The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2444-8494.htm

Prediction of cryptocurrency's price using ensemble machine learning algorithms

European Journal of Management and Business Economics

Received 1 August 2023
Revised 19 May 2024
13 August 2024
19 October 2024
24 October 2024
Accepted 24 November 2024

N.S.S. Kiranmai Balijepalli
Symbiosis Institute of Business Management, Symbiosis International University, Pune, India, and
Viswanathan Thangaraj
Symbiosis Institute of Business Management, Symbiosis International University, Pune, India and
Prin. L. N. Welingkar Institute of Management Development and Research, Bangalore, India

Abstract

Purpose – Cryptocurrency markets are gaining popularity, with over 23,000 cryptocurrencies in 2023 and a total market valuation of 870.81 billion USD in 2023. With its increasing popularity, cryptocurrencies are also susceptible to volatility. Predicting the price with the least fallacy or more accuracy has become the need of the hour as it significantly influences investment decisions.

Design/methodology/approach – This study aims to create a dynamic forecasting model using the ensemble method and test the forecasting accuracy of top 15 cryptocurrencies' prices. Statistical and econometric model prediction accuracy is examined after hyper tuning the parameters. Drawing inferences from the statistical model, an ensemble model using machine learning (ML) algorithms is developed using gradient-boosted regressor (GBR), random forest regressor (RFR), support vector regression (SVR) and multi-layer perceptron (MLP). Validation curves are utilized to optimize model parameters and boost prediction accuracy.

Findings – It is found that when the price movement exhibits autocorrelation, the autoregressive integrated moving average (ARIMA) model and the ensemble model performed better. ARIMA, simple linear regression (SLR), random forest (RF), decision tree (DT), gradient boosting (GB) and multi-model regression (MLR) ensemble models performed well with coins, showing that trends, seasonality and historical price patterns are prominent. Furthermore, the MLR approach produces more accurate predictions for coins with higher volatility and irregular price patterns.

Research limitations/implications – Although the dataset includes crisis period data, anomalies or outliers are yet to be explicitly excluded from the analysis. The models employed in this study still demonstrate high accuracy in predicting cryptocurrency prices despite these outliers, suggesting that the models are robust enough to handle unexpected fluctuations or extreme events in the market. However, the lack of specific analysis on the impact of outliers on model performance is a limitation of the study, as it needs to fully explore the resilience of the forecasting models under adverse market conditions.

Practical implications – The present study contributes to the body of literature on ensemble methods in forecasting crypto price in general, potentially influencing future studies on price forecasting. The study motivates the researchers on empirical testing of our framework on various asset classes. As a result, on the prediction ability of ensemble model, the study will significantly influence the decision-making process of traders and investors. The research benefits the traders and investors to effectively develop a model to forecast cryptocurrency price. The findings highlight the potential of ensemble model in predicting high volatile cryptocurrencies and other financial assets. Investors can design the investment strategies and asset allocation decisions by understanding the relationship between market trends and consumer behavior. Investors can enhance portfolio performance and mitigate risk by incorporating these insights into their decision-making processes. Policymakers can use this information to design more effective regulations and policies promoting economic stability and consumer welfare. The study emphasizes the need for using diversified model to understand the market dynamics and improving trading strategies.

European Journal of Management and Business Economics
Vol. 35 No. 4, 2026
pp. 501-517
Emerald Publishing Limited
e-ISSN: 2444-8494
p-ISSN: 2444-8451
DOI 10.1108/EJMBE-08-2023-0244

Originality/value – This research, to the best of our knowledge, is the first to use the above models to develop an ensemble model on the data for which the outliers have not been adjusted, and the model still outperformed the other statistical, econometric, ML and deep learning (DL) models.

Keywords Cryptocurrencies, Forecasting, Price prediction, Analytics, Machine learning, Multi-model regression model

Paper type Research paper

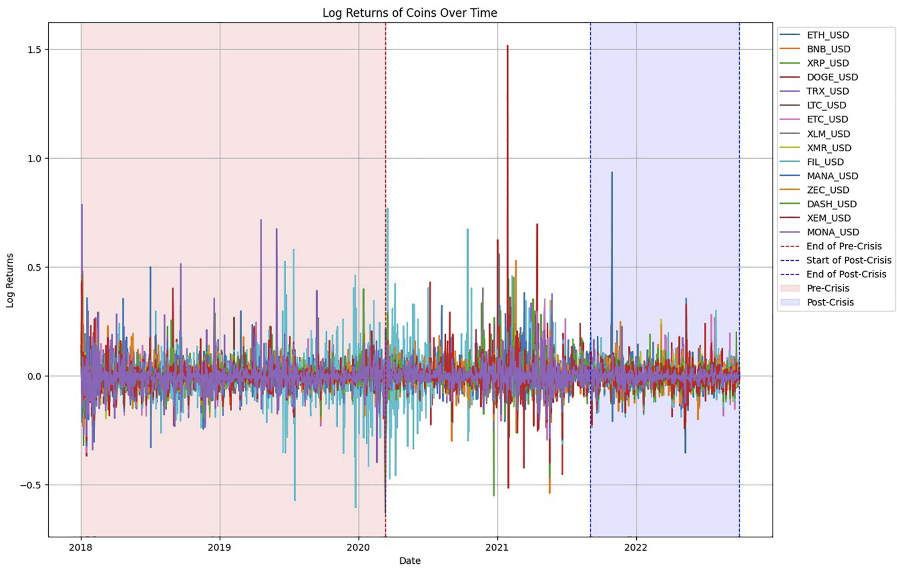
1. Introduction

A cryptocurrency is a cryptographically secured digital currency distributed on a decentralized network using blockchain technology, making them anonymous and untraceable. The advent of cryptocurrencies and blockchain technology has sparked a revolutionary shift in the financial sector (Kayani, 2023; Kayani and Hasan, 2024).

Cryptocurrencies offer substantial returns on investments besides possessing a higher risk factor. A classic example would be the Bitcoin index (in EUR), which stands at a 108.27% compound annual growth rate and a standard deviation of 156.99% between 2011 and 2024. Figure 1 shows the abnormal spike in Bitcoin return during 2021, with significantly higher fluctuations indicating heightened volatility. This calls for accurate forecasting models to combat the large fluctuations in non-stationary cryptocurrencies (Bouteska *et al.*, 2024). Prediction of cryptocurrency prices is also important for several other reasons, such as building trading strategies, risk management, price discovery, market sentiment analysis and business applications.

Fundamentalists and chartists rely on price prediction to make an informed decision. Chartists looking for price patterns and trends can refine their technical analysis based on forecasting. While fundamentalist's primary focus remains on long-term value assessment, the model's predictions could potentially serve as an additional data point for considering short-term market sentiment that might influence their investment decisions (Soltani *et al.*, 2023).

The early-stage researchers employed statistical and econometrics models to forecast the price of cryptocurrencies. To predict Bitcoin's short-term prices, the autoregressive integrated moving average (ARIMA) model was applied by Wirawan *et al.* (2019). The progressive growth



Source(s): Figure by authors

Figure 1. Historical returns of cryptocurrencies

in machine learning (ML) also motivated researchers to compare the forecasting accuracy of various models. Lyu (2022) compared the accuracy of various ML models and found gradient boosting (GB) to be the most efficient in predicting most major cryptocurrencies.

