



# Comparative Analysis of Machine Learning Algorithms for Daily Cryptocurrency Price Prediction



Timothy Kayode Samson\*

Statistics Programme, College of Agriculture, Engineering and Science, Bowen University, 540004 Iwo, Nigeria

\* Correspondence: Timothy Kayode Samson ([kayode.samson@bowen.edu.ng](mailto:kayode.samson@bowen.edu.ng))

**Received:** 01-27-2024

**Revised:** 02-28-2024

**Accepted:** 03-12-2024

**Citation:** T. K. Samson, "Comparative analysis of machine learning algorithms for daily cryptocurrency price prediction," *Inf. Dyn. Appl.*, vol. 3, no. 1, pp. 64–76, 2024. <https://doi.org/10.56578/ida030105>.



© 2024 by the author(s). Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

**Abstract:** The decentralised nature of cryptocurrency, coupled with its potential for significant financial returns, has elevated its status as a sought-after investment opportunity on a global scale. Nonetheless, the inherent unpredictability and volatility of the cryptocurrency market present considerable challenges for investors aiming to forecast price movements and secure profitable investments. In response to this challenge, the current investigation was conducted to assess the efficacy of three Machine Learning (ML) algorithms, namely, Gradient Boosting (GB), Random Forest (RF), and Bagging, in predicting the daily closing prices of six major cryptocurrencies, namely, Binance, Bitcoin, Ethereum, Solana, USD, and XRP. The study utilised historical price data spanning from January 1, 2015 to January 26, 2024 for Bitcoin, from January 1, 2018 to January 26, 2024 for Ethereum and XRP, from January 1, 2021 to January 26, 2024 for Solana, and from January 1, 2019 to January 26, 2024 for USD. A novel approach was adopted wherein the lagging prices of the cryptocurrencies were employed as features for prediction, as opposed to the conventional method of using opening, high, and low prices, which are not predictive in nature. The data set was divided into a training set (80%) and a testing set (20%) for the evaluation of the algorithms. The performance of these ML algorithms was systematically compared using a suite of metrics, including  $R^2$ , adjusted  $R^2$ , Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The findings revealed that the GB algorithm exhibited superior performance in predicting the prices of Bitcoin and Solana, whereas the RF algorithm demonstrated greater efficacy for Ethereum, USD, and XRP. This comparative analysis underscores the relative advantages of RF over GB and Bagging algorithms in the context of cryptocurrency price prediction. The outcomes of this study not only contribute to the existing body of knowledge on the application of ML algorithms in financial markets but also provide actionable insights for investors navigating the volatile cryptocurrency market.

**Keywords:** Cryptocurrencies; Machine learning; Gradient boosting; Random forest; Bagging

## 1 Introduction

Cryptocurrencies are digital currencies that employ cryptographic techniques based on blockchain technology. It is a decentralized digital currency that allows users to send and receive currency on a peer-to-peer network (Nakamoto [1]) using blockchain technology. The origin of cryptocurrencies and blockchain technology started in 2008 when pseudonymous Satoshi Nakamoto introduced Bitcoin and blockchain technology (a technology that underlines its peer-to-peer global payment system). This development ushered in a myriad of other cryptocurrencies. According to the CoinMarketCap report, the cryptocurrency market capitalization stands at \$1.1 trillion with approximately 22,932 cryptocurrencies. Among these 22,932 cryptocurrencies, Bitcoin has the highest market capitalization of 1,013,198,281,381, followed by Ethereum with a market capitalization of 358,599,912,591.

Cryptocurrencies now serve as a medium of exchange for daily payments, speculation, and payment rail for non-expensive cross-border money transfers and other non-monetary uses. Cryptocurrency is a digital medium of payment that crosses boundaries, though it is not regulated by the government. Farrell [2] observed that cryptocurrency, as a digital currency, was used as an instrument for making payments. Cryptocurrencies have been recognized globally in the economy, and they have begun to be used as speculative investment assets. Historically, the first transaction in cryptocurrency occurred on January 2, 2009, between Hal Finney and Nakamoto, which was done using Bitcoin. The use of cryptocurrency in transactions has spread to several countries around the world, as cryptocurrency exchanges

are found in many countries, including Singapore, Switzerland, Australia, the United States of America (USA), the Republic of Korea, and Nigeria, among other countries with over 425 million users in the world.

As posited by Lahmiri and Bekiros [3], there is an increased public interest in cryptocurrencies, because this market is considered by the public as a way of amassing wealth within a very short period of time. The strengths of these currencies over traditional currencies include a decentralized peer-to-peer system, high liquidity, high returns, anonymity, and lower transaction costs, among others. Despite these anticipated returns, cryptocurrencies exhibit higher volatility, marked by large jumps in prices and shocks, than traditional currencies, making them a very risky investment. For instance, the largest cryptocurrency, Bitcoin, was over \$64,000 in the first half of 2021, and by September 2, 2022, it had dropped to \$20331.28 (a 68.23% drop in value). This problem persists now. As of February 23, 2024, the value of Bitcoin stands at \$51319.50 which is also relatively low compared with its performance in the first half of 2021. Other cryptocurrencies have also suffered significant drops in prices, and as it stands now, the future of this market is only based on speculation as investors are still counting losses.

The use of GB, RF, and Bagging regression in predicting the price of financial series has gained popularity, probably because these approaches show some robustness against overfitting compared to the use of conventional regression algorithms. Derbentsev et al. [4] explored the use of these algorithms and found that RF regression performed better than other ensemble methods. Similarly, Farouk et al. [5] compared the performance of RF and boosting regression with other algorithms in predicting the price of Bitcoin, and found that the RF regression performed better. RF regression performed better than other ensemble methods. Similarly, Farouk et al. [5] used open, high, close, and low prices as features, but in this study, the features are past lags of the closing price. One of the major weaknesses of using open, high, close, and low prices as features is that they cannot be used in forecasting since these prices are not available ahead of time, but using the past lag values of closing prices helps in addressing this challenge. This study is very significant, especially to investors, intending investors, and practitioners in the crypto market, as a reliable prediction of the future prices of these cryptocurrencies could help in decision-making. Investing in cryptocurrency is very risky, and hence having a reliable forecast of prices could help suggest when to buy or sell these currencies, thereby minimizing the huge loss that can be incurred as a result of a poor investment decision. Therefore, this study leverages the use of ML algorithms to predict the daily closing price of six cryptocurrencies (Binance, Bitcoin, Ethereum, Solana, USD, and XRP).

