International Review of Financial Analysis 92 (2024) 103055

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods

Ahmed Bouteska, Mohammad Zoynul Abedin *, Petr Hajek, Kunpeng Yuan

[a] Finance, Faculty of Economics and Management of Tunis, University of Tunis El Manar, Tunisia
[b] School of Management, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, Wales, United Kingdom
[c] Science and Research Centre, Faculty of Economics and Administration, University of Pardubice, Studentska 84, 532 10 Pardubice, Czech Republic
[d] School of Economics and Management, Dalian University of Technology, No.2, Linggong Road, Liaoning, 116024 Dalian, China

ARTICLE INFO

Keywords:
Cryptocurrency
Bitcoin
Forecasting
Ensemble learning
Deep learning
Neural networks

ABSTRACT

Cryptocurrency price forecasting is attracting considerable interest due to its crucial decision support role in investment strategies. Large fluctuations in non-stationary cryptocurrency prices motivate the urgent need for accurate forecasting models. The lack of seasonal effects and the need to meet a number of unrealistic requirements make it difficult to make accurate forecasts using traditional statistical methods, leaving machine learning, particularly ensemble and deep learning, as the best technology in the area of cryptocurrency price forecasting. This is the first work to provide a comprehensive comparative analysis of ensemble learning and deep learning forecasting models, examining their relative performance on various cryptocurrencies (Bitcoin, Ethereum, Ripple, and Litecoin) and exploring their potential trading applications. The results of this study reveal that gated recurrent unit, simple recurrent neural network, and LightGBM methods outperform other machine learning methods, as well as the naive buy-and-hold and random walk strategies. This can effectively guide investors in the cryptocurrency markets.

1. Introduction

The area of cryptocurrencies has attracted growing attention from investors and regulators since Bitcoin was introduced in 2008 (Corbet et al., 2019). This growing popularity of cryptocurrencies is related to their different characteristics from other traditional financial assets. Their value is based on the confidence of the underlying algorithm, rather than on any tangible asset, allowing cryptocurrencies to be independent of any higher authority. This is what eventually leads to low transaction costs and government-independent secure peer-to-peer payments.

Extant research recognizes cryptocurrencies as an investment asset (Bouri et al., 2017; Corbet et al., 2019; Ji et al., 2018). In response, a nascent strand of literature has appeared to explore the potential synergies between cryptocurrencies and other investment assets, such as commodities (Das et al., 2020), equities (Jiang et al., 2021), and conventional currencies (Shahzad et al., 2022). Notably, Guesmi et al. (2019) underlined how Bitcoin allows hedging investment strategy against various investment assets, including gold, oil, and equities, due to its high return and low correlation with the other investment assets. Therefore, cryptocurrencies provide investors with diversification and hedging opportunities. As of 2022, there are >20,000 cryptocurrencies, but only the top 20 account for nearly 90% of the total market. The global cryptocurrency market capitalization was 1.06 trillion USD, with >300 million cryptocurrency users around the world in 2022 (Tuwiner, 2022).

Large price fluctuations of cryptocurrencies generate huge profit opportunities for high-frequency traders, including algorithmic trading bots (Chu et al., 2019; Chu et al., 2020; Patel et al., 2015). It is estimated that more than half of the trading volume is accounted for by these bots, making it increasingly difficult for human traders to make profit when trading during short periods (Ibrahim et al., 2021). These bots are aided by increasingly complex machine learning methods, frequently backed by deep learning (Rahmani Cherati et al., 2021).

The purpose of developing cryptocurrency price forecasting systems is to develop a model that can guide the algorithmic/human trader in trading decisions to increase the chances of making profits when trading cryptocurrencies. Different cryptocurrency price forecasting methods can be divided into traditional statistical methods and machine learning methods (Chen et al., 2021).

* Corresponding author.
E-mail address: m.z.abedin@swansea.ac.uk (M.Z. Abedin).

https://doi.org/10.1016/j.irfa.2023.103055
Received 17 December 2022; Received in revised form 28 November 2023; Accepted 18 December 2023
Available online 27 December 2023
1057-5219/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Early work in this area focused primarily on traditional statistical methods, such as ARIMA (autoregressive integrated moving average) (Ibrahim et al., 2021) and GARCH (Baur et al., 2018; Fakhfekh and Jeribi, 2020). However, these approaches only capture linear patterns in the time series of cryptocurrencies and furthermore assume a normal distribution of variables, which is unrealistic in the case of cryptocurrencies (Chen et al., 2021; Khedr et al., 2021).

Machine learning approaches can extract nonlinear patterns and also benefit from large datasets without assuming any prior understanding of the data. However, even traditional machine learning methods, such as multilayer perceptron (MLP) neural networks (Kristjanpoller and Minutolo, 2018) or support vector machines (SVM) (Hajek et al., 2023; Moula et al., 2017), suffer from some problems such as susceptibility to overfitting and do not fully exploit the potential of extracting high-level hidden patterns from cryptocurrency sequential data. To overcome these problems, deep learning-based forecasting models have been used, having the capacity to outperform traditional machine learning methods (Chen et al., 2021; Cui et al., 2022; Liu et al., 2021; Ortu et al., 2022). The recent work of Murray et al. (2023) substantiates this finding, demonstrating that long short-term memory (LSTM) and gated recurrent unit (GRU) neural networks outperform various other statistical and machine learning methods in terms of forecasting error. This includes not only traditional models such as ARIMA and SVM but also the more contemporary temporal fusion transformer (TFT). Another stream of research has focused on the capacity of ensemble learning approaches to reduce variance and bias by combining a set of diverse weak learning models (Aggarwal et al., 2020). In their widely acclaimed work, Sun et al. (2020) showed that ensemble learning forecasting models outperform individual machine learning models and that gradient boosting demonstrates better accuracy and robustness compared with the well-known random forest approach.

Not only does the existing literature fail to provide a comprehensive comparison of the latest machine learning methods, but previous studies also suggest that different methods may perform differently for different cryptocurrencies (Yang et al., 2020; Zhang et al., 2021). Moreover, no comparative study has been found that examines the financial performance of cryptocurrency investors from the perspective of different trading strategies over different time periods. To bridge this gap, this work aims to assess the performance of state-of-the-art deep learning and ensemble learning approaches in forecasting the prices of four major cryptocurrencies, namely Bitcoin, Ethereum, Ripple, and Litecoin. The selection of these four cryptocurrencies is not only consistent with those in previous related studies (Altan et al., 2019; Cheng, 2023) but it also covers a wide range of technologies, applications and market positions, making them ideal subjects for comprehensive analysis and forecasting. In particular, predicting the price of Ethereum can provide insights into a wider range of blockchain applications, predicting the price of Ripple can provide valuable insights into the integration of cryptocurrency technologies into traditional banking systems, and predicting the price of Litecoin alongside Bitcoin can reveal how changes in blockchain technology affect cryptocurrency market performance. This diversity ensures that the study's findings will be broadly relevant and provide valuable insights into the dynamics of the cryptocurrency sector. Furthermore, in contrast to earlier research that has tended to evaluate the forecasting performance in terms of forecasting errors (Chen et al., 2021; Murray et al., 2023), here we focus on investor performance by simulating the buy & sell, long and short trading strategies. To this end, two distinct sub-periods are considered in this study, before and after Covid-19. Indeed, the Covid-19 pandemic has had a significant impact on the cryptocurrency market, including changes in market efficiency, peak performance of some cryptocurrencies, and increases in market capitalization (El Montasser et al., 2022; Jalan et al., 2021). To this line of research, this study adds an analysis of the predictability of cryptocurrencies in the pre- and post-pandemic period. Unlike previous comparative studies that have focused only on the combinations of deep learning models (Murray et al., 2023), this paper also examines the performance of the hybrid two-step forecasting method (Efat et al., 2022) that combines ARIMA with deep learning methods, thus capturing linear and nonlinear patterns in the cryptocurrency time-series data. More importantly, the comparative analysis includes the financial performance of trading strategies based on the machine learning methods utilized. Thus, this work provides valuable insights into the performance of different machine learning models for predicting cryptocurrency prices, and their potential applications in trading strategies. The results suggest that these models can enable investors to make more informed decisions in the cryptocurrency markets, ultimately leading to better investment outcomes.

The remainder of the article is organized as follows. Section 2 provides an overview of previous literature on cryptocurrency price forecasting. Section 3 presents the research methodology employed and Section 4 shows the results. This is followed by Section 5, which discusses the results. Section 6 concludes the study with some future research directions.

2. Literature review

In theory, the value of cryptocurrencies is a reflection of their utility as a medium of exchange, which considerably increased over the last ten years. Given the increasing importance of cryptocurrencies for financial systems, early work in this area focused primarily on cryptocurrency volatilities, which have proved to be large (Klein et al., 2018) and difficult to predict so far (Fang et al., 2020; Walther et al., 2019). Moreover, empirical evidence also suggests that (Zhang et al., 2018): (1) cryptocurrency returns have heavily tailed distributions, (2) autocorrelations for relative and absolute returns decay at different rates; (3) cryptocurrencies exhibit a strong leverage effect and volatility clustering; (4) volatility and returns show the long-range dependence; and (5) volatility and price are power-law correlated. These characteristics make cryptocurrency price forecasting challenging and investments in cryptocurrencies much riskier than investments in traditional financial assets. Fluctuations in the value of cryptocurrency assets have been difficult to predict because they are not related to any fundamentals, which leads to the hypothesis that the value is mainly influenced by the sentiment of the cryptocurrency market. As shown in the literature, the price of Bitcoins and many other cryptocurrencies has displayed cyclicality patterns (also referred to as bubbles) in recent years (Dong et al., 2022; Kyriazis et al., 2020).

