forecasting

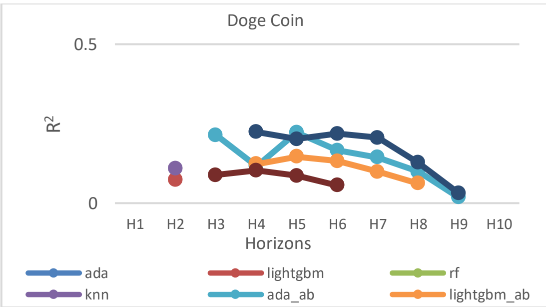


Figure 6. $R^2$ values for DOGE/USDT pair close price log return forecasting

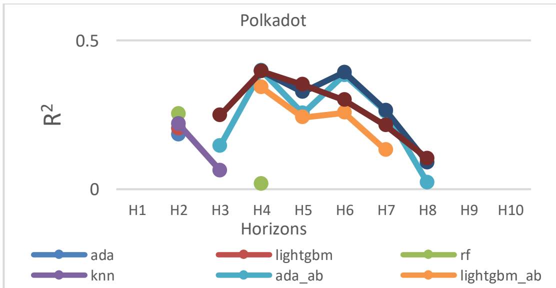


Figure 7. $R^2$ values for DOT/USDT pair close price log return forecasting

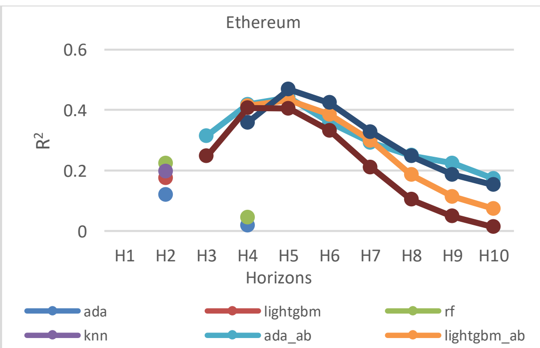


Figure 8. $R^2$ values for BTC/USDT pair close price log return forecasting

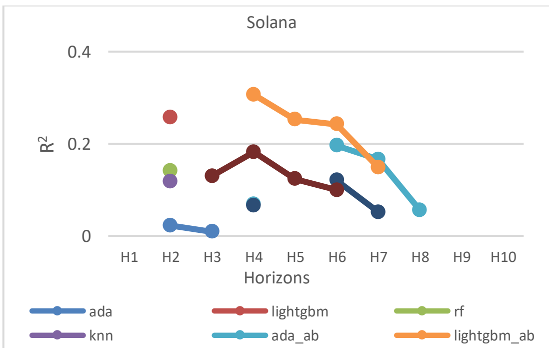


Figure 9. $R^2$ values for SOL/USDT pair close price log return forecasting

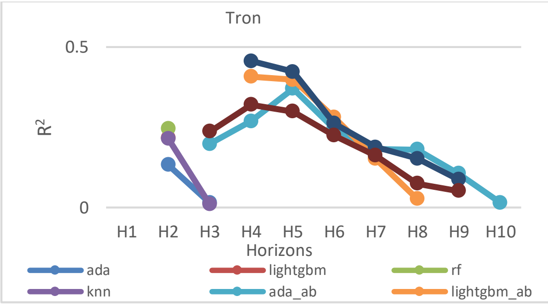


Figure 10. $R^2$ values for TRX/USDT pair close price log return forecasting

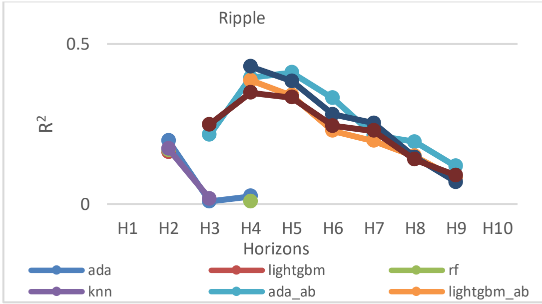


Figure 11. $R^2$ values for XRP/USDT pair close price log return forecasting

## 5. CONCLUSION

This study explored the performance of four forecasting algorithms: AdaBoost, LightGBM, RF, and KNN regressors. We used Kline OHLC data for one set of features and Heikin-ashi features for another. These regression algorithms were applied to forecast the future daily log return of the closing price over various horizons (H1 to H10). Key insights emerged from evaluating multiple cryptocurrencies and forecasting horizons. Firstly, feature selection is crucial for predictive accuracy. While Heikin-Ashi features yield better results for longer horizons in forecasting cryptocurrency prices, their superiority is not universal across all cryptocurrencies and forecast horizons. They perform best at the third horizon and then vary for longer horizons, indicating that feature effectiveness depends on the specific characteristics of each cryptocurrency's price behavior and the forecast horizon. Secondly, the performance of forecasting algorithms varies significantly. AdaBoost and RF models show competitive performance across multiple cryptocurrencies and horizons, while LightGBM produces promising results for specific cryptocurrencies and horizons, highlighting its efficacy in modeling nonlinear relationships. Lastly, the impact of forecast horizons on predictive performance emphasizes the importance of tailoring forecasting models to specific timeframes and market conditions. Short-term horizons may favor algorithms that adapt to rapid price fluctuations, whereas longer-term horizons may require models that identify sustained trends in cryptocurrency prices. Our findings emphasize the importance of incorporating averaged bars (Heikin-Ashi) and considering forecast horizons in cryptocurrency forecasting research. Further research is required to refine these methodologies and develop robust models that can capture the complexities of cryptocurrency price dynamics. It is important to note the limitations of our study, including the exclusive use of historical data from Binance and the focus on a specific set of cryptocurrencies and forecast horizons. To validate these results, future research should extend to a broader range of cryptocurrencies and explore different time samplings beyond the 1-day candlesticks used in this study. Additionally, future studies might investigate more sophisticated feature engineering methods and the application of alternative regression algorithms to further enhance predictive accuracy.

## FUNDING INFORMATION

This research wasn't funded by any grant.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ahmed EL Youssefi | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| Abdelaaziz Hessane | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| Imad Zeroual | ✓ | ✓ | | ✓ | | | | | ✓ | | | ✓ | | |
| Yousef Farhaoui | ✓ | ✓ | | ✓ | | | | | ✓ | | | ✓ | | |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

## CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

## DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

## REFERENCES

[1] B. Chen and Y. Sun, "Risk characteristics and connectedness in cryptocurrency markets: new evidence from a non-linear framework," *The North American Journal of Economics and Finance*, vol. 69, Jan. 2024, doi: 10.1016/j.najef.2023.102036.
[2] S. Corbet, V. Eraslan, B. Lucey, and A. Sensoy, "The effectiveness of technical trading rules in cryptocurrency markets," *Finance Research Letters*, vol. 31, pp. 32–37, Dec. 2019, doi: 10.1016/j.frl.2019.04.027.
[3] I. Lyukevich, I. Gorbatenko, and E. Bessonova, "Cryptocurrency market: choice of technical indicators in trading strategies of individual investors," *in 3rd International Scientific Conference on Innovations in Digital Economy*, Saint-Petersburg Russian Federation: ACM, Oct. 2021, pp. 408–416, doi: 10.1145/3527049.3527089.
[4] A. Thakkar and K. Chaudhari, "A comprehensive survey on portfolio optimization, stock price and trend prediction using particle swarm optimization," *Arch Computat Methods Engineering*, vol. 28, no. 4, pp. 2133–2164, 2021, doi: 10.1007/s11831-020-09448-8.
[5] A. El Youssefi, A. Hessane, I. Zeroual, and Y. Farhaoui, "Utilizing machine learning and deep learning for predicting crypto-currency trends," *Salud, Ciencia y Tecnología - Serie de Conferencias*, vol. 3, Mar. 2024, doi: 10.56294/sctconf2024638.
[6] G. Dudek, P. Fiszeder, P. Kubus, and W. Orzeszko, "Forecasting cryptocurrencies volatility using statistical and machine learning methods: a comparative study," *SSRN*, 2023, doi: 10.2139/ssrn.4409549.
[7] B. Y. Almansour, M. M. Alshater, and A. Y. Almansour, "Performance of ARCH and GARCH models in forecasting cryptocurrency market volatility," *Industrial Engineering & Management Systems,* vol. 20, no. 2, pp. 130–139, Jun. 2021, doi: 10.7232/iems.2021.20.2.130.
[8] E. Akyildirim, A. Goncu, and A. Sensoy, "Prediction of cryptocurrency returns using machine learning," *Annals of Operations Research*, vol. 297, no. 1–2, pp. 3–36, Feb. 2021, doi: 10.1007/s10479-020-03575-y.
[9] Z. Wang, Z. Yang, Z. Zheng, and Y. Zhu, "Bitcoin price forecasting based on arima model and multifactorial linear regression," *in 2023 International Conference on Networking, Informatics and Computing (ICNETIC),* Palermo, Italy: IEEE, May 2023, pp. 15–20, doi: 10.1109/ICNETIC59568.2023.00009.
[10] A. El Youssefi, A. Hessane, Y. Farhaoui, and I. Zeroual, "Cryptocurrency returns clustering using japanese candlesticks: towards a programmatic trading system," *in Advanced Technology for Smart Environment and Energy*, Cham: Springer International Publishing, 2023, pp. 93–103, doi: 10.1007/978-3-031-25662-2_8.
[11] R. K. Alkhodhairi, S. R. Aljalhami, N. K. Rusayni, J. F. Alshobaili, A. A. Al-Shargabi, and A. Alabdulatif, "Bitcoin candlestick prediction with deep neural networks based on real time data," *Computers, Materials & Continua*, vol. 68, no. 3, pp. 3215–3233, 2021, doi: 10.32604/cmc.2021.016881.