## 2 Literature Review

Several studies have been carried out on the use of ML algorithms for predicting cryptocurrencies. Alshehri [6] made use of both classifications of regression machine models in predicting the returns of Bitcoin. The study considered logistic regression, Linear Discriminant Analysis (LDA), K-Nearest Neighbor (KNN), Decision Tree (DT), Gaussian Naive Bayes (NB), Support Vector Machine (SVM), RF, Light Gradient-Boosting Machine (LGBM) and eXtreme Gradient Boosting (XGBoost). The study found that the XGBoost regressor performed better than other ML algorithms in foretelling the return movement of Bitcoin. Shilpa et al. [7] explored the use of Long Short-Term Memory (LSTM), Multi-Layer Perception (MLP), and Recurrent Neural Network (RNN) in predicting cryptocurrency prices. Jaquart et al. [8] employed various ML models to predict the binary daily market movement of the 100 largest cryptocurrencies. The findings indicated that these models provided reliable predictions for these cryptocurrencies. These results indicated that there is a challenge to weak cryptocurrency market efficiency, although the influence of certain limits on arbitrage cannot be entirely ruled out.

Pan [9] compared the performance of the Autoregressive Integrated Moving Average Model (ARIMA), RF, and LSTM algorithms of deep learning in predicting the price of three cryptocurrencies (Bitcoin, Ether, and Dogecoin) between 2018 and 2022. The performance of these algorithms was evaluated using MSE, RMSE, MAE, and  $R^2$ . Basher and Sadorsky [10] employed RF and Bagging in forecasting Bitcoin price direction using interest rates, inflation, and market volatility. Findings showed that RFs predict Bitcoin and gold price directions with a higher degree of accuracy than logit models. Prediction accuracy for bagging and RFs was found to be between 75% and 80% for a five-day prediction. For 10-day to 20-day forecasts, bagging and RFs record accuracies greater than 85%.

Srivastava et al. [11] used the regression algorithm and Particle Swarm Optimization (PSO) with the XGBoost algorithm for the prediction of the prices of three cryptocurrencies (Bitcoin, Dogecoin, and Ethereum). Findings revealed that the proposed method gave lower RMSE, MAE, and MSE values compared to the existing system.

Yan [12] employed a combination of statistical models and ML algorithms, namely, precisely Linear Regression (LR), GB, and RF, in forecasting the high-frequency time series (Limit Order Book) of Bitcoin. The findings by Yan [12] established the superiority of the LR algorithm over RF and GB. Saad et al. [13] compared the performance of LR, RF, and GB in predicting the prices of Bitcoin and Ethereum and found that the LR algorithm performed best with 10% of the data while GB and RF performed best with 5% of the data. Turukmane et al. [14] examined the capabilities of LSTM and XGBoost in forecasting the value of Bitcoin and found that the LSTM, which is a deep learning algorithm, performed better than the XGBoost algorithm.

Sakran [15] evaluated the performance of various ML algorithms: LR, DT regression, RF regression, support

vector regression (SVR), GB regression, AdaBoost regression, extreme GB regression (XGBR), Light GB Machine (LGBM), KNN regression, ridge, and lasso. In addition, Sakran [15] considered two deep learning algorithms, i.e., Artificial Neural Network (ANN) and Convolutional Neural Network (CNN), to forecast the daily prices of Bitcoin. The predictive performances of these algorithms were evaluated using the RMSE, MAE, and correlation coefficient (R). The study found that CNN demonstrated the highest effectiveness in predicting Bitcoin prices.

Some of the major gaps identified in this study are that most of the reviewed studies focus more on cryptocurrencies, while this study considered another trader cryptocurrency in addition to Bitcoin. The modeling approach of this study is also very different from that of other review studies, as this study makes use of the previous lag values of the closing price as features in building the ML algorithms other than using other variables because the cryptocurrency series is time series data. A review of related studies has also shown conflicting findings, as some of the studies favoured RF regression while others indicated the superiority of other ensemble ML algorithms. The present realities in the crypto market also necessitated the need for a recent study on predicting the closing price of cryptocurrencies.

### 3 Methodology

#### 3.1 Data Source and Preprocessing

Data on the closing prices of Binance, Bitcoin, Ethereum, Solana, USD, and XRP were obtained from the Yahoo Finance website (www.yahoofinance.com). This study used data spanning from January 1, 2015 to January 26, 2024 for Bitcoin, from January 1, 2018 to January 26, 2024 for Ethereum and XRP, from January 1, 2021 to January 26, 2024 for Solana, and from January 1, 2019 to January 26, 2024 for USD. These five cryptocurrencies were selected given the fact that they are among the top ten most traded cryptocurrencies in the world. As part of the process, the data was preprocessed and checked for duplicates and missing values. The six-day previous closing price was used in predicting the present-day closing price. Therefore, this makes the former the feature and the latter the target variable. This is used mainly because of the nature of the cryptocurrency series, which is time-series-based and has some unique features of cryptocurrencies as it is unregulated. Therefore, the study believes that using the previous data to predict the present price produces a reliable forecast as it utilises inherent information for prediction. Data preprocessing is an integral process in ML projects. It is a process of transforming the data in a way that is suitable for the intended machine-learning techniques. As part of the data preprocessing, the data was also normalized.

#### 3.2 Data Normalization

The time-series data for cryptocurrencies is scaled to the same value without altering the variations in the price range. The StandardScaler in Python was used for this. The range [0, 1] is created from the data using StandardScaler.

$$P_{t- \text{normalize}} = \frac{P_t - \min(P_t)}{\max(P_t) - \min(P_t)} \quad (1)$$

$$P_t = P_{t- \text{normalize}} [\max(P_t) - \min(P_t)] + \min(P_t) \quad (2)$$

#### 3.3 ML Models Used

##### 3.3.1 GB regression

This is one of the ensemble-supervised ML frameworks that make use of multiple weak DTs. By refining the model's weights based on the errors of prior iterations, the GB approach aims to significantly reduce prediction errors and increase the model's accuracy while improving overall predictive performance. Typically, the GB regression trains each subsequent model sequentially to correct its predecessor. A schematic diagram of the gradient-boosted regression tree (GBRT) is presented in Figure 1.

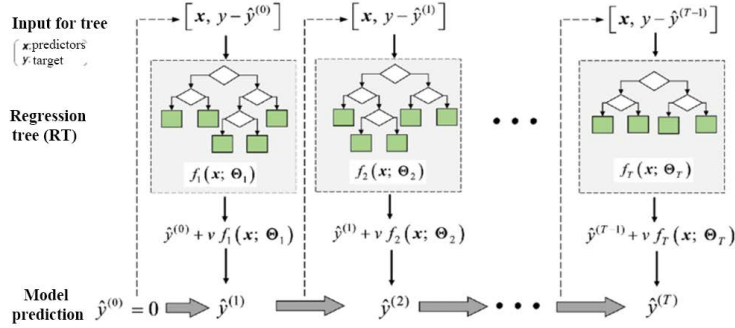
##### 3.3.2 Bagging regression

Combining bootstrap and aggregation results in bagging. This ML method trains many regression models via the bagging technique, and aggregates them to create a final model that is more reliable and accurate. The final predictions are obtained by averaging the estimates from base estimators.