Most of the research on cryptocurrency price forecasting has focused on conventional statistical methods. Catania et al. (2019) used a battery of univariate and multivariate vector autoregression (VAR) models for predicting four major cryptocurrencies: Bitcoin, Ripple, Litecoin, and Ethereum. Notably, significant improvements in forecasting accuracy were reported for the combinations of various univariate forecasting models. Conrad et al. (2018) analyzed the volatility of cryptocurrencies through the lens of GARCH-MIDAS model to extract the long and short-term volatility components, finding that S&P 500 volatility significantly affected long-term Bitcoin volatility. Likewise, Walther et al. (2019) applied the GARCH-MIDAS framework to forecast the volatilities of five highly capitalized cryptocurrencies as well as the CRIX cryptocurrency index, investigating the effect of Global Real Economic Activity as a major driver of long-term cryptocurrency volatility. Results reported by Walther et al. (2019) also suggest that the traditional GARCH model performs poorly in predicting cryptocurrency volatility during bear markets, being surpassed even by models based on individual exogenous variables.

Over the last five years, the focus of cryptocurrency price forecasting has shifted to machine learning methods. The work of Kristjanpoller and Minutolo (2018) has made a significant contribution to the field by proposing a hybrid MLP neural network-GARCH model to forecast the price volatility of Bitcoin. The results of a thorough analysis of different GARCH models revealed the benefits of combining linear and nonlinear models for predicting Bitcoin price volatility. MLP neural network was also employed by Nakano et al. (2018) for predicting Bitcoin returns based on a set of technical indicators. Experimental evidence showed that the MLP forecasting model outperforms the baseline buy-and-hold strategy. MLP also performed well when comparing its movement direction performance against ARIMA, Prophet, and random forest (Ibrahim et al., 2021). More recently, recurrent neural networks have been utilized, such LSTM and GRU, to automatically extract high-level temporal patterns from cryptocurrency time series. These advanced neural networks with deep learning were specifically developed to handle complex sequential data and, therefore, it was not surprising that MLP and other conventional machine learning methods were outperformed by LSTM in several studies (Chen et al., 2021; Lahmiri and Bekiros, 2019; Li and Dai, 2020).

GRU also produced excellent forecasting performance for four major cryptocurrency prices (Zhang et al., 2021), outperforming not only traditional machine learning methods but also LSTM-based models. However, these deep learning-based models have been shown to work effectively, especially in univariate settings (Uras et al., 2020) because they are not equipped with a feature selection component and therefore can easily become too complex to learn more challenging temporal patterns (Fu et al., 2022).

Ensemble learning methods represent a viable alternative to deep learning models due to their capacity to reduce the bias (boosting methods) or variance (bagging methods such as random forest) of individual machine learning methods (Derbentsev et al., 2020). The model based on LightGBM (light gradient boosting machine) demonstrated the capacity to outperform the random forest model in forecasting the price direction of the cryptocurrency market (Sun et al., 2020), thus suggesting that bias reduction is more relevant in the case of cryptocurrency prices than variance reduction. Overall, the above studies indicate that the machine learning-based forecasting models outperform those using conventional statistical methods. This is attributed to the capacity of machine learning models to construct generic models easily capturing nonlinear complex patterns in cryptocurrency data. Recently, there have been two attempts to systematically review the performance of machine learning methods for cryptocurrency price forecasting (Khedr et al., 2021; Ren et al., 2022). Khedr et al. (2021) concluded that LSTM is considered to be the best method for predicting cryptocurrency price time series due to its ability to recognize long-term time-series associations. Ren et al. (2022) also valued the predictive performance of LSTM while highlighting that combining different machine learning methods has now become a hot research area. While these survey studies focus on providing an overview of existing machine learning methods used for cryptocurrency price forecasting, this study seeks to conduct a comparative empirical analysis of state-of-the-art deep learning and ensemble learning methods to provide support for profitable algorithmic trading.

Algorithmic trading has been actively developing in recent decades due to a combination of factors: the rapid development of machine learning methods, the development of technologies for working with data and its analysis, the growth of storage and processing capabilities for large amounts of data. In addition, the complexity of trading system algorithms used by market participants is growing, since they compete not only with those who do not use automated systems, but also with each other. In connection with these trends, the study of the applicability of various machine learning algorithms to algorithmic trading problems is an urgent task. This is important not only for companies engaged in algorithmic trading, such as hedge funds, but also from the scientific community because the application of state-of-the-art machine learning algorithms to the area under consideration can bring new knowledge to the development of automated trading systems for cryptocurrency markets. This paper is devoted to the application of ensemble learning and deep learning to forecast cryptocurrency prices. In cryptocurrency market trading, both the base predictors in ensembles and neural networks with deep learning mimic the actions of trading agents on the cryptocurrency market. This study was carried out to investigate the relevance of ensemble learning and deep learning for automatic cryptocurrency trading.

3. Research methodology

In this section, the methods used for the construction of forecasting models are introduced, together with their specifications. The machine learning methods employed in this study include boosting-based ensemble methods, recurrent deep neural networks, and hybrid two-stage methods integrating ARIMA with recurrent deep neural networks.

3.1. Boosting-based ensemble methods

Given that bagging-based ensemble methods, including random forest, have not performed well in earlier research (Ibrahim et al., 2021; Sun et al., 2020), we decided to examine the performance of boosting-based ensemble methods in the current study. The ultimate aim of boosting is to enhance the accuracy of a sequence of weak prediction models, where each model in the sequence compensates for the errors of its predecessors. As a result, a strong model is produced representing a highly accurate combination of weak models. This approach not only proved to be effective compared with individual and other ensemble learning methods, but also outperformed deep learning models in recent investigations (Manchanda and Aggarwal, 2021). Noteworthy, AdaBoost, a traditional boosting approach, exceeded the forecasting performance of LSTM and other machine learning methods, including MLP and ELM (extreme learning machines) (Manchanda and Aggarwal, 2021).

The idea of AdaBoost is that the weights of the data instances that are accurately predicted by the preceding weak regressor are decreased while the weights of the instances where forecasts deviated from the actual cryptocurrency prices are increased. Thus, successive forecasting models increasingly focus on poorly forecasted data instances, and the performance of the overall model is iteratively improved. In other words, AdaBoost generates an additive model while the value of loss function (bias) is reduced in each iteration.

LightGBM is an enhanced version of AdaBoost, allowing for the computationally efficient minimization of an arbitrary differentiable loss function. Similarly, as AdaBoost, regression trees are employed as weak learners in LightGBM. In contrast, the fast and highly efficient training capacity of LightGBM allows for dealing with large datasets. This is enabled by exploiting the exclusive feature bundling (into a single feature and thus reducing data dimensionality) and gradient-based one-side sampling (by randomly dropping instances with small gradients). At the same time, the advantages of the well-known XGBoost (extreme gradient boosting) are retained, including parallel training, sparse optimization, multiple loss functions, early stopping, and regularization. The main difference is that LightGBM grows regression trees leafwise, and not level-wise like traditional boosting methods. The objective function of LightGBM is defined as follows:

$$G = {}^1\!/_2\left(\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} - \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda}\right) \tag{1}$$

where I_R and I_L are the sets of instances of the right and left branches, respectively; g_i and h_i represent the loss gradient statistics of the first and second order, respectively; and λ is a regularization parameter.

3.2. Recurrent deep neural networks

Recurrent neural networks (RNNs) are types of neural networks in which links between units generate a controlled sequence, which allows for processing sequential data. In contrast to MLP, RNN can process arbitrary length sequences with its internal memory. Various RNN architectures, ranging from simple to complex, have been introduced. For cryptocurrency price forecasting, the LSTM and GRU neural networks are the most widely used. RNNs, equipped with a self-feedback mechanism, have the capacity to handle long-term dependencies in cryptocurrency time-series data. The vanishing gradient represents a major limitation of RNNs. To overcome this problem, LSTM neural networks were introduced (Yu et al., 2019). Each unit of LSTM is composed of memory cells that store information updated through the input, forget, and output gate. At day t, x_t represents the input cryptocurrency price data of the LSTM cell whose output at the previous day is denoted as h_{t-1}, c_t stands for the memory cell value. The calculation process of the LSTM unit is conducted as follows:

$$i_t = (Wx_i x_t + Wh_i h_{t-1} + Wc_i c_{t-1} + b_i) \quad (2)$$

$$f_t = \left(W_{xfxt} + Wh_f h_{t-1} + W_c fc_{t-1} + b_f\right) \quad (3)$$

$$c_t = f_t c_{t-1} + i_t \, tanh \, (W_{xc} x_t + W_{hc} h_{t-1} + b_c) \quad (4)$$