Recent studies on forecasting started applying deep learning (DL) and hybrid models, combining classical models such as the ARIMA model and artificial neural networks (Suhartono *et al.*, 2017). Li *et al.* (2022) found that the hybrid model improved the forecasting accuracy of foreign currency and demonstrated that the hybrid data decomposition model outperforms econometric, ML and deep-learning models. Chen (2023) predicted the Bitcoin price of the next day and found that random forest (RF) gives better accuracy than long short-term memory (LSTM). An extensive literature review revealed limited studies using ensemble methods to predict the prices of multiple cryptocurrencies and hyper-tune the parameters to improve forecasting accuracy (Derbentseva *et al.*, 2021).

Furthermore, the existing literature predominantly focuses mostly on Bitcoin. Empirical research is necessary to test the forecasting models in a broader range of cryptocurrencies that display diverse price dynamics and volatility, which have evolved over the past few years (Piryonesi and El-Diraby, 2020). ML has the power to identify complex patterns in price movements and overcome the limitations of traditional models to improve forecasting accuracy. In contrast, traditional models like ARIMA may exhibit reliability in some cases (Nakano *et al.*, 2018). Ensemble model combine the output from multiple models, hyper-tune the parameters and improve the forecasting accuracy. A systematic, three-stage approach is followed to create an ensemble model for predicting cryptocurrency prices. First, a dataset of historical prices of the 15 cryptocurrencies selected based on their market capitalization. Then, the base models, random forest regressor (RFR), gradient boosting regressor (GBR), support vector regression (SVR) and MLP regressor, are trained on the training set, optimizing their hyper-parameters. With the base models trained, predictions on the test set are made using each model. Finally, these individual predictions are combined using stacking to create the ensemble model. Root mean squared error (RMSE), mean squared error (MSE), mean absolute error (MAE) and R2 are used to evaluate the ensemble model's performance on the 15 cryptocurrencies.

The study addresses the critical gaps in the existing literature. First, the existing studies have explored the statistical or ML models in silos. The traditional statistical models fail to capture the non-linear and rapid price movement of cryptocurrencies. Further it carries limitations such as hyperparameter tuning and lacks computational efficiency. The standalone ML models suffer from issues such as overfitting or underfitting adversely affecting the forecasting accuracy. This study bridges the gap by applying a robust ensemble machine leveraging the potential of ML to improve forecasting accuracy. Further the existing studies on forecasting, focus mostly on a single cryptocurrency, Bitcoin in general and few others like Ethereum. The broader cryptocurrency market remains largely unresearched. The study provides a methodological framework to empirically test the forecasting accuracy of ensemble model on a broader set of cryptocurrencies. Ensemble models are flexible in capturing non-linear relationships, complex patterns in the data and provide superior forecasting performance as compared to the standalone models. This paper contributes to the existing theories on predictive analytics, providing empirical evidence by creating an ensemble ML model. In addition, it contributes to the body of knowledge on how to leverage bias-variance trade-offs in forecasting. To fill the research gap, the study aims to address the following questions: (1) How to develop an ensemble model to accurately predict the prices of multiple cryptocurrencies? (2) Does the ensemble model outperform the statistical and econometric model in improving forecasting accuracy? (3) Does the ensemble model provide a better estimate of future price on a broader set of multiple cryptocurrencies that differ in return and volatility dynamics?

The rest of the paper is structured as follows. A brief literature review is detailed in Section 2, outlining the previous papers that used various models to predict the price of cryptocurrencies. Section 3 explores the methodology and the ensemble prediction approach of this study, whereas Section 4 presents the empirical findings, including a comparison of the proposed ensemble method with econometric and ML models. The paper concludes with Section 5, providing avenues for future research in forecasting.

2. Literature review

Over the last decade, cryptocurrencies have gone up from an obscure asset to a preferred investment due to millennials' popularity, increased risk-taking propensity, perceived profitability, decentralization and acceptance of cryptocurrency as an investment option.

Cryptocurrency price forecasting is considered one of the financial domain's most difficult predictions (Livieris *et al.*, 2021). Challenges, complexity and interpretability of models have been classified as a problem of time series forecasting by most successful researchers (Adegboruwa *et al.*, 2019; Hamayel and Owda, 2021; Lahmiri and Bekiros, 2019; Patel *et al.*, 2020; Tandon *et al.*, 2019; Wirawan *et al.*, 2019).

This literature review examines the existing research studies on cryptocurrency price forecasting and the various possible methodologies for the same. Traditional statistical and econometric models have been considered for the analysis, which goes on to assess promising ML techniques. Finally, the potential of combining these methodologies to create hybrid models is analyzed.

2.1 Existing approaches to cryptocurrency price forecasting

2.1.1 Traditional statistical, econometric and ML models. The early research on cryptocurrency price forecasting relied on traditional econometric and statistical models. Efficient and immediate solutions for time series problems can be achieved through statistical models (Meenu *et al.*, 2020). Econometric approaches that combine statistical and economic principles, according to Alahmari (2019) are crucial to estimating and predicting economic variables relevant to cryptocurrency prices. Time-series analysis often employs the ARIMA model, which, as discussed above, is a widespread technique (El-Bannany *et al.*, 2020). Yet, the statistical and econometric models in forecasting come along with limitations. The statistical models rely on specific assumptions such as the normal distribution of data, the linear relationship between the variables and the stability of parameters over time. Thus, statistical models perform poorly either with noisy data or with changes in the trends and patterns.

Nikou *et al.* (2019) found that the ARIMA model, while suitable for price prediction within specific periods, struggles with capturing sharp price fluctuations. Similarly, Greaves and Au (2015) also observed significant prediction errors with ARIMA, particularly in the cases pertaining to volatile markets. Moreover, statistical and econometric models work within a specific framework, thus limiting the model's ability to handle complexities in real-world data. The presence of non-linear relationships between variables adversely affects the forecasting accuracy.

ML techniques have gained significant traction in cryptocurrency price forecasting as a consequence of advancements in big data technology and artificial intelligence. These techniques analyze complex data patterns and make predictions powerfully (Chevallier *et al.*, 2021; Derbentseva *et al.*, 2021).

2.1.2 Integration of techniques (ensemble models). There is a solid foundation through traditional statistical and econometric models to understand trends and relationships of financial data. However, cryptocurrency markets present the complexity of non-linear patterns that traditional models might not capture. ML techniques have the ability to identify such patterns with ease, only when provided with large datasets and careful tuning to avoid overfitting.

The benefits of statistical models lie in providing a foundational framework and distinguishing patterns and trends. For ML, the same reflects in uncovering the intricacies of complex non-linear relationships within the data. Through a combined approach, it is possible to tackle the limitations of individual models and optimize prediction accuracy (Madan *et al.*, 2015).

Hybrid models that combine traditional and ML approaches can outperform classical ML models in terms of accuracy (Alahmari, 2019). The strengths of both methodologies come together to optimize accuracy in such a hybrid model.