##### 3.3.3 RF

A supervised learning approach called RF regression makes use of both bagging and boosting strategies. In RFs, the trees grow in parallel; therefore, there is no interaction between them as they grow. Since RFs perform well with high-dimensional data, missing values, and outliers, they are regarded as incredibly strong and powerful ML models. They also don't require a lot of hyperparameter tweaking and are comparatively simple to utilize. A RF is created in RF regression by building multiple trees in an arbitrary manner. Every tree is made from a distinct sample of rows,

and for every node, a distinct sample of characteristics is chosen for division. Every tree provides a forecast, which is then averaged to yield a single outcome. Because of the averaging, a RF performs better than a single DT, which enhances the accuracy of the prediction generated by RF regression. Several metrics were used in evaluating the performance of each ML algorithm, namely, RMSE, MAE, coefficient of determination, and adjusted coefficient of determination.



**Figure 1.** Schematic diagram of the GBRT

source: Wang et al. [16]

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_t - \hat{p}_t)^2}{n}} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_t - \hat{p}_t| \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (p_t - \hat{p}_t)^2}{\sum_{i=1}^n (p_t - \bar{p}_t)^2} \quad (5)$$

$$R_{adj.}^2 = 1 - \left[ \frac{(1 - R^2)(n - 1)}{(n - k - 1)} \right] \quad (6)$$

where,  $p_t$  is the actual closing price,  $\hat{p}_t$  is the predicted closing price,  $k$  is the number of parameters, and  $n$  is the number of observations. The data were split into a testing and validation set, with 80% of the data as the test set and 20% as the validation set. Also, from `sklearn.model_selection`, `train_test_split`, `cross_val_score` and `KFold` were imported. To improve the performance of these ML algorithms, hyperparameter tuning was carried out using the `GridSearch` algorithm. For each of the algorithms, the maximum depth was [5, 6], while the number of estimators was [300, 500, 900, 1000]. Using the `GridSearchCV` from `sklearn.model_selection`, one of the libraries in Python, the optimal hyperparameters were obtained. The `GridSearchCV` method searches for the best set of hyperparameters from a grid of hyperparameter values.

#### 4 Results and Discussion

The Results section may be divided into subsections. It should describe the results concisely and precisely, provide their interpretation, and draw possible conclusions from the results.

The result in Table 1 presents the summary descriptive statistics for the six selected cryptocurrencies. The minimum price for Binance was 4.189971, while for Bitcoin, Ethereum, Solana, USD, and XRP, it was 171.509995, 82.829887, 1.502038, 0.877400, and 0.115093, respectively. Among the six cryptocurrencies, Bitcoin reported the highest standard deviation (15925.273191), indicating that it had the highest risk level compared with other cryptocurrencies. The least standard deviation was obtained by USD (0.006155), indicating its price was more consistent than other cryptocurrencies. The Coefficient of Variation (COV) for both Binance (100.8949%) and Bitcoin (107.1229%) were both above 100%, indicating the standard deviation exceeded the mean, while the COV was less than 100% for other cryptocurrencies.

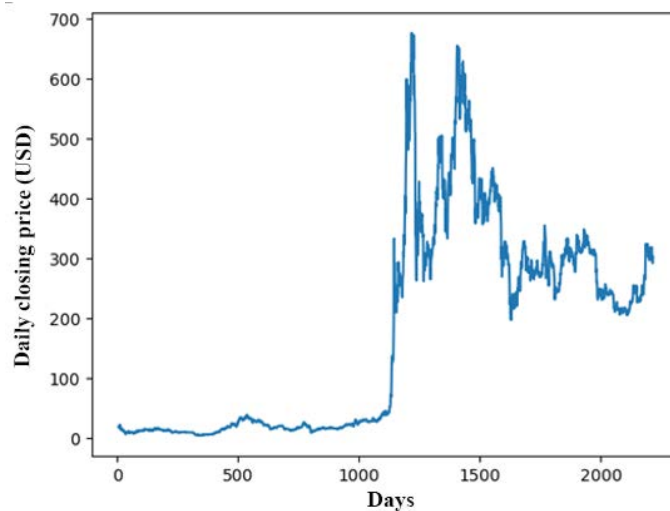
The COV obtained for XRP was lower than that of other cryptocurrencies, implying that the XRP price series was more homogenous than others. Bitcoin obtained the highest COV, indicating more variation than Binance,

Ethereum, Solana, USD and XRP (Table 1). The time plots in Figure 2, Figure 3, Figure 4, Figure 5, Figure 6 and Figure 7 show evidence of rising prices of Binance (Figure 2), Bitcoin (Figure 3), and Ethereum (Figure 4) towards the end of the series, while declining prices can be observed in Solana (Figure 5) and XRP (Figure 6). Figure 7 reveals that for USD, the price was almost the same towards the end of the series. From Figure 2 to Figure 7, it can be deduced that among the cryptocurrencies, Solana has had a significant upward movement in price compared with other cryptocurrencies. However, both Bitcoin and Ethereum have experienced a significant and noticeable surge in price after 1,000 days. Figure 8, Figure 9, Figure 10, Figure 11, Figure 12 and Figure 13 show the histogram plots for each cryptocurrency. These plots reveal that all the cryptocurrencies, excluding USD, are positively skewed (skewed to the right), while USD is symmetric (Figure 12).