$$o_t = (W_{xo} x_t + Who h_t - 1 + W_{co} c_{t-1} + b_o) \quad (5)$$

$$h_t = o_t \, tanh \, (c_t) \quad (6)$$

where i_t, f_t and o_t is the output of the input forget, and output gate, respectively and their corresponding weight matrices are W_i, W_f and W_o; c_t is the state of the memory cell; b is bias; and σ and tanh represent sigmoid and hyperbolic tangent activation functions, respectively. Benefitting from memory cells and control gates, LSTM builds a long-term delay between input and feedback. The internal state of memory cells retains a continuous flow of error, without the gradient exploding or disappearing. Similarly, GRU consists of the update and reset gates and a memory cell, whose outputs u_t, r_t and c_t, respectively, can be obtained as follows:

$$u_t = (W_{ux} x_t + W_{uc} c_{t-1} + b_u) \quad (7)$$

$$r_t = (W_{rc} x_t + W_{rc} c_{t-1} + b_r) \quad (8)$$

$$\widehat{c}_t = tanh(W_{cx} h_t + W_{rc}(r_t c_{t-1}) + b_c) \quad (9)$$

$$c_t = (1 - u_t) c_{t-1} + u_t \widehat{c}_t \quad (10)$$

where $\widehat{c}_t$ is a candidate state of the memory cell. In GRU, the reset gate is used to select the optimal time lag. Having fewer computational parameters than LSTM, GRU proved to be more effective when handling less frequent and smaller datasets, such as cryptocurrency price timeseries (Hansun et al., 2022). Inspired by this finding, we also consider a simple RNN in this study to exploit its computational efficiency. The simple RNN is defined as follows:

$$h_t = tanh(W_{hx} h_{t-1} + W_{hh} x_{t-1} + b_h) \quad (11)$$

$$y_t = tanh\left(W_{hy} h_t + b_y\right) \quad (12)$$

where W_h represents the weight matrix in the hidden layer, and h_t and y_t are the outputs of the hidden and output layer, respectively.

3.3. Hybrid two-stage models

The idea of a hybrid two-stage cryptocurrency price forecasting model has its roots in the seminal paper of Zhang (2003). The idea is to use a deep neural network to estimate the residuals of the ARIMA model, with ARIMA capturing the linear patterns of the cryptocurrency price data while a deep neural network LSTM is used to model the remaining nonlinear patterns, thus improving the accuracy of forecasts.

The ARIMA model is effective in detecting linear patterns in time-series data. The assumption of a linear data generation process is unrealistic for cryptocurrency time-series data, and it is a major limitation of the ARIMA model. At the same time, despite their rapid development and reasonably successful application to real-world forecasting problems, machine learning methods tend to be complex and may lack the transparency needed to receive widespread acceptance. Classical MLP neural network models are also not good enough to capture both linear and nonlinear patterns equally well. Therefore, here we combine ARIMA for estimating the linear component L_t and the above-mentioned LSTM recurrent deep neural network for estimating the residual e_t from the linear model, that is, the nonlinear component N_t.

Table 1
Values and ranges of model hyper-parameters.

Models	Hyper-parameters
MLP	Hidden layers: [1,2]; the number of hidden units: [10, 20, 30, 40, 50, 100]; activation function: ['relu', 'tanh']; optimizer solver: ['sgd', 'adam']; regularization alpha: [0.0001], learning rate for sgd optimizer: ['constant' , 'invscaling', 'adaptive'].
LSTM	Hidden layers: [1, 2]; the number of epochs: [3, 5, 10, 50, 100, 300]; the number of hidden units [4, 8, 16, 32, 64, 128, 256]; learning rate [0.001, 0.01, 0.1]; lag days [3, 5, 8, 10]
AdaBoost	The number of estimators: [10, 20, 30, 40, …, 100]; learning rate: [0.001, 0.01, 0.1, 1.0]; loss functions: [linear, square, exponential]
LightGBM	The number of estimators: [10, 20, 30, 40, …, 100]; learning rate: [0.001, 0.01, 0.1, 1.0].
RNN	The number of units: [4, 16, 32, 64, 128].
GRU	Hidden layers: [1, 2]; the number of epochs: [3, 5, 10, 50, 100, 300]; the number of hidden units [4, 8, 16, 32, 64, 128, 256]; learning rate [0.001, 0.01, 0.1].

3.4. Specifications of models

For both the boosting-based ensemble methods (AdaBoost and LightGBM regressor methods) and recurrent deep neural networks (simple RNN, GRU, and LSTM), grid search with a rolling window cross-validation (Bhattacharjee et al., 2022; Fuss and Koller, 2016) under minimizing the mean square error was used to find the optimal values of the hyper-parameters, as is shown in Table 1.[1]

For both boosting methods, two hyper-parameters were examined, the number of estimators and learning rate, for the LightGBM method, three types of loss functions were considered. For the simple RNN, different numbers of units were tested from. For the GRU and LSTM models, the following values of hyper-parameters were examined: hidden layers, number of epochs, number of hidden units, and learning rate. Overall, the least complex models were generated for Litecoin (with one hidden layer and 8 units) while most complex models were optimal for the Ethereum data (using 128 or 256 units in the hidden layers). Generally, 5 epochs were enough to train the deep learning-based forecasting models.

In the hybrid two-stage models, the parameters of ARIMA were selected semi-automatically by using the smallest value of BIC (Bayesian information criterion) on training data. For Ripple and Litecoin, the ARIMA models were ARIMA(1,0,1) and ARIMA(2,0,0), respectively. When the best model was represented by white noise, the restricted AR model was chosen based on the value of PACF (partial autocorrelation function). Hence, AR (6) and AR (2) were the best models for Bitcoin and Ethereum, respectively.

4. Empirical results

The daily cryptocurrency time-series data were collected from https:/www.investing.com.[2] To investigate the predictability of the

[1] The models were implemented in Python using the Scikit-Learn (ensemble methods), Statsmodels (ARIMA), and Keras (recurrent deep neural networks) libraries. The code is provided in the supplementary material.

[2] All data used are freely accessible and downloadable at https://www.investing.com/crypto/. The data used in this study are also available in the supplementary material.

Table 2
Data description.

Cryptocurrency	Starting date	Midpoint date (Covid-19 timepoint)	Ending date	No. of observations before midpoint date (before Covid-19)	No. of observations after midpoint date (after Covid-19)
Bitcoin	April 1, 2016	January 1, 2020	August 31, 2023	1370	1339
Ethereum	April 1, 2016	January 1, 2020	August 31, 2023	1370	1339
Litecoin	April 1, 2016	January 1, 2020	August 31, 2023	1370	1339
Ripple	April 1, 2016	January 1, 2020	August 31, 2023	1370	1339

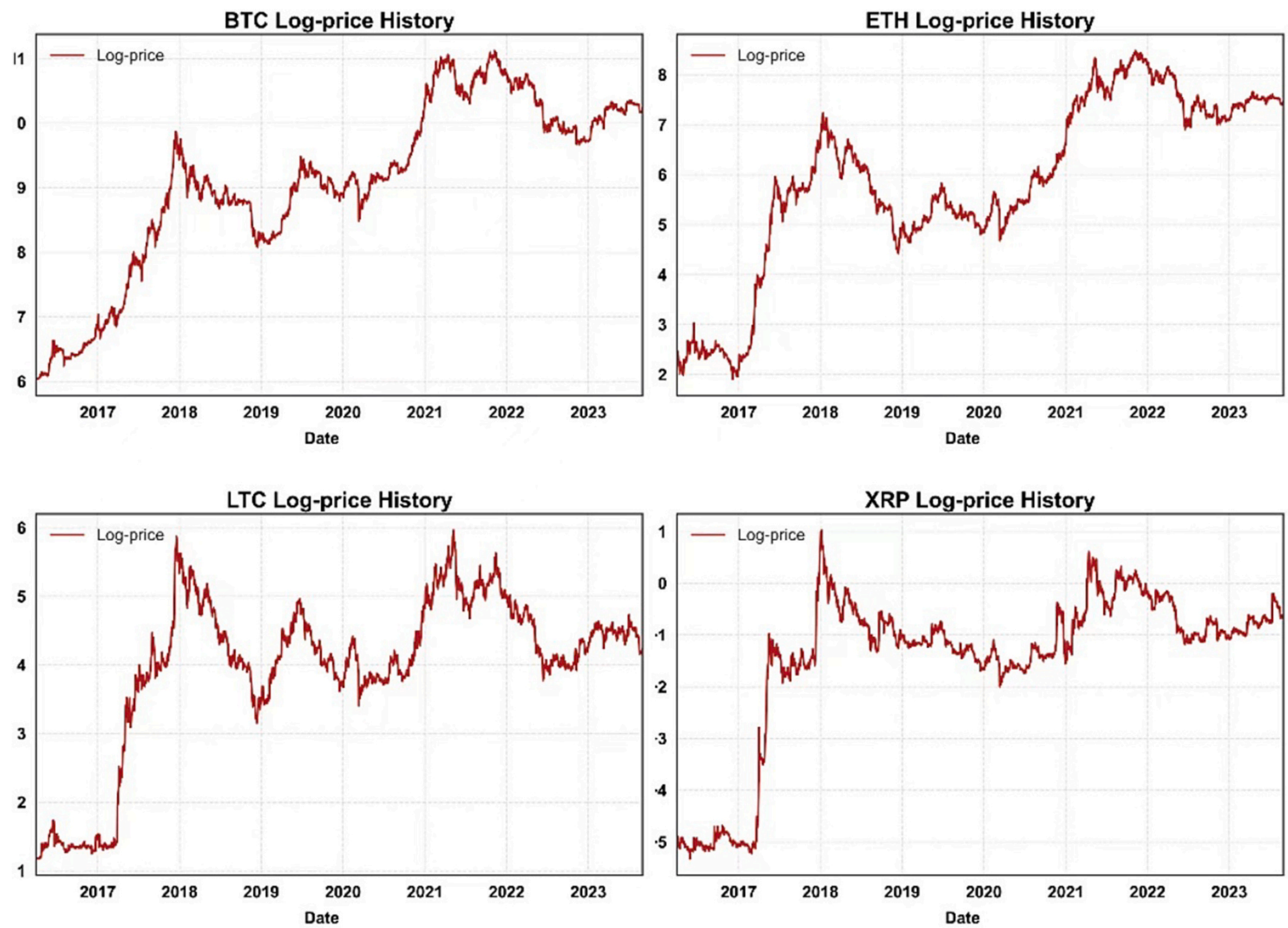

Fig. 1. Logarithmic cryptocurrency prices.