2.1.3 Effects of information shocks on cryptocurrency prices. Wang *et al.* (2020) studied the relationship between economic policy uncertainty (EPU) and Bitcoin prices. Their findings suggest that EPU has a negative impact on Bitcoin prices in the short term. This effect gradually diminishes over time and has no notable long-term effects. Similarly, positive information shocks lead to higher volatility in the short run, while negative shocks have left no significant impact on cryptocurrency prices (Yin *et al.*, 2021)

3. Theoretical framework

Cryptocurrency markets are known for their intense fluctuations and unpredictability. A predictive model which can efficiently navigate price fluctuations is required for effective cryptocurrency price prediction. Comparative studies of statistical models and their strengths and weaknesses are lacking (Lobell and Burke, 2010). ML models are dynamic and can deal with complexities to predict the price with a high degree of accuracy. Researchers used ML models such as DT, GB, SVR and MLP. Contradicting the researchers who supported ML models for price forecasting, Cohn *et al.* (1996) have concluded that statistically-based learning architectures combined with a mixture of Gaussians and locally weighted regression improve forecasting accuracy. Feature engineering and feature extraction using ML models can improve forecasting accuracy.

Ensemble model synthesize the output from multiple ML models to make an accurate forecast. It works like seeking advice from many sources thus improving the forecasting accuracy. The model improves the forecasting accuracy by merging predictions from multiple models. The learning rate of ensemble model can be effectively improved through optimization and hyperparameter tuning. By combining diverse models, ensemble forecasting accuracy can be improved using features such as maximum vote, stacking, blending, bagging and boosting. The combination of multiple algorithms in an ensemble model makes it capable of outperforming single-base models in predicting cryptocurrency (Yang *et al.*, 2022).

4. Methodology

4.1 Research approach and data source

Primarily differing in their approach to value stability and underlying mechanisms, stablecoins and cryptocurrencies are quite different. Stablecoins are engineered to mitigate price volatility by pegging their worth to a stable asset or fiat currency such as the US dollar or Euro. This ensures their consistency in value. This research paper focuses on the prediction of cryptocurrency prices due to their volatility when compared to stable asset prices. By focusing solely on non-stablecoin cryptocurrencies, our research avoids potential complexities introduced by stablecoins, whose value is pegged to external assets and may exhibit less price volatility compared to other cryptocurrencies.

Daily price data of 15 coins, based on market capitalization, is collected to build the forecasting model starting from January 2018 to September 2022. The coins selected for the study are Ethereum (ETH), Binance (BNB), Ripple (XRP), Dogecoin (DOGE), TRON (TRX), Litecoin (LTC), Ethereum Classic (ETC), Stellar (XLM), Monero (XMR), File coin (FIL), Decentraland (MANA), ZCash (ZEC), Dash (DASH), NEM coin (XEM) and Mona coin (MONA), all priced in USD. Data regarding the prices is obtained from Coinmarketcap as it offers comprehensive coverage and is widely used in cryptocurrency research (Vidal-Tomás, 2022).

The research methodology is divided into four stages. The first stage is to examine the forecasting accuracy of statistical models and refine them further to improve the forecasting accuracy. We evaluate the forecasting accuracy of ARIMA in stage two. In Stage Three, ML algorithms, including GB, DT, RF, SVR and MLP, are employed.

The final stage combines the best-performing ML models to optimize the forecasting accuracy further and build the model. The three models are RF With eXtreme Gradient

Boosting (XGBoost), Ensemble model with RF, GB, DT and multi-model regression (MLR). *R*-squared, RMSE, MSE and mean absolute error were employed to evaluate the performance of the models.

4.2 Models employed for forecasting

4.2.1 Simple linear regression (SLR). A SLR model is applied to forecast future prices when there is a linear relationship between dependent and independent variables. The equation has the form Y = a + bX, where Y is the dependent variable, i.e. the price of cryptocurrencies and X is the independent variable (time), b is the slope of the line and a is the y-intercept.

4.2.2 Auto-regressive integrated moving average (ARIMA). The ARIMA model, a stochastic process, includes sums of autoregressive and moving average components. The model is applicable when there is autocorrelation among the residuals. The model includes three parameters, i.e. p, d and q.

p-auto-regressive component

d-order of differencing

q-moving average component.

ARIMA model specification

$$\mathrm{Yt}' = \mathrm{C} + \varnothing 1\mathrm{Yt} - 1' + \ldots + \varnothing \mathrm{pYt} - \mathrm{p}' + \varnothing 1\mathrm{t} - 1 + \ldots + \varnothing \mathrm{qt} - \mathrm{q} + \mathrm{t} \tag{1}$$

4.2.3 Random forest (RF). RF, a supervised learning algorithm, is applicable to solving classification and regression problems. The first work on RF goes back to Ho (1995). Breiman (2001) further developed the idea and presented the RF algorithm in 2001. Various DTs are created on randomly selected samples, the forecasted results are produced from each tree and then the best tree is chosen by voting. The key benefits of RF are its generalization capability and minimal sensitivity to hyperparameters (Naing and Htike, 2015). In the prediction of cryptocurrency prices, RF has been used for BTC forecasting (Chevallier *et al.*, 2021) and BTC, ETH and XRP in (Derbentseva *et al.*, 2021).

4.2.4 Decision tree. DT, a type of supervised ML model, made of decision nodes and leaf nodes, resembling a tree-like structure with branches and leaves. The decision node represents features and attributes, and the leaf node represents the outcome. Inducing DT is one of the oldest and most popular techniques for learning discriminatory models, which has been developed independently in the statistical (Breiman, 2017; Kass, 1980) and ML (Hunt *et al.*, 1966; Quinlan, 1983, 1986) communities.

4.2.5 Gradient boosting (GB). GB, an ML technique, is used to solve regression and classification issues. A collection of weak prediction models, often DT, are produced as outcomes (Madeh and El-Diraby, 2021). GB frequently outperforms RF (Madeh and El-Diraby, 2021; Piryonesi and El-Diraby, 2020), allowing optimization of any differentiable loss function.

The loss function and the base-learner models can be arbitrarily specified on demand. Given a loss function denoted by $\Psi(y, f)$ and/or a custom base-learner $h(x, \theta)$, obtaining a solution to the parameter estimates can be difficult.

4.2.6 Support vector regressor (SVR). SVR is used for regression tasks to minimize forecasting errors. It operates by identifying a hyperplane with maximum margin from the training data. Support vectors, kernel functions and hyperparameter tuning are employed to capture complex relationships and handle outliers. This algorithm can powerfully predict continuous numerical values. Built on support vector machines for classification, SVR enables both linear and non-linear regressions. Examples of SVR usage in forecasting cryptocurrency prices can be found in (Chevallier *et al.*, 2021; Khedr *et al.*, 2021).

4.2.7 Multi-layer perceptron (MLP) regressor. MLP regressor processes input data with multiple layers of interconnected nodes (neurons) to form non-linear transformations.

It employs the backpropagation algorithm to adjust the weights and biases of the network, optimizing the model to minimize the error. Many complicated problems are not linearly separable, so the single hidden layer perceptron, which can only solve linearly separable problems, is not effective (Rosenblatt, 1958). These issues require a solution with multiple hidden layers (Rumelhart *et al.*, 1986). The MPLNN, a feedforward neural network with multiple layers, comprises three layers: an input layer, one or more hidden layers and an output layer.