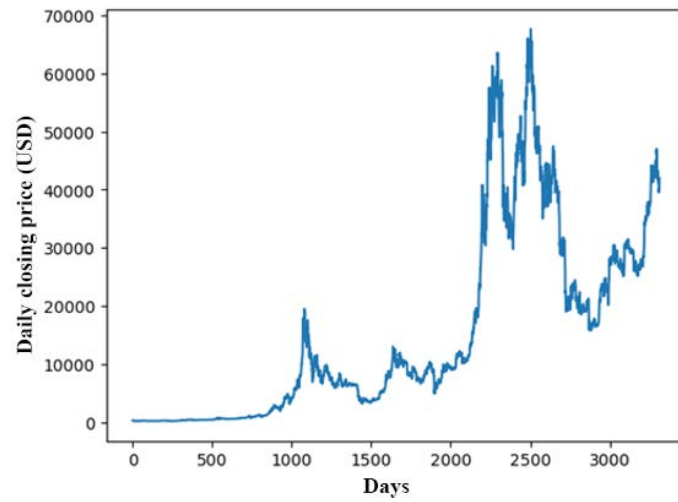
The comparative performance of these three ML algorithms is presented in Table 2. For Binance, GB, RF and Bagging, the  $R^2$  of 0.991852, 0.992792 and 0.992300 were obtained, and the adjusted  $R^2$  were 0.991740, 0.992693 and 0.992194, respectively. The RF obtained the highest  $R^2$  and adjusted  $R^2$  compared with other ML algorithms. In terms of forecasting performance, RF also obtained the lowest RMSE (15.067578) and MAE (6.497206) compared to GB and Bagging. For Bitcoin, GB obtained the highest  $R^2$  (0.997332), adjusted  $R^2$  (0.997308) and the lowest RMSE (837.293506) and MAE (422.809066) compared with other competing algorithms, thereby becoming the most suitable algorithm for predicting Bitcoin price. The result also shows that for Ethereum, USD and XRP, the RF algorithm outperformed other ML algorithms both in terms of fitness performance ( $R^2$  and adjusted  $R^2$ ) and forecasting performance (RMSE and MAE). For Solana, GB performed better than RF and Bagging (Table 2). The plot of the actual and predicted price of these cryptocurrencies for 30 days out of the sample was plotted based on the most suitable ML algorithms, and the figures obtained are shown in Figure 14, Figure 15, Figure 16, Figure 17, Figure 18 and Figure 19. The plots show a very close agreement between the actual and predicted price, validating the predictability power of these ML algorithms in the daily closing price prediction of these cryptocurrencies.

**Table 1.** Descriptive statistics for the cryptocurrencies

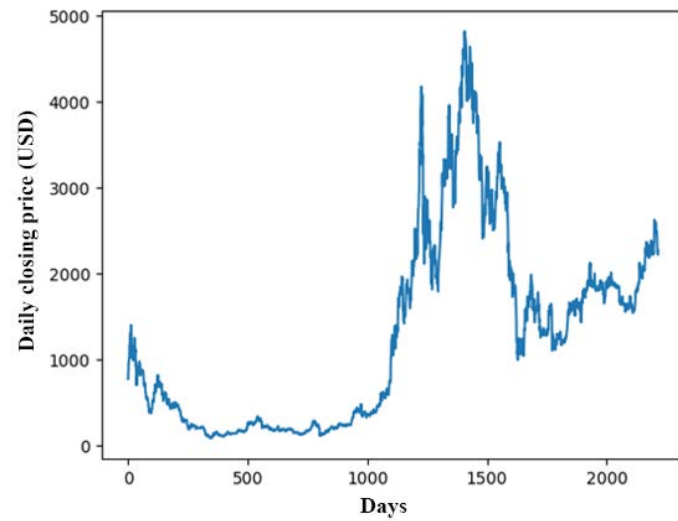
Statistics	Binance	Bitcoin	Ethereum	Solana	USD	XRP
N	2217	3313	2217	1121	1852	2217
Min.	4.189971	171.509995	82.829887	1.502038	0.877400	0.115093
Max.	634.549500	66382.062500	4718.039063	246.122421	1.023058	3.117340
Std.	167.976391	15925.273191	1089.302540	52.328679	0.006155	0.303340
Mean	166.486461	14866.360135	1228.212832	54.060337	0.998702	0.500553
25%	15.645951	1193.770020	224.641891	20.451468	0.998218	0.298669
50%	39.656357	8492.932617	1058.969971	31.525139	0.999529	0.424843
75%	296.519989	25677.480470	1851.828369	80.722099	0.999800	0.609635
COV (%)	100.8949	107.1229	88.69005	96.79681	0.6163	60.60098
Skewness	0.6584	1.1326	0.8894	1.5914	2.7075	2.5728
Kurtosis	-0.6789	0.2545	0.0268	1.7167	18.5583	12.21873



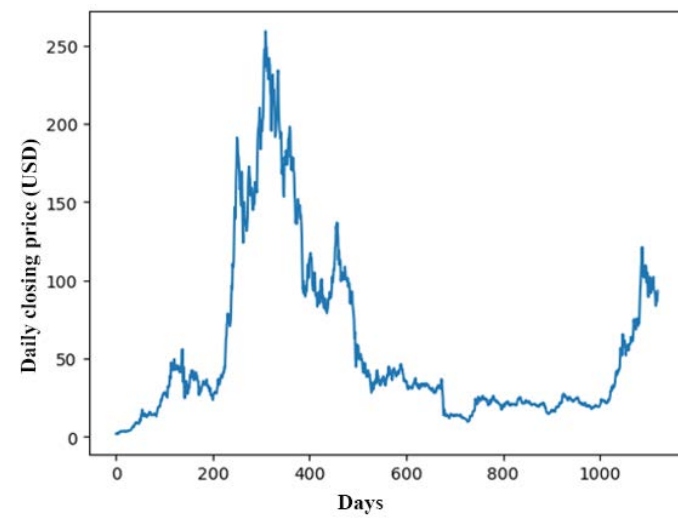
**Figure 2.** Time plot for the daily closing price of Binance



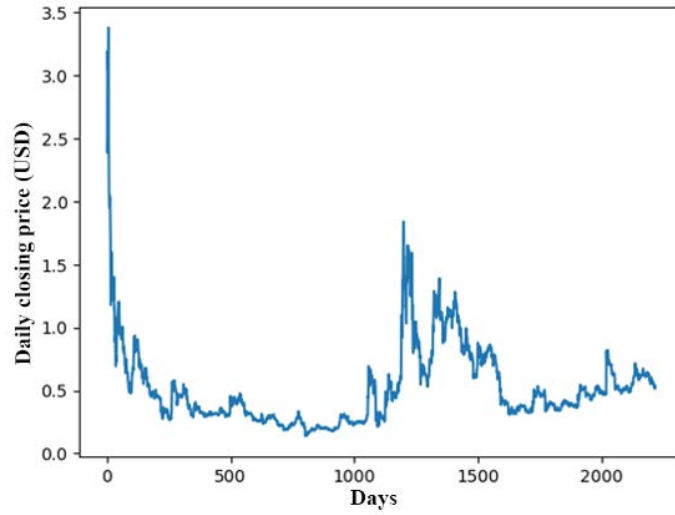
**Figure 3.** Time plot for the daily closing price of Bitcoin