most relevant cryptocurrencies in terms of volume, four most popular cryptocurrencies were selected for our comparative analysis, namely BTC/USD, ETH/USD, LTC/USD, and XRP/USD. The trading data was obtained up to August 31, 2023. Table 2 presents the basic characteristics of the cryptocurrency data, and Fig. 1 shows the fluctuation of the logarithmic cryptocurrency prices. Note that each time series was split into two different sub-periods (denoted as before Covid-19 and after Covid-19), given the considerable effect of Covid-19 on cryptocurrency markets.

entire available period was considered as presented in Table 2.

In the experimental setting, the rolling window cross-validation approach (Bhattacharjee et al., 2022; Fuss and Koller, 2016) was used to split the time series into the training set immediately followed by the testing set. Specifically, the cryptocurrency time-series data were partitioned into a training set and testing set following a rolling window to match the structure of the time series data, the number of rolling test samples is shown in Table 2. Consistent with earlier studies (Corbet et al., 2022; Livieris et al., 2021), this paper applied the first-order differences of daily cryptocurrency logarithmic prices (Fig. 2). It should be noted that while differencing can be a useful approach to dealing with non-stationarity in time series data, it does not necessarily eliminate the need for complex machine learning models. Complex machine learning models are often able to capture more complex patterns, handle larger datasets, and automatically extract relevant features from the data. In addition, they can be effective in dealing with complex relationships and non-linear dynamics that may not be adequately captured by traditional quantitative techniques (Shajalal et al., 2023).

The summary descriptive statistics for the cryptocurrency time series

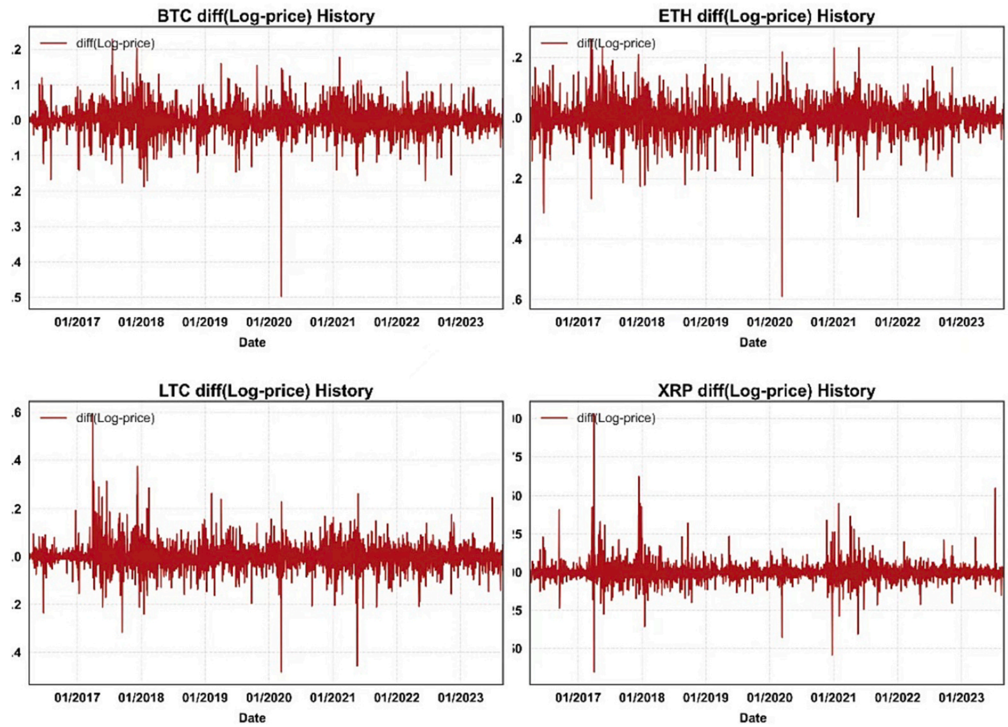

Fig. 2. First-order differences of daily logarithmic cryptocurrency prices.

Table 3
Summary descriptive statistics for cryptocurrency data.

Cryptocurrency	Mean	Minimum	Maximum	St.Dev.	Skewness	Kurtosis
Bitcoin	0.001526	−0.49728	0.227602	0.038855	−0.84402	13.01061
Ethereum	0.001829	−0.58964	0.258599	0.053355	−0.59340	9.05271
Litecoin	0.001102	−0.48208	0.591547	0.056149	0.404124	12.22713
Ripple	0.001559	−0.65299	1.027995	0.068411	2.056983	33.64264

are presented in Table 3. As can be seen in Table 3, the average daily price differences were positive for all the four cryptocurrencies, with Ethereum showing the highest returns, while Ripple showed the highest standard deviation. In addition, the high kurtosis of all cryptocurrency data indicates leptokurtic time series, with Ethereum showing the highest excess kurtosis. Finally, while the price differences of Bitcoin and Ethereum were negatively skewed (with a longer left tail), the opposite result can be observed for the price differences of Litecoin and Ripple.

To consider the stochastic nature of neural networks, hereinafter the results for the used neural networks are reported as an average of 50 simulation runs. As the datasets varied in terms of sample sizes (Table 2), the comparison of the forecasting performance of the used methods between different cryptocurrencies is difficult. To consider this limitation into account, three naive algorithms were employed to represent benchmarks. To this end, in agreement with previous studies (Akyildirim et al., 2021; Caporale et al., 2018; Oyedele et al., 2023), the following methods were used: random walk (RW), white noise (WN), and buy & sell (B&S). The random walk method is based on the Martingale assumption, hence using the last value of cryptocurrency price as the forecast for the next day value. The white noise method relies on a randomly generated cryptocurrency time series with normal distribution; that is, with the mean calculated as the mean of the training sample and the variance being equal to the variance of the training sample. The buy & sell method replicates a simple strategy of buying a cryptocurrency for a fixed amount of money every day and selling it at the end of the day.

To comprehensively evaluate the forecasting performance of the ensemble learning and deep learning methods, two sets of metrics were used. First, commonly used regression metrics were employed as follows: MAPE (mean absolute percentage error), ME (mean error), MAE (mean absolute error), MPE (mean percentage error), RMSE (root mean square error), R (correlation coefficient), and MIN-MAX error. Second, a set of metrics useful for evaluating investor performance was used, namely scalar product (SP), return score (Return), long return (Return_long), short return (Return_short), mean directional accuracy (MDA), mean directional accuracy positive (MDA+), and mean directional accuracy negative (MDA-). The SP of the actual and forecast values was used to simulate the buy & sell trading strategy, where the amount of investment is proportional to the forecast signal. The return score was used to simulate the trading strategy based on the signals of the used ensemble learning and deep learning methods. The return score was calculated as the sum of the returns of a particular trading strategy. The long (short) return simulated the return obtained using the long (short) trading strategy. MDA compares the predicted price direction (upward or downward) to the actual cryptocurrency price direction, while MDA+ and MDA- evaluate the upward and downward directional accuracy, respectively.

In addition to the three naive methods, several other baseline methods were used to demonstrate the efficiency of the deep learning methods, including the traditional ARIMA, MLP, and hybrid two-stage ARIMA+MLP methods. Five and ten previous cryptocurrency prices were examined in the experiments and the best results are presented hereinafter.

The results of the experiments are presented in Tables 4–7. From Tables 4–7, it can be noted that the compared methods performed

Table 4
Results of Bitcoin forecasting performance – regression and investor metrics.