4.2.8 Ensemble models. Three ensemble models have been created and tested to enhance the accuracy of this study.

4.2.8.1 RF ensembles with XGBoost. GB is used in this approach to train an RF ensemble. The boosting technique is applied to the DTs in RF ensembles by updating the weights of the observations based on the ensemble residuals. Then, each tree in the ensemble is trained to correct the errors of the previous trees. This is done by using the updated weights of the observations. Such an iterative process continues as long as the ensemble's performance continues to improve.

The final ensemble is expected to be more accurate than a simple RF, as the GB algorithm is used to train the DTs.

4.2.8.2 Averaging ensemble model with RF, GB and DT. This method integrated multiple ML model forecasts to increase the forecast's overall accuracy. This ensemble technique is built to individually develop each of the three models, namely RF, GB and DT, which will then return the average of their predictions. Each model may cause errors, so this method can balance them. This practice increases the overall accuracy of the ensemble model's predictions. Each model is trained with a distinct set of algorithms and parameters on the same data set.

4.2.8.3 Multi-model regression model (MLR). This approach employs the RFR, the GB, the SVR and the MLP regressor to form an MLR. Each model's capabilities enhance this ensemble model to increase forecast accuracy. A meta-features data frame is generated by the MLR, which trains a linear regression meta-model using the ensemble of these models. This effectively captures varied patterns and correlations in cryptocurrency data. The MLR technique's success in lowering prediction errors and offering insights into model performance is validated through the RMSE, MSE, MAE and R2 assessment measures.

5. Analysis

5.1 Descriptive statistics

Table 1 shows the descriptive statistics of 15 cryptocurrencies. Generally, cryptocurrencies show high volatility and also create volatility clustering. A perfect normal distribution data series has a skewness value of 0, and kurtosis equals 3. The skewness, kurtosis and Jarque–Bera results indicate that cryptocurrencies' price movements do not follow a normal distribution. This makes it evident that investors face few larger gains and many periods of loss.

5.2 Mean absolute error

Table 2 shows the MAE of cryptocurrency prices. Upon examining the results, it is evident that the performance of the models varies across different coins. Starting with XRP, the RF model exhibits the lowest MAE of 0.019, outperforming other models such as ARIMA, SLR, DT, SVR and MLP. The RF model consistently demonstrates superior accuracy with comparatively lower MAE values than other models for BNB, MONA, XLM, DASH, XMR, FIL and TRX. However, the MLP model shows the lowest MAE values for ETH, LTC, ZEC, MANA and ETC, indicating its effectiveness in predicting the prices. On the other hand, SLR has high MAE values for most coins except for XRP. XRP shows a relatively low MAE compared to other models through SLR. Both RF ensemble with XGBoost and averaging

Table 1. Descriptive statistics

Coins	Skewness	Kurtosis	Jarque-Bera
ETH_USD	1.148557	3.094957	381.8956
BNB_USD	1.073268	2.716866	338.6924
XRP_USD	2.295876	12.24073	7692.837
DOGE_USD	2.087496	7.819337	2937.438
TRX_USD	1.208705	4.660242	621.3699
LTC_USD	1.159686	4.068617	471.173
ETC_USD	1.676378	6.375899	1635.572
XLM_USD	1.186234	4.21978	514.1647
XMR_USD	0.94309	3.44275	271.205
FIL_USD	2.70873	11.43144	7256.654
MANA_USD	2.252884	7.757234	3101.925
ZEC_USD	2.687686	13.10948	9471.696
DASH_USD	3.352199	17.53833	18518.55
XEM_USD	4.245989	27.40811	48253.59
MONA_USD	2.835728	13.97474	11026.09

Note(s): Table showing Descriptive statistics of cryptocurrencies
Source(s): Table by authors

Table 2. Mean absolute error

Coin	ARIMA	SLR	RF	DT	GB	SVR	MLP	RF with XGBoost	Avg RF, GB and DT
XRP	98.288	0.264	0.019	0.024	0.021	0.108	0.183	0.05579	0.008
BNB	13.192	97.665	4.61	5.671	5.696	60.99	111.3	213.3256	44.659
MONA	0.031	0.713	0.054	0.069	0.062	0.312	0.579	25.63429	5.134
XLM	0.009	0.113	0.008	0.009	0.008	0.059	0.129	20.20679	4.259
DASH	0.003	95.75	5.879	7.211	6.531	56.31	84	0.01516	0.004
ETH	5.454	739	38.614	46.389	41.215	724.5	775.9	41.7622	6.805
XMR	2.068	65.669	4.916	6.257	5.879	35.54	61.33	0.01283	0.002
FIL	0.01	18.503	1.1	1.346	1.167	9.257	19.05	0.11622	0.023
LTC	8.742	51.08	3.8	4.499	4.229	35.34	50.07	3.88621	1.086
DOGE	2.294	0.06	0.004	0.004	0.004	0.07	0.077	24.22042	5.544
TRX	0.835	0.02	0.002	0.002	0.002	0.069	0.067	0.37456	0.062
XEM	6.395	0.111	0.006	0.008	0.007	0.062	0.065	23.80469	5.818
ZEC	6.212	68.536	4.934	5.842	5.572	35.45	59.54	0.04309	0.009
MANA	0.006	0.505	0.026	0.033	0.031	0.254	0.558	0.08772	0.029
ETC	0.038	13.167	1.035	1.162	0.975	7.478	11.62	6.77117	1.212

Note(s): Table showing the Mean Absolute Error of cryptocurrencies
Source(s): Table by authors

ensemble of RF, GB and DT show low MAE demonstrating that the model performs well for most cryptocurrencies.

5.3 Mean squared error

According to Table 3 on the MSE of cryptocurrency prices, the results vary in accordance with the performance of models across different coins. First, the RF model outperforms other models such as ARIMA, SLR, DT, GB, RF ensemble with XGBoost and averaging ensemble of RF, GB and DT for XRP, demonstrating the lowest MSE of 0.002. However, it is essential to note that it is unusual for the MLP model to show a negative MSE value for XRP. This might indicate an issue with the model or the data. The RF model consistently exhibits the lowest

Table 3. Mean squared error

Coin	ARIMA	SLR	RF	DT	GB	SVR	MLP	RF with XGBoost	Avg RF, GB and DT
XRP	18931	0.105	0.002	0.003	0.002	0.58	−1.67	0.0077	0.000
BNB	393.7	15799	117.392	197.296	213.091	0.09	−0.91	126782.8	6882.085
MONA	0.003	0.933	0.014	0.03	0.018	0.75	0.02	2553.567	140.136
XLM	0	0.019	0	0	0	0.62	−10.21	959.3232	61.044
DASH	0	14219	161.275	239.256	156.187	−4.5	−24.59	0.00127	0.000
ETH	99.239	815760.5	5419.305	8791.746	6711	−23.35	−0.79	3470.993	186.544
XMR	16.705	6763.324	77.521	142.697	117.923	−0.23	−1.27	0.0003	0.000
FIL	0	878.676	7.574	12.603	9.638	−1.23	−3.53	0.02645	0.002
LTC	191.46	4001.806	53.799	79.782	69.582	−2.23	−3.27	67.33126	14.105
DOGE	17.443	0.009	0	0	0	0.4	−52.48	1185.888	91.333
TRX	1.592	0.001	0	0	0	−593	−2.07	0.26202	0.017
XEM	108.51	0.021	0	0.001	0	0.67	−0.03	1243.139	112.427
ZEC	126.38	7103.623	97.882	139.455	112.793	−1.23	−11.31	0.00409	0.000
MANA	0	0.564	0.005	0.008	0.008	0.44	−5.23	0.03888	0.006
ETC	0.007	329.031	14.379	15.213	12.707	−0.93	−3.74	235.6092	9.808