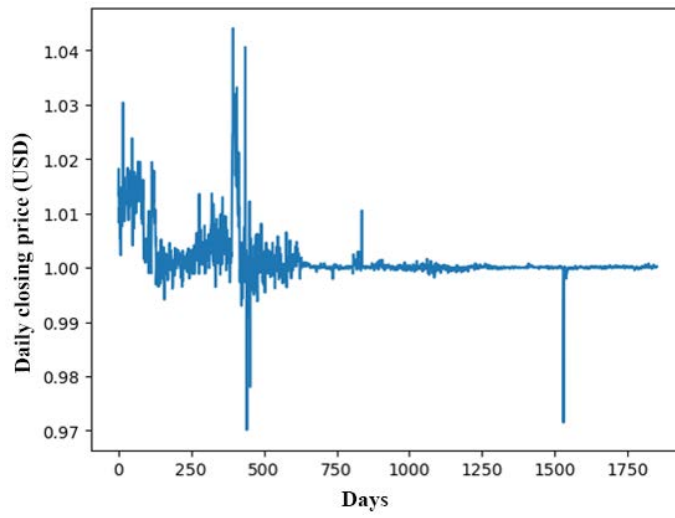
**Figure 4.** Time plot for the daily closing price of Ethereum



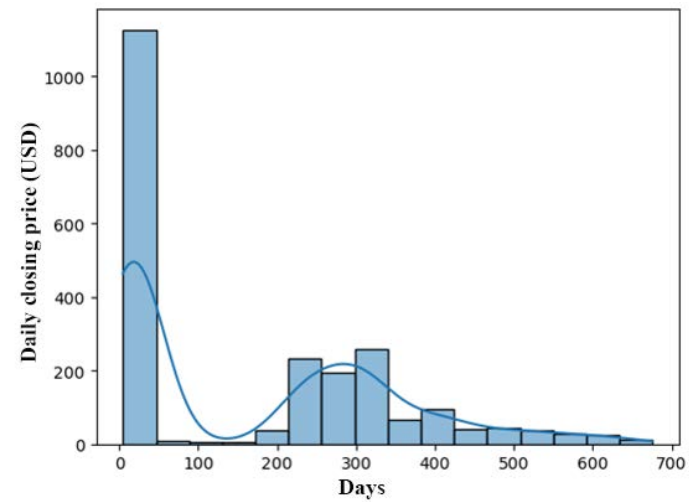
**Figure 5.** Time plot for the daily closing price of Solana



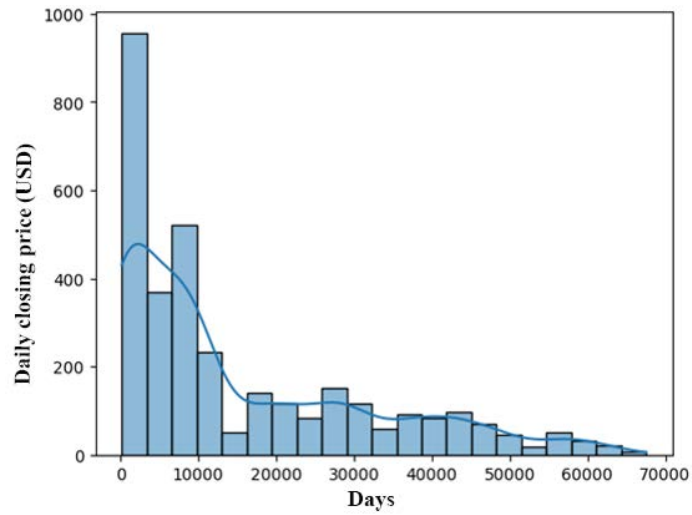
**Figure 6.** Time plot for the daily closing price of XRP



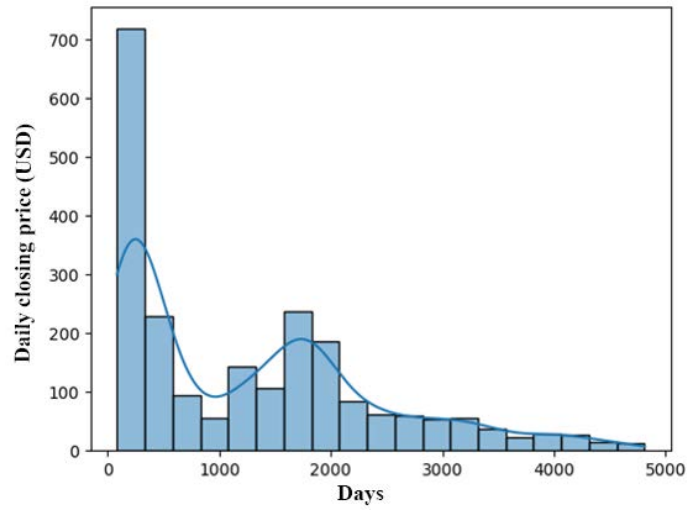
**Figure 7.** Time plot for the daily closing price of USD



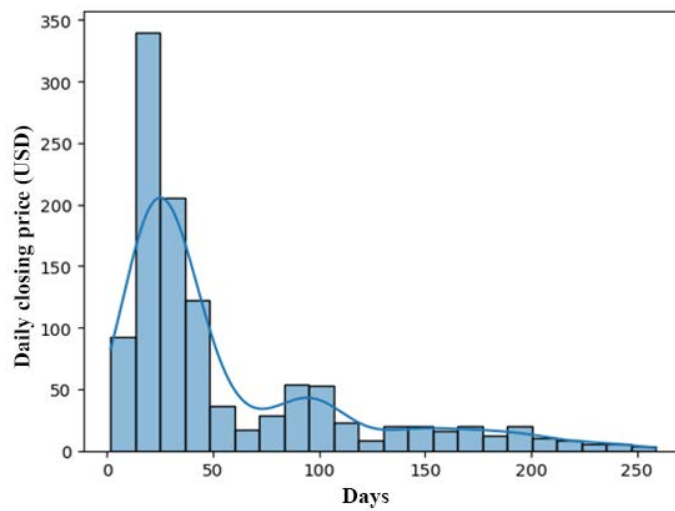
**Figure 8.** Histogram for the daily closing price of Binance