Before Covid-19 (2016-04-01 to 2019-12-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	2.3206	0.0018	0.0326	1.2303	0.0427	−0.0028	0.0536
RW	4.2513	0.0027	0.0421	**0.0653**	0.0518	−0.0339	−0.0089
B&S	–	–	–	–	–	–	–
ARIMA	2.0730	0.0016	0.0314	1.4435	0.0398	−0.0392	0.0425
MLP	1.7589	0.0009	0.0233	1.0032	0.0307	0.4884	0.0341
LSTM	**1.0865**	0.0013	0.0272	1.0103	0.0361	−0.2352	**0.1010**
ARIMA+MLP	2.5503	0.0013	0.0286	1.3667	0.0364	0.2944	0.0489
ARIMA+LSTM	2.2950	0.0014	0.0301	1.4338	0.0380	0.1745	0.0556
AdaBoost	1.5506	0.0009	0.0221	0.9405	0.0294	0.5522	0.0639
LightGBM	1.5608	**0.0008**	**0.0205**	0.7688	**0.0280**	**0.6067**	0.0618
Simple RNN	1.6677	0.0012	0.0258	1.1060	0.0343	0.2629	0.0583
GRU	1.6831	0.0011	0.0250	0.9890	0.0328	0.3766	0.0427
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.0074	−0.2021	−0.1136	−0.0886	52.06	52.63	51.52
RW	0.3703	−0.4271	−0.2260	−0.2010	49.48	49.47	49.49
B&S	−0.0250	0.1703	**2.5727**	−2.5977	53.21	**100.00**	0.00
ARIMA	0.6979	0.0879	0.0314	0.0565	46.39	50.53	42.42
MLP	0.2905	2.4917	1.2333	1.2584	63.40	65.26	61.62
LSTM	−0.2798	−1.0300	−0.5275	−0.5025	43.81	36.84	50.51
ARIMA+MLP	0.3896	1.6641	0.8196	0.8446	58.76	62.11	55.56
ARIMA+LSTM	**1.9158**	0.4368	0.2059	0.2309	49.48	66.32	33.33
AdaBoost	−0.3511	2.4100	1.1925	1.2175	63.40	60.00	66.67
LightGBM	−0.3891	**3.1973**	1.5861	**1.6111**	**70.62**	67.37	**73.74**
Simple RNN	0.5778	1.2719	0.6234	0.6484	54.12	57.89	50.51
GRU	0.0796	1.8492	0.9121	0.9371	61.34	63.16	59.60

After Covid-19 (2020-01-01 to 2023-08-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	2.3029	0.0010	0.0251	1.2868	0.0323	0.0854	0.0429
RW	3.4430	0.0016	0.0324	0.8488	0.0406	0.0469	−0.0028
B&S	–	–	–	–	–	–	–
ARIMA	1.5634	0.0009	0.0233	1.0570	0.0301	−0.0489	0.0486
MLP	1.7886	0.0010	0.0248	0.5186	0.0322	−0.1178	0.0368
LSTM	**1.2102**	0.0009	0.0231	1.0695	0.0295	−0.2766	**0.0691**
ARIMA+MLP	1.5049	0.0007	0.0202	0.4260	0.0260	0.4178	0.0200
ARIMA+LSTM	1.5597	0.0008	0.0219	0.9105	0.0281	0.2511	0.0416
AdaBoost	1.2122	**0.0005**	0.0170	0.3978	0.0223	0.6160	0.0476
LightGBM	1.4518	**0.0005**	**0.0168**	**0.2916**	**0.0221**	**0.6234**	0.0486
Simple RNN	2.1846	0.0008	0.0223	0.4604	0.0286	0.2939	0.0180
GRU	2.1729	0.0013	0.0271	0.8790	0.0360	−0.0443	0.0072
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	0.2854	0.6972	0.3356	0.3615	57.67	65.63	49.46
RW	0.3509	0.0025	−0.0117	0.0142	47.62	48.96	46.24
B&S	−0.0259	0.1392	**2.0566**	−2.0825	51.12	**100.00**	0.00
ARIMA	0.1363	−0.2319	−0.1289	−0.1030	49.21	51.04	47.31
MLP	0.0875	−0.5718	−0.2988	−0.2729	49.21	51.04	47.31
LSTM	−1.3067	−1.2879	−0.6569	−0.6310	37.04	19.79	54.84
ARIMA+MLP	−0.7757	1.9676	0.9709	0.9967	68.25	57.29	79.57
ARIMA+LSTM	**1.5317**	0.6455	0.3098	0.3357	55.03	70.83	38.71
AdaBoost	0.1852	**2.6144**	1.2942	**1.3201**	**75.13**	82.29	67.74
LightGBM	−0.4340	2.3641	1.1691	1.1950	71.43	66.67	**76.34**
Simple RNN	0.5629	1.2358	0.6050	0.6308	61.38	67.71	54.84
GRU	−0.6326	0.2702	0.1222	0.1481	55.03	51.04	59.14

differently not only on different cryptocurrencies, but also in terms of regression and investor statistics. As for the regression metrics, LightGBM and AdaBoost performed best for Bitcoin and Ethereum across the sub-periods studied. Different results were obtained for the two remaining cryptocurrencies with lower market capitalization. For Ripple, LightGBM surpassed the remaining methods during the pre-Covid-19 sub-period, while MLP and GRU outperformed the other methods in the period following the emergence of Covid-19. Similarly, different results are apparent for Litecoin, with the MLP and Simple RNN demonstrating superior performance prior to and following the appearance of Covid-19, respectively. What is striking here is that, except Bitcoin and Ethereum, these methods did not perform similarly well in terms of investor statistics, suggesting that although achieving low forecast deviations, these methods failed to capture the direction of the next day's price change. Regarding the investor metrics, Simple RNN performed best for Ripple and LightGBM showed superior performance for Litecoin. For the pre-Covid-19 sub-period, the returns of the best performing methods ranged from 2.68 for Litecoin using AdaBoost to 4.93 for Ripple (Simple RNN). For the period following the emergence of Covid-19, the returns ranged from 2.61 (for Bitcoin using AdaBoost) to 3.39 (for Litecoin using LightGBM). Exceptional MDA was obtained for all cryptocurrencies, ranging from 67.5% for Ethereum to 75.1% for Bitcoin. Generally, there was a greater MDA across cryptocurrencies during the post-Covid-19 pandemic, resulting in improved predictability of cryptocurrency price trends. This is a rather remarkable result when considering balanced performance achieved in all cases in terms of upward and downward trend prediction. To compare the investor performance statistically, we conducted a nonparametric Friedman test across the investor metrics. The test uses the Friedman statistics to rank the forecasting models across the two sub-periods. The Friedman *p*-value

Table 5
Results of Ethereum forecasting performance – regression and investor metrics.

Before Covid-19 (2016-04-01 to 2019-12-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	9.1019	0.0026	0.0408	7.7331	0.0507	0.0279	−0.0096
RW	6.4210	0.0032	0.0446	2.9206	0.0569	0.0019	−0.0186
B&S	–	–	–	–	–	–	–
ARIMA	3.3405	0.0018	0.0331	1.1155	0.0429	0.0584	0.0245
MLP	2.0626	0.0014	0.0276	**0.5063**	0.0371	0.4027	0.0192
LSTM	**1.2799**	0.0017	0.0301	1.0748	0.0406	−0.0113	0.0818
ARIMA+MLP	3.5920	0.0015	0.0289	2.1820	0.0385	0.3814	0.0198
ARIMA+LSTM	3.0440	0.0017	0.0310	1.3090	0.0409	0.2626	0.0315
AdaBoost	3.2182	0.0012	0.0257	2.5112	0.0348	0.5126	0.0281
LightGBM	1.7602	**0.0011**	**0.0237**	1.3562	**0.0334**	**0.5690**	**0.0441**
Simple RNN	2.4482	0.0012	0.0260	1.4288	0.0347	0.5268	0.0016
GRU	2.1173	0.0012	0.0261	1.3738	0.0348	0.5225	−0.0126
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.1181	0.3712	0.1729	0.1983	48.45	46.32	50.51
RW	−0.6564	0.2324	0.1035	0.1289	48.97	44.21	53.54
B&S	−0.0254	0.7985	**2.8866**	−2.9119	52.11	**100.00**	0.00
ARIMA	0.7521	−0.0252	−0.0253	0.0001	48.45	49.47	47.47
MLP	0.2905	2.4304	1.2025	1.2279	62.89	64.21	61.62
LSTM	−0.8632	−0.5598	−0.2926	−0.2672	47.94	48.42	47.47
ARIMA+MLP	0.7856	1.7765	0.8756	0.9009	57.22	58.95	55.56
ARIMA+LSTM	**2.3276**	0.9257	0.4502	0.4755	53.09	70.53	36.36
AdaBoost	0.6268	3.2507	1.6127	1.6380	64.43	66.32	62.63
LightGBM	0.6585	**3.5207**	1.7477	**1.7730**	**67.53**	70.53	**64.65**
Simple RNN	0.4325	3.5054	1.7400	1.7654	65.98	68.42	63.64
GRU	0.4298	3.3323	1.6535	1.6788	63.40	65.26	61.62
After Covid-19 (2020-01-01 to 2023-08-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	2.4354	0.0022	0.0350	0.6965	0.0465	0.0243	0.0841
RW	3.7473	0.0027	0.0410	−0.1798	0.0522	0.1100	0.0203
B&S	–	–	–	–	–	–	–
ARIMA	1.1594	0.0016	0.0287	0.7467	0.0397	0.0353	**0.1312**
MLP	1.6599	0.0011	0.0240	1.4470	0.0324	0.5900	0.0743
LSTM	1.1763	0.0016	0.0290	0.9329	0.0400	−0.0930	0.1270
ARIMA+MLP	1.4629	0.0011	0.0250	1.0449	0.0336	0.5132	0.1086
ARIMA+LSTM	**1.1498**	0.0013	0.0266	0.5360	0.0366	0.3767	0.1312
AdaBoost	1.3491	**0.0009**	0.0230	1.1058	**0.0294**	**0.6661**	0.0497
LightGBM	1.4344	**0.0009**	**0.0219**	1.0440	0.0300	0.6437	0.1044
Simple RNN	2.0241	0.0019	0.0312	**0.9193**	0.0438	0.0566	0.0791
GRU	1.4885	0.0014	0.0277	1.1412	0.0380	0.2405	0.1175
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.0661	−0.4673	−0.2550	−0.2124	43.39	48.51	37.50
RW	−0.5820	1.1031	0.5303	0.5729	56.08	50.50	62.50
B&S	−0.0426	0.4142	**2.6858**	−2.7284	52.84	**100.00**	0.00
ARIMA	0.5349	−0.1435	−0.0930	−0.0504	53.97	55.45	52.27
MLP	−1.3096	3.0747	1.5161	1.5586	67.72	52.48	85.23
LSTM	−1.1820	−0.1093	−0.0759	−0.0333	51.85	44.55	60.23
ARIMA+MLP	−1.1435	2.5770	1.2672	1.3098	64.55	54.46	76.14
ARIMA+LSTM	**1.9223**	0.8068	0.3821	0.4247	60.32	72.28	46.59
AdaBoost	−0.6104	3.1822	1.5698	1.6124	66.67	62.38	71.59
LightGBM	−0.1119	**3.3438**	1.6506	**1.6932**	**71.43**	68.32	**75.00**
Simple RNN	0.0379	1.0002	0.4788	0.5214	56.08	55.45	56.82
GRU	−0.5029	1.1768	0.5671	0.6097	55.03	51.49	59.09