Note(s): Table showing mean squared error of cryptocurrencies
Source(s): Table by authors

MSE values for BNB, MONA, XLM, DASH, XMR, FIL and TRX, indicating its superior accuracy when compared to other models. For ETH, LTC, ZEC, MANA and ETC, the MLP model, conversely, showcases the lowest MSE values among the models considered and stands out as the most accurate. It is worth mentioning that the SLR model shows a relatively low MSE value for XRP and performs poorly for most coins, resulting in a significant rise in MSE values compared to other models.

5.4 Root mean squared error

Table 4 results derive variations in the performance of models among different coins. The RF model outperforms other models such as ARIMA, SLR, DT, GB, SVR and MLP and achieves the lowest RMSE for XRP at 0.052. However, it is essential to note that the SLR model presents a relatively low RMSE for XRP. The RF model consistently exhibits the lowest RMSE values for BNB, MONA, XLM, DASH, XMR, FIL, TRX and ZEC, indicating its superior accuracy compared to other models. For ETH, LTC, MANA and ETC, the MLP model conversely stands out as the most accurate. It showcases the lowest RMSE values among the models considered. The ARIMA model shows relatively high RMSE values across the board, while the SVR model performs moderately well for most coins. Additionally, except for XRP and TRX, where it performs relatively better, the SLR model exhibits higher RMSE values for most coins. RF ensemble with XGBoost and Averaging ensemble of RF, GB and DT has relatively low RMSE demonstrating that the model performs well for most cryptocurrencies.

5.5 R-*squared*

The results of Table 5 show that the RF, DT and GB models have strong predictive power, due to high *R*-squared values across most coins. Their ability to account for a considerable proportion of the fluctuations in cryptocurrency prices is evident in these models. The ARIMA model has limited explanatory capability for cryptocurrency price movements, as it exhibits comparatively lower *R*-squared values for all coins. Some coins have relatively higher *R*-squared values through the SLR model, while others have significantly negative values.

Table 4. Root mean squared error

Coin	ARIMA	SLR	RF	DT	GB	SVR	MLP	RF with XGBoost	Avg RF, GB and DT
XRP	137.59	0.323	0.052	0.043	0.041	0.171	0.274	0.08777	0.020
BNB	19.842	125.7	14.046	14.598	10.835	109.593	136.629	356.0657	82.958
MONA	0.054	0.966	0.173	0.135	0.118	0.475	0.695	50.53283	11.838
XLM	0.02	0.137	0.021	0.017	0.016	0.075	0.151	30.97294	7.813
DASH	0.004	119.25	15.468	12.497	12.699	100.591	145.217	0.03559	0.015
ETH	9.962	903.2	93.764	81.922	73.616	1212.689	925.488	58.91513	13.658
XMR	4.087	82.239	11.946	10.859	8.805	56.477	85.682	0.01724	0.004
FIL	0.019	29.642	3.55	3.105	2.752	22.931	29.521	0.16264	0.044
LTC	13.837	63.26	8.932	8.342	7.335	49.155	69.146	8.20556	3.756
DOGE	4.177	0.094	0.015	0.014	0.016	0.085	0.106	34.43673	9.557
TRX	1.262	0.026	0.004	0.004	0.003	0.074	0.072	0.51188	0.129
XEM	10.417	0.143	0.024	0.02	0.019	0.079	0.103	35.25817	10.603
ZEC	11.242	84.283	11.809	10.62	9.894	59.686	97.813	0.06396	0.018
MANA	0.01	0.751	0.089	0.088	0.073	0.506	0.806	0.19719	0.080
ETC	0.084	18.139	3.9	3.565	3.792	15.187	17.31	15.34957	3.132

Note(s): Table showing root mean squared error of cryptocurrencies
Source(s): Table by authors

Table 5. *R*-Squared

Coin	ARIMA	SLR	RF	DT	GB	SVR	MLP	RF with XG Boost	Avg RF, GB and DT
XRP	0	−356.35	0.98	0.98	0.98	0.58	−1.67	0.25	0.980
BNB	0.03	0.14	0.99	0.99	1	0.09	−0.91	0.91	1.000
MONA	0.03	−1.44	0.98	0.97	0.99	0.75	0.02	0.92	1.000
XLM	0.01	−1029.32	0.98	0.98	0.99	0.62	−10.21	0.67	0.980
DASH	0.13	−1.55	0.99	0.98	0.99	−4.5	−24.59	0.87	0.980
ETH	0	−0.26	1	0.99	1	−23.35	−0.79	0.7	0.990
XMR	0.01	−15.39	0.98	0.98	0.99	−0.23	−1.27	0.45	0.980
FIL	0.01	−6.3	0.99	0.99	0.99	−1.23	−3.53	0.62	0.980
LTC	0.01	−71.37	0.98	0.98	0.99	−2.23	−3.27	0.78	0.960
DOGE	0.08	−1.89	0.98	0.98	0.98	0.4	−52.48	0.79	0.990
TRX	0.99	−3.05	0.99	0.98	0.99	−593.2	−2.07	0.64	0.980
XEM	0.01	−5.6	0.98	0.97	0.98	0.67	−0.03	0.75	0.980
ZEC	0.01	−5.44	0.98	0.98	0.99	−1.23	−11.31	0.7	0.980
MANA	0.05	−0.92	0.99	0.99	0.99	0.44	−5.23	0.95	0.990
ETC	0.02	−4.2	0.97	0.96	0.96	−0.93	−3.74	0.68	0.990

Note(s): Table showing *R* Squared of cryptocurrencies
Source(s): Table by authors

Hence, the results are mixed. It is important to note that the SLR model performs worse than a horizontal line (mean) in fitting the data when *R*-squared values are negative. This implies that the relationships between those specific coins' dependent and independent variables cannot be captured adequately by the SLR model. There is varied performance across different coins with the MLP model as well. It achieves high *R*-squared values for some coins, such as BNB and ETH, and presents negative *R*-squared values for others. This suggests that for those particular coins, the MLP model may not effectively capture the underlying patterns in cryptocurrency prices. For RF ensemble with XGBoost and averaging ensemble model with RF, GB and DT show reasonably high R2 values indicate that the model can explain a considerable percentage of the data variation. However, when certain cryptocurrencies, such as XRP, have lower R2 values for RF ensemble with XGBoost, it is inferred that the model cannot explain a large variation in the data for this coin. Thus, this model's prediction accuracy is quite poor for this currency. The averaging ensemble model with RF, GB and DT model has R2 values are typically near one, indicating that the model can accurately predict the target variable.