[12] M. Sheraz, S. Dedu, and V. Preda, "Volatility dynamics of non-linear volatile time series and analysis of information flow: evidence from cryptocurrency data," *Entropy*, vol. 24, no. 10, Oct. 2022, doi: 10.3390/e24101410.
[13] S. Simtharakao and D. Sutivong, "Exploring normalization techniques in neural networks for bitcoin candlestick price prediction," *in 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC*), Bali, Indonesia: IEEE, Feb. 2023, pp. 483–488, doi: 10.1109/ICAIIC57133.2023.10067086.
[14] S. Sivaprakkash and S. Vevek, "Price volatility in cryptocurrencies: a modelling approach," *in Advances in Finance, Accounting, and Economics,* IGI Global, 2023, pp. 29–43, doi: 10.4018/978-1-6684-5691-0.ch002.
[15] O. A. Hassen, S. M. Darwish, N. A. Abu, and Z. Z. Abidin, "Application of cloud model in qualitative forecasting for stock market trends," *Entropy,* vol. 22, no. 9, Sep. 2020, doi: 10.3390/e22090991.
[16] T. Shalini, S. Pranav, and S. Utkarsh, "Picking buy-sell signals: a practitioner's perspective on key technical indicators for selected indian firms," *Studies in Business and Economics*, vol. 14, no. 3, pp. 205–219, Dec. 2019, doi: 10.2478/sbe-2019-0054.
[17] K. Piasecki and A. Ł. -Hanćkowiak, "Heikin-ashi technique with use of oriented fuzzy numbers," *Uncertainty and Imprecision in Decision Making and Decision Support: New Advances, Challenges, and Perspectives*, Cham: Springer International Publishing, 2022, pp. 60–71, doi: 10.1007/978-3-030-95929-6_5.
[18] M. M. Madbouly, M. Elkholy, Y. M. Gharib, and S. M. Darwish, "Predicting stock market trends for japanese candlestick using cloud model," *in Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020*), Cham: Springer International Publishing, 2020, pp. 628–645, doi: 10.1007/978-3-030-44289-7_59.
[19] A. El Youssefi, A. Hessane, A. El Allaoui, I. Zeroual, and Y. Farhaoui, "Heikin ashi candlesticks for cryptocurrency returns clustering," *Artificial Intelligence and Smart Environment,* Cham: Springer International Publishing, 2023, pp. 481–485, doi: 10.1007/978-3-031-26254-8_69.
[20] "Yahoo Finance-stock market live, quotes, business and finance news," *Yahoo Finance.* Accessed: Mar. 22, 2024. [Online]. Available: https://finance.yahoo.com/
[21] "PyCaret: An open source, low-code machine learning library in Python," *PyCaret*, 2020. Accessed: Nov. 23, 2023. [Online]. Available: https://pycaret.org/
[22] F. Pedregosa *et al*., "Scikit-learn: machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[23] "1.11. Ensembles: Gradient boosting, random forests, bagging, voting, stacking-scikit-learn 1.4.0 documentation," *Scikit Learn*, Accessed: Jan. 25, 2024. [Online]. Available: https://scikit-learn.org/stable/modules/ensemble.html#adaboost
[24] G. Ke *et al*., "LightGBM: a highly efficient gradient boosting decision tree," *in Advances in Neural Information Processing Systems*, Curran Associates Inc., pp. 1-9, 2017.
[25] Y. Jiang, H. Li, G. Yang, C. Zhang, and K. Zhao, "Machine learning-driven ontological knowledge base for bridge corrosion evaluation," *IEEE Access*, vol. 11, pp. 144735–144746, 2023, doi: 10.1109/ACCESS.2023.3344320.
[26] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, Jul. 2021, doi: 10.7717/peerj-cs.623.