**Figure 9.** Histogram for the daily closing price of Bitcoin

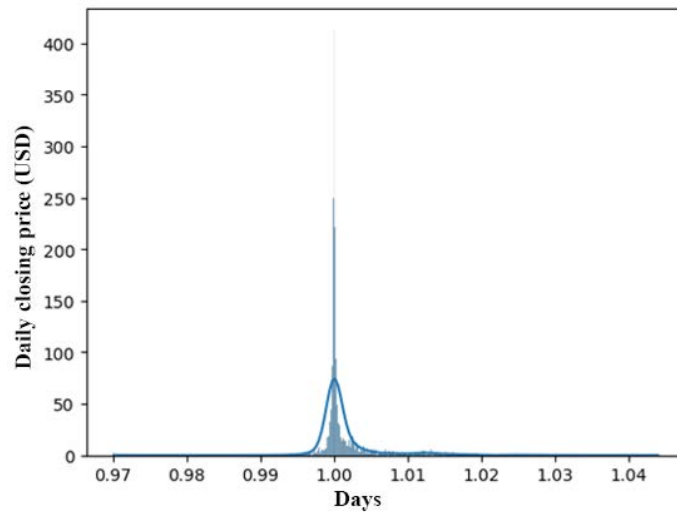


**Figure 10.** Histogram for the daily closing price of Ethereum

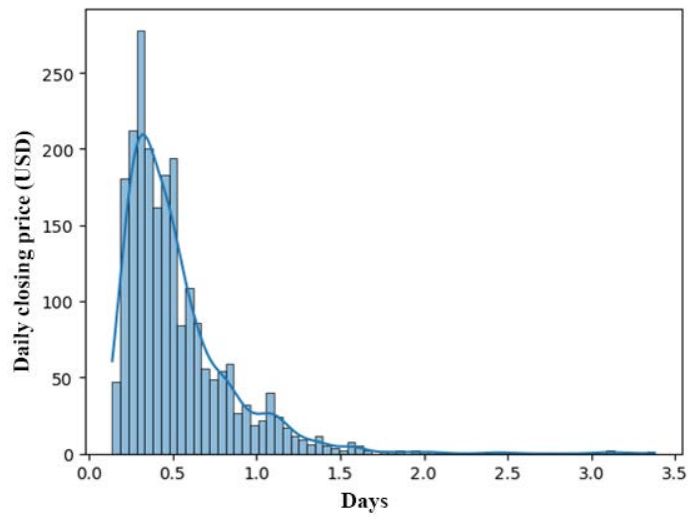


**Figure 11.** Histogram for the daily closing price of Solana

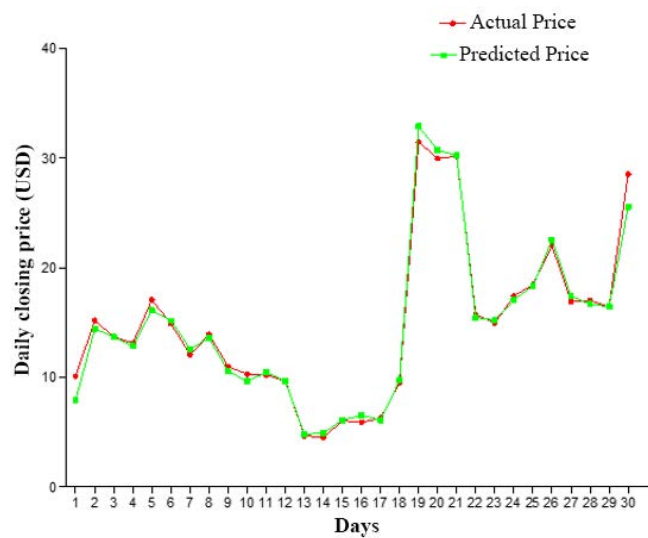




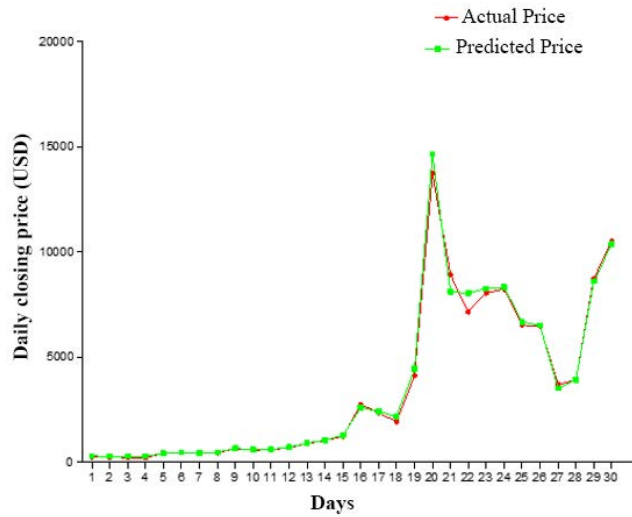
**Figure 12.** Histogram for the daily closing price of USD



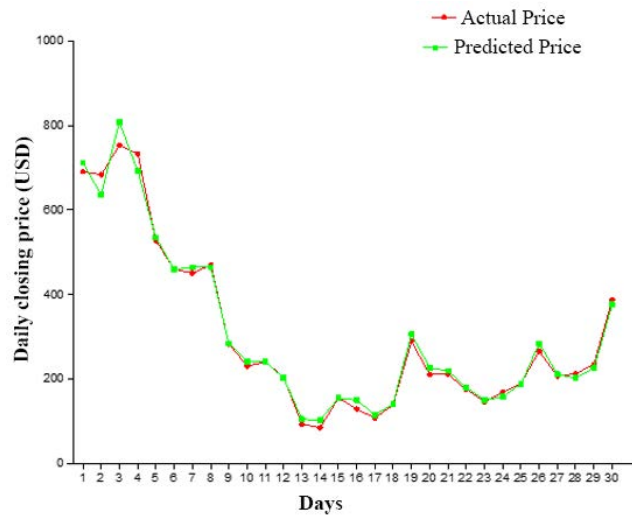
**Figure 13.** Histogram for the daily closing price of XRP



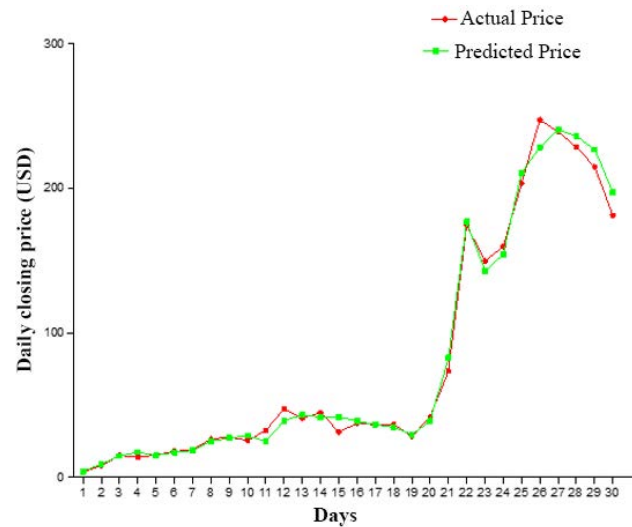
**Figure 14.** Graph of the 30-day actual and predicted closing price of Binance