5.6 Multi-model regression (MLR) model

Table 6 and Figure 2 shows the forecasting results of MLR for pre and post-COVID crisis. The results show that the model consistently achieves exceptional performance across all coins. Extremely low values of MAE, MSE and RMSE indicate its high accuracy in predicting cryptocurrency prices covering pre and post-crisis. Additionally, it can be concluded that the model can perfectly explain the variance in cryptocurrency prices, indicating its strong explanatory power, given that the R2 value of all coins is 1.00. The significant performance of the MLR in accurately forecasting cryptocurrency prices across different coins, as the table highlights, is overall indicative of its effectiveness.

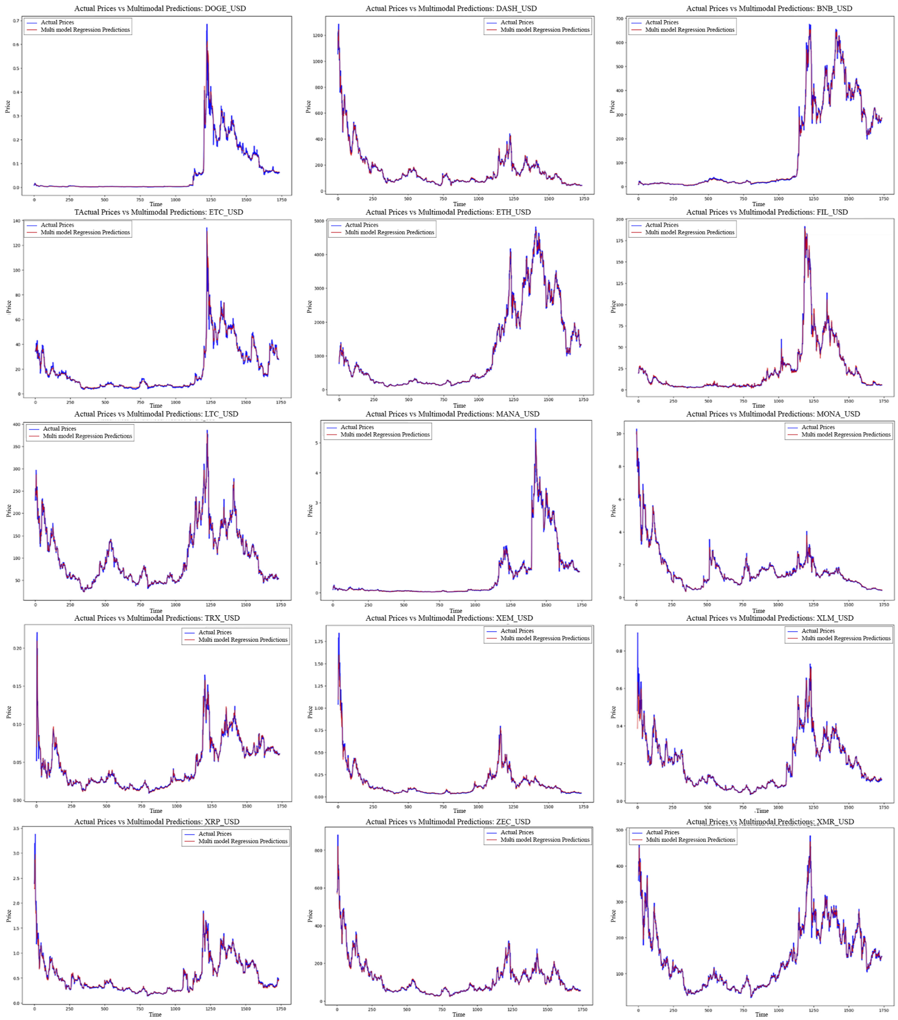
6. Findings

The descriptive statistics analysis, helped understand the nature of cryptocurrency price distributions. Cryptocurrencies are known for their high volatility and the presence of volatility clustering, as noted in the literature review. Aligning with findings from the previous

Table 6. Multi-model regression model

Coin	Pre MAE	Pre MSE	Pre RMSE	Pre R2	Post MAE	Post MSE	Post RMSE	Post R2	Con MAE	Con MSE	Con RMSE	Con R2
XRP	0.02	0.001	0.035	0.991	0.007	0.000	0.010	0.999	1.035	8.828	2.971	1
BNB	0.252	0.151	0.388	0.997	3.267	20.167	4.491	0.998	0.198	0.381	0.617	1
MONA	0.069	0.014	0.119	0.995	0.006	0.000	0.011	0.999	0.001	0	0.003	1
XLM	0.008	0	0.015	0.987	0.002	0.000	0.003	0.999	0.002	0	0.004	0.99
DASH	7.662	198.404	14.086	0.995	1.173	3.345	1.829	0.999	0	0	0	1
ETH	7.223	181.374	13.468	0.997	24.651	1107.4	33.277	0.999	0.334	0.439	0.663	1
XMR	3.372	34.607	5.883	0.995	1.950	7.281	2.698	0.997	0.295	0.266	0.516	0.99
FIL	0.19	0.108	0.329	0.996	0.375	0.432	0.657	0.999	0	0	0.002	1
LTC	2.371	13.674	3.698	0.994	1.231	3.650	1.910	0.999	0.124	0.095	0.308	1
DOGE	0	0	0	0.991	0.001	0.000	0.002	0.999	0.067	0.041	0.203	1
TRX	0.001	0	0.003	0.978	0.001	0.000	0.001	0.996	0.005	0	0.012	1
XEM	0.013	0.001	0.029	0.987	0.001	0.000	0.002	0.999	0.425	0.626	0.791	1
ZEC	5.533	118.976	10.908	0.993	1.525	5.479	2.341	0.998	0.786	1.722	1.312	1
MANA	0.002	0	0.004	0.989	0.031	0.003	0.052	0.998	0.001	0	0.003	1
ETC	0.263	0.364	0.603	0.994	0.370	0.300	0.550	1.000	0.004	0	0.009	1

Note(s): Table showing multi-model regression model for all cryptocurrencies
Source(s): Table by authors



Source(s): Figure by authors

Figure 2. Figure showing the multi-modal regression model performance for all cryptocurrencies

research, the skewness and kurtosis values indicate deviations from a normal distribution (Nikou *et al.*, 2019). This supports the need for robust forecasting models capable of handling such volatile and skewed data.

ARIMA, SLR, RF, DT, GB and MLR ensemble models performed well with coins such as XLM, FIL, MANA, DASH, XMR and ZEC. These models adequately reflect the above coins' trends, seasonality and historical price patterns. Furthermore, MLR demonstrates its advantages when applied to coins like BNB and MONA. SLR has performed well for cryptocurrencies that have significant market capitalization, strong liquidity and price movements that have historically been consistent, like XRP, TRX and DOGE. The MLR ensemble approach produces more accurate predictions for coins with higher volatility and irregular price patterns, such as ETH and ETC, by combining ML models to remove errors.