<0.01 (the Friedman statistics ranged from 35.3 to 56.2) indicates significant differences between the compared forecasting models for all the four cryptocurrencies. Among the forecasting models, the Simple RNN ranked first for Ripple, while LightGBM ranked first for Ethereum, Bitcoin and Litecoin.

5. Discussion

Overall, three different patterns were observed in our forecasting results, with Bitcoin/Ethereum and Ripple/Litecoin representing these patterns. This finding is not surprising given the descriptive statistics of their time series. While the expected finding was that ensemble learning and deep learning methods outperform the conventional statistical methods and shallow neural networks, this study showed that, at least in terms of point forecasts, less complex conventional models can be more effective for some cryptocurrency time series. This finding is in contrast to recent review studies (Khedr et al., 2021; Ren et al., 2022), which highlighted the dominance of LSTM models. This may be because conventional models are just as effective as deep learning models, particularly when the data are univariate and there is no need to deal with additional variables or complex relationships. Traditional computational models are based on statistical principles and assumptions that are appropriate for univariate data, as these models take into account factors such as autocorrelation, seasonality and trend. Therefore, in the context of univariate time series analysis, where there are no additional variables or complex relationships to consider, traditional quantitative techniques can often be sufficient (Castán-Lascorz et al., 2022).

We have shown that LSTM models can be overcome by GRU models, even when LSTM is combined with ARIMA. One reasonable explanation for this decrease is that the ensemble learning and deep learning

Table 6
Results of Ripple forecasting performance – regression and investor metrics.

Before Covid-19 (2016-04-01 to 2019-12-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	7.2913	0.0019	0.0342	−2.3405	0.0441	0.0001	0.0637
RW	9.2486	0.0028	0.0421	−4.3758	0.0527	−0.0828	0.0299
B&S	–	–	–	–	–	–	–
ARIMA	**0.9997**	0.0013	0.0255	0.9997	0.0358	0.0892	0.1486
MLP	1.5667	0.0014	0.0267	1.2493	0.0377	−0.3116	0.1290
LSTM	1.8834	0.0013	0.0261	0.3450	0.0366	−0.2481	0.1388
ARIMA+MLP	3.6238	0.0014	0.0278	2.5354	0.0368	0.1993	0.0868
ARIMA+LSTM	2.6338	0.0014	0.0293	0.8290	0.0380	0.1868	0.0908
AdaBoost	3.5384	**0.0009**	0.0220	2.7337	0.0305	0.5321	0.0829
LightGBM	2.8600	0.0008	**0.0197**	2.3387	**0.0285**	**0.6182**	0.1078
Simple RNN	3.6358	0.0027	0.0348	**0.7220**	0.0523	0.5112	0.1761
GRU	3.8229	0.0039	0.0447	1.7255	0.0627	0.1772	**0.2071**
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	0.0156	0.0054	0.0048	0.0006	45.36	42.39	48.04
RW	−0.4854	−0.2821	−0.1390	−0.1431	47.42	45.65	49.02
B&S	0.0041	0.9591	2.4816	−2.4775	52.01	**100.00**	0.00
ARIMA	−4.7639	−0.0265	−0.0112	−0.0153	52.58	1.09	**99.02**
MLP	−2.5181	−1.5176	−0.7567	−0.7608	44.33	11.96	73.53
LSTM	−1.1531	−1.5355	−0.7657	−0.7698	40.72	35.87	45.10
ARIMA+MLP	0.4418	0.7275	0.3658	0.3617	51.55	60.87	43.14
ARIMA+LSTM	**2.0134**	1.0897	0.5469	0.5428	52.06	72.83	33.33
AdaBoost	0.7292	3.2439	1.6240	1.6199	69.07	78.26	60.78
LightGBM	0.5372	3.4476	1.7258	1.7217	71.13	75.00	67.65
Simple RNN	−0.2844	**4.9318**	**2.4881**	**2.4437**	**71.65**	65.69	78.26
GRU	−0.2782	1.3665	0.7055	0.6610	54.12	51.96	56.52
After Covid-19 (2020-01-01 to 2023-08-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	5.5345	0.0040	0.0460	−0.4069	0.0630	−0.0075	0.0569
RW	7.0524	0.0054	0.0542	1.3672	0.0734	−0.0037	0.0136
B&S	–	–	–	–	–	–	–
ARIMA	**0.9998**	0.0027	0.0301	0.9998	0.0517	−0.1947	0.1724
MLP	1.9275	**0.0019**	0.0262	1.0353	0.0440	**0.5252**	0.0052
LSTM	1.5035	0.0030	0.0321	1.1093	0.0545	−0.4740	**0.1443**
ARIMA+MLP	2.4915	0.0032	0.0333	0.7799	0.0563	0.1655	0.0586
ARIMA+LSTM	2.5516	0.0035	0.0364	1.3587	0.0594	−0.0738	0.1335
AdaBoost	1.1980	0.0022	0.0262	0.8145	0.0464	0.4479	0.1284
LightGBM	1.4219	0.0021	**0.0252**	0.7566	0.0459	0.4602	0.1004
Simple RNN	1.6898	0.0021	0.0267	**0.6932**	0.0457	0.4750	−0.0053
GRU	1.4707	0.0020	0.0256	0.7454	**0.0442**	0.5184	0.0356
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.9814	0.7209	0.3550	0.3659	50.79	39.08	60.78
RW	−0.2337	−0.1856	−0.0982	−0.0874	55.03	51.72	57.84
B&S	−0.0108	0.6931	**2.8411**	−2.8520	53.52	**100.00**	0.00
ARIMA	−5.3480	−0.3342	−0.1725	−0.1617	53.44	0.00	**99.02**
MLP	0.9483	2.5124	1.2508	1.2616	66.67	81.61	53.92
LSTM	1.9768	−1.8340	−0.9224	−0.9116	41.27	78.16	9.80
ARIMA+MLP	0.2686	1.6379	0.8135	0.8244	59.26	58.62	59.80
ARIMA+LSTM	**2.5810**	−1.2910	−0.6509	−0.6401	45.50	67.82	26.47
AdaBoost	1.0692	2.3798	1.1845	1.1953	66.67	82.76	52.94
LightGBM	0.6917	**2.9912**	1.4902	**1.5010**	**72.49**	77.01	68.63
Simple RNN	0.3032	2.7200	1.3546	1.3654	69.31	71.26	67.65
GRU	0.5134	2.7135	1.3513	1.3621	69.84	74.71	65.69

methods were overfitted for the less complex time series. The poor performance of the hybrid models might be due to the fact that there is a lot of noise in the residuals, so the hybrid models get overfitted by the noise.

Different results were observed for the investor metrics, suggesting that more complex machine learning methods are needed to adequately perform in terms of forecasting cryptocurrency market direction. We have shown that remarkable returns can be achieved by following the trading strategy based on the forecasts produced by the LightGBM models. This remarkable performance can be attributed to effectively managing large datasets while exploiting its regularization mechanism that helps prevent overfitting, making it more robust for cryptocurrency price forecasting (Sun et al., 2020). The highest returns could be obtained for Ripple in the pre-Covid-19 period and for Litecoin in the post-Covid-19 period, indicating that these cryptocurrencies are the most inefficient cryptocurrency markets, whereas Bitcoin appears to be the least inefficient market. This complexity effect might also be related to greater liquidity in the Bitcoin market (Al-Yahyaee et al., 2020).

Our results indicate that the trading strategies based on deep learning (for Ripple) or ensemble learning (for Bitcoin, Ethereum, and Litecoin) could allow cryptocurrency investors to effectively predict market development, particularly in less complex cryptocurrency markets. The study's findings provide cryptocurrency investors with valuable insights into effective trading strategies, adjusting their investment strategy to either take a long position or a short position. The demonstrated financial effectiveness of deep and ensemble learning techniques in cryptocurrency trading also offers new tools for financial analysts, enhancing their ability to predict market movements. Nonetheless, policymakers ought to implement financial market interventions with a view to enhancing the level of transparency and efficiency in these

Table 7
Results of Litecoin forecasting performance – regression and investor metrics.