**BIOGRAPHIES OF AUTHORS**

**Ahmed El Youssefi** earned his Ph.D. in software engineering and artificial intelligence in 2025. He is a computer science inspector working at the Ministry of Education, Morocco, since 2020. His research focuses on technical analysis, optimized labeling of financial data and artificial intelligence applied to cryptocurrency price forecasting. He actively contributes to the scientific community as a reviewer and serves on the scientific committees of various international conferences. He can be contacted at email: ah.elyoussefi@edu.umi.ac.ma.

**Abdelaaziz Hessane** is an Assistant Professor at the Faculty of Sciences, Moulay Ismail University of Meknès, Morocco. He earned his Ph.D. in science and techniques in 2024 from the Faculty of Sciences and Techniques of Errachidia, specializing in computer science, software engineering, and artificial intelligence. He also holds an M.S. in business intelligence and image processing from the same institution (2020) and has experience teaching computer science at the high school level. His research focuses on artificial intelligence applications in precision agriculture. He actively contributes to the scientific community as a reviewer for several esteemed journals and serves on the scientific committees of various international conferences. He can be contacted at email: a.hessane@edu.umi.ac.ma.

**Imad Zeroual** is currently an Associate Professor in the Department of Computer Science, Faculty of Sciences and Technics, Moulay Ismail University. He received his Ph.D. in computer science from Mohamed First University in 2018. He is also a member of several international and national scientific communities such as the International Association for Educators and Researchers (IAER/171101), London, UK, the International Association of Engineers (IAENG/206013), and the Arabic Language Engineering Society (ALESM), Rabat, Morocco. His areas of research are artificial intelligence and data science. He primarily works on natural language processing, machine learning, information retrieval/extraction, and language teaching and learning. He can be contacted at email: mr.imadine@gmail.com.

**Yousef Farhaoui** is a Professor at Moulay Ismail University of Meknès, Faculty of Sciences and Techniques, Morocco. Local publishing and research coordinator, Cambridge International Academics in United Kingdom. He obtained his Ph.D. degree in computer security from Ibn Zohr University, Faculty of Science. His research interests include learning, e-learning, computer security, big data analytics, and business intelligence. He has three books in computer science. He is a coordinator and member of the organizing committee and a member of the scientific committee of several international congresses and is a member of various international associations. He has authored 10 books and many book chapters with reputed publishers such as Springer and IGI. He served as a reviewer for IEEE, IET, Springer, Inderscience, and Elsevier journals. He is also the guest editor of many journals with Wiley, Springer, and Inderscience. He has been the general chair, session chair, and panelist in several conferences. He is senior member of IEEE, IET, ACM, and EAI Research Group. He can be contacted at email: y.farhaoui@fste.umi.ac.ma.