It is worth noting that the intricacies of each coin's behavior may not be captured as these are broad observations. Macroeconomic, attractiveness and demand and supply variables are all factors that can impact a model's efficacy. The literature review expresses concern regarding cryptocurrency price forecasting challenges, echoed by the findings, which demonstrate that different models perform variably across different cryptocurrencies. The analysis of the RF ensembles with XGBoost and averaging ensemble model with RF, GB and DT demonstrates the effectiveness of ensemble models combining multiple models for improved forecasting accuracy. The results support that ensemble models can enhance predictive performance. They often outperformed individual models across various cryptocurrencies in the current study.

7. Conclusion

Cryptocurrency price forecasting carries a significant influence on investment decisions. The present study makes contribution to the theoretical and practical implications on cryptocurrency investments. On the theoretical side, the study advances the conceptual framework on building predictive models by demonstrating the efficiency of ensemble model in forecasting. It shows a new dimension to predictive analytics by empirically validating the accuracy of multiple algorithms pre-COVID and post-COVID period. On the practical side, it offers a valuable forecasting tool to the traders and investors in making informed decision. The novelty of the research is in its attempt to empirically test the forecasting ability of ensemble model across time span and its comparison with econometric model on a broad-based set of cryptocurrencies. Consistent with the expectations, ensemble models created by combining RF, XGBoost, gradient boosting and decision tree (DT) reduce the variance in predicting errors indicating the forecasting accuracy. Interestingly, it is found that when the price movement exhibits autocorrelation, the ARIMA model is best applicable. Compared to other ML models applied in isolation the ensemble model performed better.

7.1 Practical implications

The present study contributes to the body of literature on ensemble methods in forecasting crypto price in general, potentially influencing future studies on price forecasting. The study motivates the researchers on empirical testing of our framework on various asset classes. As a result, on the prediction ability of ensemble model, the study will significantly influence the decision-making process of traders and investors. The research benefits the traders and investors to effectively develop a model to forecast cryptocurrency price. The findings highlight the potential of ensemble model in predicting high volatile cryptocurrencies and other financial assets. Investors can design the investment strategies and asset allocation decisions by understanding the relationship between market trends and consumer behavior. Investors can enhance portfolio performance and mitigate risk by incorporating these insights into their decision-making processes. Policymakers can use this information to design more effective regulations and policies promoting economic stability and consumer welfare. The study emphasizes the need for using diversified model to understand the market dynamics and improving trading strategies.

7.2 Limitations

Although the dataset includes crisis period data, anomalies or outliers are yet to be explicitly excluded from the analysis. The models employed in this study still demonstrate high accuracy in predicting cryptocurrency prices despite these outliers, suggesting that the models are robust enough to handle unexpected fluctuations or extreme events in the market. However, the lack of specific analysis on the impact of outliers on model performance is a limitation of the study, as it needs to fully explore the resilience of the forecasting models under adverse market conditions.

7.3 Future research

Cryptocurrency markets can be influenced by investor sentiment, which can be reflected in price movements. By analyzing these price movements, investors might gain insights into overall market sentiment. However, exploring this link further is certainly an area for future research. While our study focuses on price prediction using the ensemble model, we acknowledge the need for further exploration to definitively link price movements to specific investor behaviors. Complementary techniques, such as sentiment analysis of social media data or transaction volume analysis, could be valuable tools for future research in this area.

References

Adegboruwa, T.I., Adeshina, S.A. and Boukar, M.M. (2019), "Time series analysis and prediction of bitcoin using long short term memory neural network", *2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*, pp. 1-5, doi: 10.1109/ICECCO48375.2019.9043229.

Alahmari, S.A. (2019), "Using machine learning ARIMA to predict the price of Cryptocurrencies$", Vol. 11 No. 3, p. 6.

Bouteska, A., Abedin, M.Z., Hajek, P. and Yuan, K. (2024), "Cryptocurrency price forecasting–a comparative analysis of ensemble learning and deep learning methods", *International Review of Financial Analysis*, Vol. 92, 103055, doi: 10.1016/j.irfa.2023.103055.

Breiman, L. (2001), "Random forests", *Machine Learning*, Vol. 45 No. 1, pp. 5-32, doi: 10.1023/A:1010933404324.

Breiman, L. (2017), *Classification and Regression Trees*, Routledge, Boca Raton.

Chen, J. (2023), "Analysis of bitcoin price prediction using machine learning", *Journal of Risk and Financial Management*, Vol. 16 No. 1, p. 51, doi: 10.3390/jrfm16010051.

Chevallier, J., Guégan, D. and Goutte, S. (2021), "Is it possible to forecast the price of bitcoin?", *Forecasting*, Vol. 3 No. 2, pp. 377-420, doi: 10.3390/forecast3020024.

Cohn, D.A., Ghahramani, Z. and Jordan, M.I. (1996), "Active learning with statistical models", *Journal of Artificial Intelligence Research*, Vol. 4, pp. 129-145, doi: 10.1613/jair.295.

Derbentseva, V., Khrustalovac, S., Babenko, V., Khrustalev, K. and Obruch, H. (2021), "Comparative performance of machine learning ensemble algorithms for forecasting cryptocurrency prices", *International Journal of Engineering*, Vol. 34 No. 1, doi: 10.5829/ije.2021.34.01a.16.

El-Bannany, M., Sreedharan, M. and Khedr, A.M. (2020), "A robust deep learning model for financial distress prediction", *International Journal of Advanced Computer Science and Applications (IJACSA), The Science and Information (SAI) Organization Limited*, Vol. 11 No. 2, doi: 10.14569/IJACSA.2020.0110222.

Greaves, A. and Au, B. (2015), "Using the bitcoin transaction graph to predict the price of bitcoin", p. 8.

Hamayel, M.J. and Owda, A.Y. (2021), "A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms", *AI, Multidisciplinary Digital Publishing Institute*, Vol. 2 No. 4, pp. 477-496, doi: 10.3390/ai2040030.

Ho, T.K. (1995), "Random decision forests", *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1)–Volume 1*, USA, IEEE Computer Society, p. 278.

Hunt, E.B., Marin, J. and Stone, P.J. (1966), *Experiments in Induction*, Academic Press, Massachusetts.

Kass, G.V. (1980), "An exploratory technique for investigating large quantities of categorical data", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Vol. 29 No. 2, pp. 119-127, doi: 10.2307/2986296.

Kayani, U.N. (2023), "Exploring prospects of blockchain and fintech: using SLR approach", *Journal of Science and Technology Policy Management*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/JSTPM-01-2023-0005.

Kayani, U. and Hasan, F. (2024), "Unveiling cryptocurrency impact on financial markets and traditional banking systems: lessons for sustainable blockchain and interdisciplinary collaborations", *Journal of Risk and Financial Management*, Vol. 17 No. 2, p. 58, doi: 10.3390/jrfm17020058.

Khedr, A.M., Arif, I., El-Bannany, M., Alhashmi, S.M. and Sreedharan, M. (2021), "Cryptocurrency price prediction using traditional statistical and machine-learning techniques: a survey", *Intelligent Systems in Accounting, Finance and Management*, Vol. 28 No. 1, pp. 3-34, doi: 10.1002/isaf.1488.

Lahmiri, S. and Bekiros, S. (2019), "Cryptocurrency forecasting with deep learning chaotic neural networks", *Chaos, Solitons and Fractals*, Vol. 118, pp. 35-40, doi: 10.1016/j.chaos.2018.11.014.