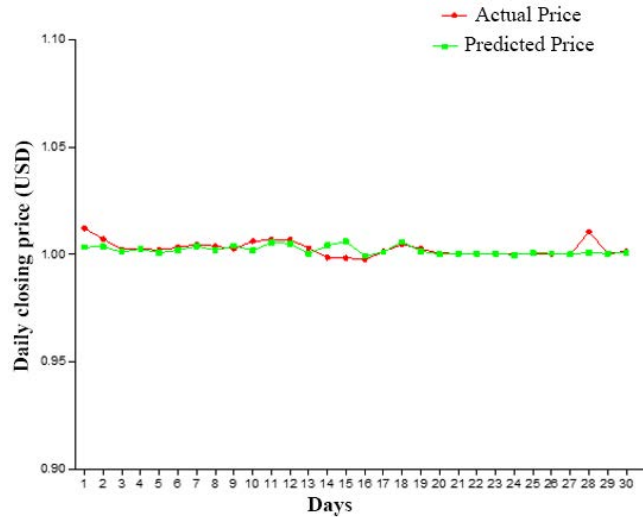
**Figure 15.** Graph of the 30-day actual and predicted closing price of Bitcoin



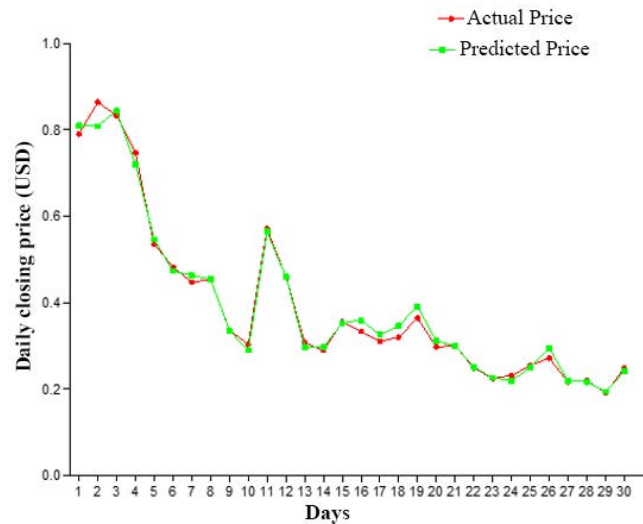
**Figure 16.** Graph of the 30-day actual and predicted closing price of Ethereum



**Figure 17.** Graph of the 30-day actual and predicted closing price of Solana



**Figure 18.** Graph of the 30-day actual and predicted closing price of USD



**Figure 19.** Graph of the 30-day actual and predicted closing price of XRP

It can be observed that Bitcoin has the highest COV and standard deviation, indicating a higher level of risk. The highest standard deviation also indicates that Bitcoin has the highest volatility among other cryptocurrencies, which is corroborated by the findings of Gupta and Vaishali [17]. This study also shows that cryptocurrencies are heavy-tailed, which is corroborated by the findings of Osterriedder [18] and Palstand and Ryden [19], where Bitcoin was found to have strong non-normal characteristics. All cryptocurrencies show a positive skewness that aligns with the findings of Karagiorgis et al. [20], Yang [21], and Liu and Tsyvinski [22].

The study also shows the superiority of RF regression over other ML algorithms, which is corroborated by Alarcon [23]. Similarly, Farouk et al. [5] also proposed that RF outperformed LR, AdaBoost, DT, KNN, GB, and neural networks in two of the datasets considered using  $R^2$ , Mean Absolute Percentage Error (MAPE) and MAE as the performance metrics. Derbentsev et al. [4] also confirmed that among the ensemble-based ML approaches, RF performed better than boosting in forecasting cryptocurrency prices. Both Bagging and GB reduce bias and enhance accuracy when dealing with complex relationships or imbalanced data. However, RF combines their strengths. Therefore, it is superior to both of them.

Boosting models are weighed based on performance. However, each model in Bagging regression receives equal weight, which is a possible reason why boosting regression outperforms Bagging regression. Boosting combines the predictions of weak learners to create a strong learner. However, Boosting models are trained sequentially, and each new model corrects errors made by the previous ones. This may lead to the superiority of boosting regression over Bagging regression. RF regression does not depend on the order or number of trees and is less prone to overfitting since it uses averaging and feature sampling to reduce the complexity and variance of the ensemble. This is a possible

reason why RF regression is superior to boosting regression. The trees in RF are independent and their output can be determined in any order, unlike boosting regression, which builds trees one at a time. In addition, RF combines results at the end of the process by averaging, while boosting combines results along the way. These could have given RF regression an edge over boosting regression in predicting the daily closing price of these cryptocurrencies. The implications of these findings for investors and the broader financial community are that using RF regression with the previous four-day lag values of cryptocurrency prices reliably estimates the prediction of their daily closing price. This algorithm could help guide investors and the financial community in decision-making with regard to the crypto market.

**Table 2.** Comparative performance of GB, RF and Bagging algorithms

Cryptocurrencies	ML Algorithms	R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE	MAE
Binance	GB	0.991852	0.991740	16.019852	6.692864
	RF	<b>0.992792</b>	<b>0.992693</b>	<b>15.067578</b>	<b>6.497206</b>
	Bagging	0.992300	0.992194	15.573332	6.820010
Bitcoin	GB	<b>0.997332</b>	<b>0.997308</b>	<b>837.293506</b>	<b>422.809066</b>
	RF	0.997075	0.997048	876.798318	440.305001
	Bagging	0.996906	0.996877	901.804698	460.708941
Ethereum	GB	0.993212	0.993119	<b>96.756653</b>	<b>49.565825</b>
	RF	<b>0.997075</b>	<b>0.997048</b>	876.798318	440.305001
	Bagging	0.992704	0.992603	100.315682	53.394860
Solana	GB	<b>0.989007</b>	<b>0.988701</b>	<b>5.791002</b>	<b>3.210980</b>
	RF	0.988687	0.988373	5.874518	3.223499
	Bagging	0.986669	0.986298	6.377180	3.454734
USD	GB	0.641312	0.635383	0.002437	0.001085
	RF	<b>0.665794</b>	<b>0.634944</b>	<b>0.000302</b>	<b>0.000222</b>
	Bagging	0.595462	0.588776	0.002588	0.001156
XRP	GB	0.956477	0.955878	0.062877	0.023789
	RF	<b>0.969956</b>	<b>0.969542</b>	<b>0.052242</b>	<b>0.024776</b>
	Bagging	0.962607	0.962093	0.058281	0.025386

## 5 Conclusions

This study explored the use of three different ML algorithms (i.e., GB, RF, and Bagging) in predicting the daily closing price of six cryptocurrencies. Results showed that the RF regression outperformed other ML algorithms for Binance, Ethereum, USD and XRP, while GB outperformed other ML algorithms for Bitcoin. This study shows the superiority of the RF regression in predicting the closing price of most of these cryptocurrencies. The RF regression is superior to the GB and Bagging regression algorithms because it combines their strengths in prediction.