Before Covid-19 (2016-04-01 to 2019-12-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	12.7863	0.0035	0.0432	4.6522	0.0592	0.0114	0.1668
RW	13.6119	0.0042	0.0494	−2.8895	0.0651	0.0034	0.1641
B&S	–	–	–	–	–	–	–
ARIMA	**1.0051**	0.0020	0.0235	0.9942	0.0444	−0.0340	**0.2863**
MLP	2.2688	0.0012	0.0196	**−0.0703**	**0.0351**	**0.6398**	0.2427
LSTM	1.2487	0.0021	0.0243	0.8344	0.0458	−0.2085	0.2534
ARIMA+MLP	4.9575	0.0018	0.0266	1.8160	0.0428	0.3245	0.2269
ARIMA+LSTM	3.6209	0.0018	0.0257	1.3296	0.0424	0.3529	0.2300
AdaBoost	3.1387	0.0015	0.0216	0.2177	0.0386	0.5106	0.2526
LightGBM	2.4554	**0.0014**	**0.0196**	−0.1327	0.0371	0.5663	0.2406
Simple RNN	2.5022	**0.0014**	0.0222	−0.2846	0.0379	0.5342	0.2245
GRU	2.8096	**0.0014**	0.0217	−0.3606	0.0369	0.5598	0.2114
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.6098	0.1510	0.0696	0.0814	49.48	46.46	52.63
RW	0.0433	0.3973	0.1927	0.2045	53.09	51.52	54.74
B&S	−0.0118	0.5278	**2.2580**	−2.2698	50.29	**100.00**	0.00
ARIMA	3.6668	0.0241	0.0062	0.0179	51.03	95.96	4.21
MLP	0.1161	2.5623	1.2752	1.2870	67.01	66.67	67.37
LSTM	−1.7866	−1.2000	−0.6059	−0.5941	47.42	29.29	66.32
ARIMA+MLP	−0.0941	2.5942	1.2912	1.3030	60.31	52.53	68.42
ARIMA+LSTM	**2.7781**	0.4407	0.2145	0.2263	56.70	70.71	42.11
AdaBoost	−0.2214	**2.6827**	1.3354	**1.3472**	66.49	54.55	**78.95**
LightGBM	−0.2178	2.6739	1.3311	1.3428	**68.04**	63.64	72.63
Simple RNN	−0.3447	2.2188	1.1035	1.1153	64.95	62.63	67.37
GRU	−0.1891	2.4517	1.2200	1.2317	65.98	65.66	66.32
After Covid-19 (2020-01-01 to 2023-08-31)							
Method	MAPE	ME	MAE	MPE	RMSE	R	MIN-MAX
WN	5.3427	0.0031	0.0397	3.2827	0.0557	−0.1341	0.1282
RW	14.4445	0.0036	0.0459	−8.4577	0.0604	0.0856	0.1208
B&S	–	–	–	–	–	–	–
ARIMA	**0.9989**	0.0020	0.0290	0.9989	0.0444	0.0633	0.2086
MLP	2.6556	0.0016	0.0277	**0.4133**	0.0403	0.4409	0.1507
LSTM	1.5458	0.0021	0.0298	1.3016	0.0453	−0.0191	**0.1783**
ARIMA+MLP	1.2830	0.0012	0.0274	0.9887	0.0347	−0.5306	0.0890
ARIMA+LSTM	1.1576	0.0011	0.0263	0.8930	0.0333	0.5702	0.0972
AdaBoost	2.9554	0.0010	0.0229	1.1805	0.0319	**0.6975**	0.0000
LightGBM	3.4677	0.0012	0.0224	1.4952	0.0353	0.6198	0.1566
Simple RNN	2.2888	**0.0008**	**0.0216**	1.0384	**0.0287**	0.5444	0.0120
GRU	2.2607	0.0018	0.0278	0.4635	0.0425	0.3090	0.1680
Method	**SP**	**Return**	**Return_long**	**Return_short**	**MDA [%]**	**MDA+ [%]**	**MDA- [%]**
WN	−0.5277	−0.3957	−0.1953	−0.2005	47.09	43.62	50.53
RW	−0.0799	0.9031	0.4542	0.4490	52.91	57.45	48.42
B&S	0.0052	0.5019	**2.7536**	−2.7484	51.36	**100.00**	0.00
ARIMA	**5.2291**	0.2005	0.1028	0.0976	50.26	97.87	3.16
MLP	−0.5707	2.0535	1.0293	1.0241	58.73	54.26	63.16
LSTM	0.4309	−0.8283	0.4116	0.4168	48.15	50.00	46.32
ARIMA+MLP	1.1153	2.6573	1.3402	1.3171	48.04	52.50	44.09
ARIMA+LSTM	2.0505	0.9450	−0.4840	−0.4610	41.27	15.63	67.74
AdaBoost	0.7442	2.9949	1.5000	1.4948	66.14	78.72	53.68
LightGBM	0.3632	**3.3869**	1.6960	**1.6908**	**74.60**	76.60	**72.63**
Simple RNN	0.3974	2.6986	1.3552	1.3435	68.25	74.19	62.50
GRU	−0.7234	1.5542	0.7797	0.7745	63.49	57.45	69.47

markets. According to our findings on the predictability of cryptocurrency price trends, this seems particularly important in the post-pandemic period.

Given the study design, caution must be exercised when interpreting the results, as the second sub-period period included the Covid-19 pandemic period, which positively affected cryptocurrency market efficiency (Mnif et al., 2020), as well as its role as a store of value (Corbet et al., 2020). Although we used the most recent data available, more research is needed to validate our findings for the post-pandemic data. This study also failed to account for the portfolio returns, suggesting that future studies should focus on the construction of trading strategies for portfolio investors.

6. Conclusion

In this study, we have provided a comparative study of univariate ensemble learning and deep learning models for forecasting cryptocurrency prices. We have conducted extensive experiments using historical time-series data from four major cryptocurrencies. For the regression results, the results show the higher effectiveness of complex machine learning methods for all four cryptocurrency time series. More importantly, we investigated the efficacy of trading strategies based on forecasting models, showing that LightGBM may provide highly profitable trading strategies for investors in the Bitcoin, Ethereum and Litecoin markets. For the Ripple market, Simple RNN is recommended as the best forecasting model for investors. Strikingly, these findings appear to be robust to the sub-periods studied. Taken together, these findings suggest that ensemble learning and deep learning models can effectively guide investors in their trading decisions. In the future, we will examine more frequent real-time data to better exploit the advantages of deep learning models. In a similar manner, we seek to utilize multivariate data, including the determinants of cryptocurrency supply and demand, in order to compare the performance of univariate and multivariate models. In this way, we would also be able to interpret the models more effectively in terms of the contribution of the determinants.

Data availability

Data will be made available on request.

Acknowledgements

This research paper was made possible by a grant from the Czech Sciences Foundation (No. 22-22586S).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.irfa.2023.103055.

References

Aggarwal, D., Chandrasekaran, S., & Annamalai, B. (2020). A complete empirical ensemble mode decomposition and support vector machine-based approach to predict bitcoin prices. *Journal of Behavioral and Experimental Finance, 27*, Article 100335.

Akyildirim, E., Goncu, A., & Sensoy, A. (2021). Prediction of cryptocurrency returns using machine learning. *Annals of Operations Research, 297*(1), 3–36.

Altan, A., Karasu, S., & Bekiros, S. (2019). Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos, Solitons & Fractals, 126*, 325–336.

Al-Yahyaee, K. H., Mensi, W., Ko, H. U., Yoon, S. M., & Kang, S. H. (2020). Why cryptocurrency markets are inefficient: The impact of liquidity and volatility. *The North American Journal of Economics and Finance, 52*, Article 101168.

Baur, D. G., Dimpfl, T., & Kuck, K. (2018). Bitcoin, gold and the US dollar–a replication and extension. *Finance Research Letters, 25*, 103–110.

Bhattacharjee, B., Kumar, R., & Senthilkumar, A. (2022). Unidirectional and bidirectional LSTM models for edge weight predictions in dynamic cross-market equity networks. *International Review of Financial Analysis, 84*, Article 102384.

Bouri, E., Gupta, R., Tiwari, A. K., & Roubaud, D. (2017). Does bitcoin hedge global uncertainty? Evidence from wavelet-based quantile-in-quantile regressions. *Finance Research Letters, 23*, 87–95.

Caporale, G. M., Gil-Alana, L., & Plastun, A. (2018). Persistence in the cryptocurrency market. *Research in International Business and Finance, 46*, 141–148.

Castán-Lascorz, M. A., Jiménez-Herrera, P., Troncoso, A., & Asencio-Cortés, G. (2022). A new hybrid method for predicting univariate and multivariate time series based on pattern forecasting. *Information Sciences, 586*, 611–627.

Catania, L., Grassi, S., & Ravazzolo, F. (2019). Forecasting cryptocurrencies under model and parameter instability. *International Journal of Forecasting, 35*(2), 485–501.

Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for bitcoin exchange rate prediction using economic and technology determinants. *International Journal of Forecasting, 37*(1), 28–43.

Cheng, J. (2023). Modelling and forecasting risk dependence and portfolio VaR for cryptocurrencies. *Empirical Economics, 65*, 899–924.