Li, Y., Jiang, S., Li, X. and Wang, S. (2022), "Hybrid data decomposition-based deep learning for Bitcoin prediction and algorithm trading", *Financial Innovation*, Vol. 8 No. 1, p. 31, doi: 10.1186/s40854-022-00336-7.

Livieris, I.E., Kiriakidou, N., Stavroyiannis, S. and Pintelas, P. (2021), "An advanced CNN-LSTM model for cryptocurrency forecasting", *Electronics*, Vol. 10 No. 3, p. 287, doi: 10.3390/electronics10030287.

Lobell, D.B. and Burke, M.B. (2010), "On the use of statistical models to predict crop yield responses to climate change", *Agricultural and Forest Meteorology*, Vol. 150 No. 11, pp. 1443-1452, doi: 10.1016/j.agrformet.2010.07.008.

Lyu, H. (2022), "Cryptocurrency price forecasting: a comparative study of machine learning model in short-term trading", *2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML)*, pp. 280-288, doi: 10.1109/CACML55074.2022.00054.

Madeh, P.S. and El-Diraby, T.E. (2021), "Using machine learning to examine impact of type of performance indicator on flexible pavement deterioration modeling", *Journal of Infrastructure Systems*, Vol. 27 No. 2, 04021005, doi: 10.1061/(ASCE)IS.1943-555X.0000602.

Madan, I., Saluja, S. and Zhao, A. (2015), "Automated bitcoin trading via machine learning algorithms", p. 5.

Meenu, S., Khedr, A.M. and El Bannany, M. (2020), "A multi-layer perceptron approach to financial distress prediction with genetic algorithm", *Automatic Control and Computer Sciences*, Vol. 54 No. 6, pp. 475-482, doi: 10.3103/S0146411620060085.

Naing, W.Y.N. and Htike, Z.Z. (2015), "Forecasting of monthly temperature variations using random forests", *ARPN Journal of Engineering and Applied Sciences*, Vol. 10 No. 21, pp. 10109-10112.

Nakano, M., Takahashi, A. and Takahashi, S. (2018), "Bitcoin technical trading with artificial neural network", *Physica A: Statistical Mechanics and its Applications*, Vol. 510, pp. 587-609, doi: 10.1016/j.physa.2018.07.017.

Nikou, M., Mansourfar, G. and Bagherzadeh, J. (2019), "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms", *Intelligent Systems in Accounting, Finance and Management*, Vol. 26 No. 4, pp. 164-174, doi: 10.1002/isaf.1459.

Patel, M.M., Tanwar, S., Gupta, R. and Kumar, N. (2020), "A deep learning-based cryptocurrency price prediction scheme for financial institutions", *Journal of Information Security and Applications*, Vol. 55, 102583, doi: 10.1016/j.jisa.2020.102583.

Piryonesi, S.M. and El-Diraby, T.E. (2020), "Data analytics in asset management: cost-effective prediction of the pavement condition index", *Journal of Infrastructure Systems*, Vol. 26 No. 1, 04019036, doi: 10.1061/(ASCE)IS.1943-555X.0000512.

Quinlan, J.R. (1983), "Learning efficient classification procedures and their application to chess end games", *Machine Learning*, pp. 463-482, doi: 10.1007/978-3-662-12405-5_15.

Quinlan, J.R. (1986), "Induction of decision trees", *Machine Learning*, Vol. 1, pp. 81-106, doi: 10.1007/bf00116251.

Rosenblatt, F. (1958), "The perceptron: a probabilistic model for information storage and organization in the brain", *Psychological Review*, Vol. 65 No. 6, pp. 386-408, doi: 10.1037/h0042519.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986), "Learning representations by back-propagating errors", *Nature*, Vol. 323 No. 6088, pp. 533-536, doi: 10.1038/323533a0.

Soltani, H., Taleb, J. and Abbes, M.B. (2023), "The directional spillover effects and time-frequency nexus between stock markets, cryptocurrency, and investor sentiment during the COVID-19 pandemic", *European Journal of Management and Business Economics*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/EJMBE-09-2022-0305.

Suhartono, Rahayu, S.P., Prastyo, D.D., Wijayanti, D.G.P. and Juliyanto (2017), "Hybrid model for forecasting time series with trend, seasonal and salendar variation patterns", *Journal of Physics: Conference Series*, Vol. 890, 012160, doi: 10.1088/1742-6596/890/1/012160.

Tandon, S., Tripathi, S., Saraswat, P. and Dabas, C. (2019), "Bitcoin price forecasting using LSTM and 10-fold cross validation", *2019 International Conference on Signal Processing and Communication (ICSC)*, pp. 323-328, doi: 10.1109/ICSC45622.2019.8938251.

Vidal-Tomás, D. (2022), "Which cryptocurrency data sources should scholars use?", *International Review of Financial Analysis*, Vol. 81, 102061, doi: 10.1016/j.irfa.2022.102061.

Wang, P., Li, X., Shen, D. and Zhang, W. (2020), "How does economic policy uncertainty affect the bitcoin market?", *Research in International Business and Finance*, Vol. 53, 101234, doi: 10.1016/j.ribaf.2020.101234.

Wirawan, I.M., Widiyaningtyas, T. and Hasan, M.M. (2019), "Short term prediction on bitcoin price using ARIMA method", *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, IEEE, pp. 260-265, doi: 10.1109/ISEMANTIC.2019.8884257.

Yang, Z., Li, L., Xu, X., Kailkhura, B., Xie, T. and Li, B. (2022), "On the certified robustness for ensemble models and beyond", *arXiv*, doi: 10.48550/arXiv.2107.10873, available at: http://arxiv.org/abs/2107.10873 (accessed 1 July 2023).

Yin, L., Nie, J. and Han, L. (2021), "Understanding cryptocurrency volatility: the role of oil market shocks", *International Review of Economics and Finance*, Vol. 72, pp. 233-253, doi: 10.1016/j.iref.2020.11.013.

About the authors

N.S.S. Kiranmai Balijepalli is Research Scholar at the Symbiosis Institute of Business Management, a constituent of Symbiosis International (Deemed) University in Bengaluru, India. Her research focuses on the intersection of technology and finance, utilizing machine learning techniques to predict future cryptocurrency prices and detect market manipulations. Kiranmai aims to advance the understanding and development of predictive models and analytical tools in the dynamic field of cryptocurrency. She has published her work in international journals and presented papers at international conferences, contributing significantly to the academic and professional discourse on cryptocurrency markets. N.S.S. Kiranmai Balijepalli is the corresponding author and can be contacted at: kiranmai@sibm.edu.in

Dr Viswanathan Thangaraj is Associate Professor at the Symbiosis Institute of Business Management, a constituent of Symbiosis International (Deemed) University in Bengaluru, India. Holding a Doctorate Degree, he is also a corporate trainer specializing in Finance and Data Science. Dr. Thangaraj's areas of expertise include data science, ML, artificial intelligence, financial econometrics, financial analytics and corporate finance. He has published several research papers in international journals and presented his work at conferences in India, Singapore, Malaysia and Thailand. His research focuses on integrating technology and finance through advanced analytical techniques.