When using these algorithms, specifically RF regression for Binance, Ethereum, USD and XRP, and GB regression for Bitcoin and Solana, it helps guide investors in making trading decisions to increase their chances of making profits. Other algorithms and deep learning algorithms, such as LSTM, also need to be considered for better prediction of these cryptocurrencies. In addition, similar studies need to be conducted for cryptocurrencies not included in this study.

## Data Availability

The data used to support the research findings are available from the corresponding author upon request.

## Conflicts of Interest

The author declares no conflict of interest.

## References

- [1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decent. Bus. Rev.*, vol. 2008, p. 21260, 2008. <https://bitcoin.org/bitcoin.pdf>
- [2] R. Farell, "An analysis of the cryptocurrency industry," no. 130, 2015. <https://www.coursehero.com/file/38603007/An-Analysis-of-the-Cryptocurrency-Industrypdf/>
- [3] S. Lahmiri and S. Bekiros, "The impact of the COVID-19 pandemic upon stability and sequential irregularity of equity and cryptocurrency markets," *Chaos Solitons Fractals*, vol. 138, p. 109936, 2020. <https://doi.org/10.1016/j.chaos.2020.109936>

- [4] V. Derbentsev, N. Datsenko, V. Babenko, O. Pushko, and O. Pursk, "Forecasting cryptocurrency prices using ensembles-based machine learning approach," in *2020 IEEE International Conference on Problems of Info Communications. Science and Technology (PIC S&T)*. Kharkiv, Ukraine, 2020. <https://doi.org/10.1109/PICS T51311.2020.9468090>
- [5] M. Farouka, N. S. Ragaba, D. Salamab, O. Elrashidya, L. Mandoura, M. Ahmed, J. Walid, M. Mesbah, R. Attia, N. Ahmeda, and R. Elazaba, "Bitcoin ML: An efficient framework for bitcoin price prediction using machine learning," *J. Comput. Commun.*, vol. 3, no. 1, pp. 70–87, 2024.
- [6] A. Alshehri, "Predicting cryptocurrency returns using classification and regression machine learning models," Master's thesis, The School of Computing Sciences of the University of East Anglia, 2022.
- [7] B. Shilpa, L. Kasal, M. Shetty, R. Pai, and T. Nayak, "Cryptocurrency price prediction using machine learning," *Int. Res. J. Modernization Eng. Technol. Sci.*, vol. 4, no. 6, pp. 3561–3565, 2022.
- [8] P. Jaquart, S. Kopke, and C. Weinhardt, "Machine learning for cryptocurrency market prediction and trading," *J. Finance Data Sci.*, vol. 8, pp. 331–352, 2022. <https://doi.org/10.1016/j.jfds.2022.12.001>
- [9] L. Pan, "Cryptocurrency price prediction based on ARIMA, random forest, and LSTM algorithm," *BCP Bus. Manag.*, vol. 38, pp. 3396–3404, 2022. <https://doi.org/10.54691/bcpbm.v38i.4313>
- [10] S. A. Basher and P. Sadorsky, "Forecasting bitcoin price direction with random forests: How important are interest rates, inflation, and market volatility?" *Mach. Learn. Appl.*, vol. 9, p. 100355, 2022. <https://doi.org/10.1016/j.mlwa.2022.100355>
- [11] V. Srivastava, V. Dwivedi, and A. Singh, "Cryptocurrency price prediction using enhanced PSO with extreme gradient boosting algorithm," *Cybern. Inf. Technol.*, vol. 23, no. 2, pp. 170–187, 2023.
- [12] X. Yan, *Forecasting Cryptocurrency Prices*. Imperial College London, 2020.
- [13] M. Saad, J. Choi, D. Nyang, J. Kim, and A. Mohaisen, "Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions," *IEEE Syst. J.*, vol. 10, no. 10, pp. 1–12, 2018. <https://doi.org/10.1109/jsyst.2019.2927707>
- [14] A. Turukmane, C. Manasa, V. Yaraswini, B. SriTonya, and A. Vineetha, "Bitcoin value prediction," *Int. J. Creat. Res. Thoughts*, vol. 11, pp. a820–a825, 2020.
- [15] F. M. Sakran, "Cryptocurrency analysis using machine learning and deep learning approaches," *J. Comp. Electr. Electron. Eng. Sci.*, vol. 2023, no. 1, pp. 29–33, 2023. <https://doi.org/10.51271/JCEES-0007>
- [16] M. X. Wang, D. Huang, G. Wang, and D. Q. Li, "SS-XGBoost: A machine learning framework for predicting newmark sliding displacements of slopes," *J. Geotech. Geoenviron. Eng.*, vol. 146, no. 9, p. 04020074, 2020. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002297](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002297)
- [17] G. Gupta and Vaishali, "Determinants of cryptocurrency: An analysis of volatility and risk-return trade-off," *Knowledgeable Res.*, vol. 1, no. 8, pp. 26–35, 2023. <https://doi.org/10.57067/pprt.2023.1.08.25-35>
- [18] J. Osterrieder, "The statistics of bitcoin and cryptocurrencies," *Adv. Econ. Bus. Manag. Res. (AEBM)*, vol. 26, pp. 285–289, 2017.
- [19] E. Parlstrand and O. Ryden, *Explaining the Market Price of Bitcoin and Other Cryptocurrencies with Statistical Analysis*. Department of Mathematics Kungliga Tekniska Högskolan, 2015. <https://www.diva-portal.org/smash/get/diva2:814478/FULLTEXT01.pdf>
- [20] A. Karagiorgis, A. Ballis, and K. Drakos, "The skewness-kurtosis plane for cryptocurrencies' universe," *Int. J. Finance Econ.*, pp. 1–13, 2023. <https://doi.org/10.1002/ijfe.2795>
- [21] T. Yang, "Skewness in the cryptocurrency market," *BCP Bus. Manag.*, vol. 21, pp. 425–432, 2022. <https://doi.org/10.54691/bcpbm.v21i.1268>
- [22] Y. Liu and A. Tsyvinski, "Risks and returns of cryptocurrency," *NBER Working Paper*, no. 24877, pp. 1–25, 2018.
- [23] E. Alarcón, *Inference and Prediction of Cryptocurrency Market Returns*. Lund University School of Economics and Management Lund, Sweden, 2020. <https://lup.lub.lu.se/luur/download?func=downloadFile&recordId=9023391&fileId=9023395>