Chu, J., Zhang, Y., & Chan, S. (2019). The adaptive market hypothesis in the high frequency cryptocurrency market. *International Review of Financial Analysis, 64*, Article 221231.

Chu, J., Chan, S., & Zhang, Y. (2020). High frequency momentum trading with cryptocurrencies. *Research in International Business and Finance, 52*, Article 101176.

Conrad, C., Custovic, A., & Ghysels, E. (2018). Long-and short-term cryptocurrency volatility components: A GARCH-MIDAS analysis. *Journal of Risk and Financial Management, 11*(2), 23.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis, 62*, 182–199.

Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., & Oxley, L. (2020). Any port in a storm: Cryptocurrency safe-havens during the COVID-19 pandemic. *Economics Letters, 194*, Article 109377.

Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., Lucey, B., & Oxley, L. (2022). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters, 45*, Article 102137.

Cui, T., Ding, S., Jin, H., & Zhang, Y. (2022). Portfolio constructions in cryptocurrency market: A CVaR-based deep reinforcement learning approach. *Economic Modelling, 106078*.

Das, D., Le Roux, C. L., Jana, R. K., & Dutta, A. (2020). Does bitcoin hedge crude oil implied volatility and structural shocks? A comparison with gold, commodity and the US Dollar. *Finance Research Letters, 36*, Article 101335.

Derbentsev, V., Matviychuk, A., & Soloviev, V. N. (2020). Forecasting of cryptocurrency prices using machine learning. In *Advanced Studies of Financial Technologies and Cryptocurrency Markets* (pp. 211–231).

Dong, F., Xu, Z., & Zhang, Y. (2022). Bubbly bitcoin. *Economic Theory, 74*(3), 9731015.

Efat, M. I. A., Petr, H., Abedin, M. Z., Azad, R. U., Jaber, M. A., Aditya, S., & Hassan, M. K. (2022). Deep learning model using hybrid adaptive trend estimated series for modelling and forecasting sales. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04838-6

El Montasser, G., Charfeddine, L., & Benhamed, A. (2022). COVID-19, cryptocurrencies bubbles and digital market efficiency: Sensitivity and similarity analysis. *Finance Research Letters, 46*, Article 102362.

Fakhfekh, M., & Jeribi, A. (2020). Volatility dynamics of crypto-currencies' returns: Evidence from asymmetric and long memory GARCH models. *Research in International Business and Finance, 51*, Article 101075.

Fang, T., Su, Z., & Yin, L. (2020). Economic fundamentals or investor perceptions? The role of uncertainty in predicting long-term cryptocurrency volatility. *International Review of Financial Analysis, 71*, Article 101566.

Fu, E., Zhang, Y., Yang, F., & Wang, S. (2022). Temporal self-attention-based ConvLSTM network for multivariate time series prediction. *Neurocomputing, 501*, 162–173.

Fuss, R., & Koller, J. (2016). The role of spatial and temporal structure for residential rent predictions. *International Journal of Forecasting, 32*(4), 1352–1368.

Guesmi, K., Saadi, S., Abid, I., & Ftiti, Z. (2019). Portfolio diversification with virtual currency: Evidence from bitcoin. *International Review of Financial Analysis, 63*, Article 431437.

Hajek, P., Hikkerova, L., & Sahut, J. M. (2023). How well do investor sentiment and ensemble learning predict bitcoin prices? *Research in International Business and Finance, 64*, Article 101836.

Hansun, S., Wicaksana, A., & Khaliq, A. Q. (2022). Multivariate cryptocurrency prediction: Comparative analysis of three recurrent neural networks approaches. *Journal of Big Data, 9*(1), 1–15.

Ibrahim, A., Kashef, R., & Corrigan, L. (2021). Predicting market movement direction for bitcoin: A comparison of time series modeling methods. *Computers and Electrical Engineering, 89*, Article 106905.

Jalan, A., Matkovskyy, R., & Yarovaya, L. (2021). "Shiny" crypto assets: A systemic look at gold-backed cryptocurrencies during the COVID-19 pandemic. *International Review of Financial Analysis, 78*, Article 101958.

Ji, Q., Bouri, E., Gupta, R., & Roubaud, D. (2018). Network causality structures among bitcoin and other financial assets: A directed acyclic graph approach. *The Quarterly Review of Economics and Finance, 70*, 203–213.

Jiang, Y., Wu, L., Tian, G., & Nie, H. (2021). Do cryptocurrencies hedge against EPU and the equity market volatility during COVID-19?–new evidence from quantile coherency analysis. *Journal of International Financial Markets Institutions and Money, 72*, Article 101324.

Khedr, A. M., Arif, I., El Bannany, M., Alhashmi, S. M., & Sreedharan, M. (2021). Cryptocurrency price prediction using traditional statistical and machine learning techniques: A survey. *Intelligent Systems in Accounting, Finance and Management, 28*(1), 3–34.

Klein, T., Thu, H. P., & Walther, T. (2018). Bitcoin is not the new gold–a comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis, 59*, 105–116.

Kristjanpoller, W., & Minutolo, M. C. (2018). A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis. *Expert Systems with Applications, 109*, 1–11.

Kyriazis, N., Papadamou, S., & Corbet, S. (2020). A systematic review of the bubble dynamics of cryptocurrency prices. *Research in International Business and Finance, 54*, Article 101254.

Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals, 118*, 35–40.

Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on cnn lstm hybrid neural network model. *The Journal of Engineering, 2020*(13), 344–347.

Liu, M., Li, G., Li, J., Zhu, X., & Yao, Y. (2021). Forecasting the price of bitcoin using deep learning. *Finance Research Letters, 40*, Article 101755.

Livieris, I. E., Stavroyiannis, S., Iliadis, L., & Pintelas, P. (2021). Smoothing and stationarity enforcement framework for deep learning time-series forecasting. *Neural Computing and Applications, 33*(20), 14021–14035.

Manchanda, H., & Aggarwal, S. (2021). Forecasting cryptocurrency time series using Adaboost-based ensemble learning techniques. In *Innovations in cyber physical systems* (pp. 207–219). Singapore: Springer.

Mnif, E., Jarboui, A., & Mouakhar, K. (2020). How the cryptocurrency market has performed during COVID 19? A multifractal analysis. *Finance Research Letters, 36*, Article 101647.

Moula, F. E., Guotai, C., & Abedin, M. Z. (2017). Credit default prediction modeling: An application of support vector machine. *Risk Management, 19*(2), 158–187.

Murray, K., Rossi, A., Carraro, D., & Visentin, A. (2023). On forecasting cryptocurrency prices: A comparison of machine learning, deep learning, and ensembles. *Forecasting, 5*(1), 196–209.

Nakano, M., Takahashi, A., & Takahashi, S. (2018). Bitcoin technical trading with artificial neural network. *Physica A: Statistical Mechanics and its Applications, 510*, Article 587609.

Ortu, M., Uras, N., Conversano, C., Bartolucci, S., & Destefanis, G. (2022). On technical trading and social media indicators for cryptocurrency price classification through deep learning. *Expert Systems with Applications, 198*, Article 116804.

Oyedele, A. A., Ajayi, A. O., Oyedele, L. O., Bello, S. A., & Jimoh, K. O. (2023). Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction. *Expert Systems with Applications, 213*, Article 119233.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications, 42*(4), 2162–2172.

Rahmani Cherati, M., Haeri, A., & Ghannadpour, S. F. (2021). Cryptocurrency direction forecasting using deep learning algorithms. *Journal of Statistical Computation and Simulation, 91*(12), 2475–2489.

Ren, Y. S., Ma, C. Q., Kong, X. L., Baltas, K., & Zureigat, Q. (2022). Past, present, and future of the application of machine learning in cryptocurrency research. *Research in International Business and Finance, 63*, Article 101799.

Shahzad, S. J. H., Balli, F., Naeem, M. A., Hasan, M., & Arif, M. (2022). Do conventional currencies hedge cryptocurrencies? *The Quarterly Review of Economics and Finance, 85*, 223–228.

Shajalal, M., Petr, H., & Abedin, M. Z. (2023). Product backorder prediction with deep neural network on imbalance data. *International Journal of Production Research, 61*(1), 302–319.

Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters, 32*, Article 101084.

Tuwiner, J. (2022). 63+ cryptocurrency statistics, facts & trends. Available at https://buybitcoinworldwide.com/cryptocurrency-statistics/.

Uras, N., Marchesi, L., Marchesi, M., & Tonelli, R. (2020). Forecasting bitcoin closing price series using linear regression and neural networks models. *PeerJ Computer Science, 6*, Article e279.

Walther, T., Klein, T., & Bouri, E. (2019). Exogenous drivers of bitcoin and cryptocurrency volatility–a mixed data sampling approach to forecasting. *Journal of International Financial Markets Institutions and Money, 63*, Article 101133.

Yang, B., Sun, Y., & Wang, S. (2020). A novel two-stage approach for cryptocurrency analysis. *International Review of Financial Analysis, 72*, Article 101567.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation, 31*(7), 1235–1270.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0

Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Some stylized facts of the cryptocurrency market. *Applied Economics, 50*(55), 5950–5965.

Zhang, Z., Dai, H. N., Zhou, J., Mondal, S. K., García, M. M., & Wang, H. (2021). Forecasting cryptocurrency price using convolutional neural networks with weighted and attentive memory channels. *Expert Systems with Applications, 183*, Article 115378.