Bitcoin price prediction based on other cryptocurrencies using machine learning and time series analysis

N. Maleki[a], A. Nikoubin[a], M. Rabbani[a,*], and Y. Zeinali[b]

a. School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, P.O. Box 11155-4563, Iran.
b. School of Industrial Engineering, Sharif University of Technology, Tehran, Iran.

Abstract. Cryptocurrencies have attracted significant awareness up to now, and Bitcoin is the most remarkable one. As they have enormous instability in their price, predicting their fluctuations is a challenging task. Although several research used traditional statistical and economic methods to discover the driving variables of cryptocurrency prices, progress in developing prediction models for decision-making tools in investing techniques is still in its early stages. Many different cryptocurrency price prediction methods cover purposes, such as forecasting a one-step approach that can be done through time series analysis, neural networks, and machine learning algorithms. However, realizing the trend of a coin in the long run is required. In this paper, we aimed to investigate and forecast Bitcoin prices based on three other well-known coins (i.e., Ethereum, Zcash, and Litecoin), assuming little information about Bitcoin prices using machine learning algorithms. Moreover, we proposed a new method to predict Bitcoin's price considering different cryptocurrencies prices. The results demonstrated that Zcash had the best performance in forecasting Bitcoin price without information on Bitcoin price fluctuations among other cryptocurrencies.

1. Introduction

Time series prediction methods are used in various areas, such as climate change, population variations, commodities and product demand, and financial markets [1]. Economic time series is one of the most prevalent and practical branches of time series prediction, which can provide accurate and precise insight into financial markets. There are many different financial

*. Corresponding author. Tel./Fax: +98 21 88350642
   E-mail addresses: maleki.negar@ut.ac.ir (N. Maleki);
   alirezanikoubin@ut.ac.ir (A. Nikoubin); mrabani@ut.ac.ir
   (M. Rabbani); yasser.zeinali@gmail.com (Y. Zeinali)

time series data, such as stock and bond prices, energy prices, etc. In the financial field, researchers investigate various assets through different aspects to discover their relationships.

Cryptocurrency is an advanced resource intended to function as a trade mechanism that utilizes solid cryptography to verify monetary exchanges; every unit can be controlled through different applications. Cryptographic forms of money use decentralized control instead of fiat currency which uses centralized financial systems [2]. The decentralized control of every digital currency works through circulated record innovation, ordinarily, a blockchain that fills in as an open budgetary exchange database. Bitcoin, first discharged as open-source programming in 2009, is commonly viewed as the principal decentralized digital money. Since

Bitcoin's introduction, more than 4,000 digital currencies have been developed, such as Ethereum, Litecoin, Zcash, etc. Unlike traditional credit currencies, Bitcoin allows people to transfer money or purchase merchandise via a peer-to-peer payment network without a centralized regulatory body [3].

Like other assets, cryptocurrency prices have statistical features such as average, variance, and covariance that may change over time. This specification is called non-stationary behavior, such as random walks, trends, cycles, or combinations of the three. Non-stationary data are unpredictable and cannot be modeled or forecasted [4]. The non-stationary time series results may be unreliable because they indicate a relationship between two variables where one does not exist. The non-stationary data should be transformed into stationary data to provide valid results. In contrast to the non-stationary process, which has a variable variance and an average that does not remain close to the historical mean of price returns, the stationary process fluctuates around a constant long-term average. It has a stable variance independent of time. To this end, the time series data should be transformed into stationary by different processes [5].

There are various purposes for using time series. Curve fitting, for example, is the way to foresee a curve or mathematical function that has the best fit to a series of information fields [6]. Function approximation, choosing a function among the characterized class that intently coordinates an objective function in an explicit errand manner, prediction, etc. Prediction and forecasting in time series are developed by various time series and Artificial Intelligence (AI) methods like Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), Autoregressive Conditional Heteroskedasticity (ARCH), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Artificial Neural Networks (ANNs).

The main contribution of this paper is forecasting the price of cryptocurrencies using a new approach based on the cointegration of different currencies to make them stationary for being predictable and determine the best coordination of cryptocurrencies. Analyzing the cointegration has two steps; first, we regress the dependent variable on the independent variable. The statistical tests should be run on equations to test the null hypothesis, the condition in which both currencies converge. Second, we estimate the best lag of models based on different performance criteria.

The rest of this paper is organized as follows: Section 2 reviews different estimating models of digital currency time series found in the literature. Section 3 exhibits the outline of what is done in this study and its methods to make them clear and understandable. Section 4 depicts the recommended approaches and their scenarios in foreseeing prices. The best empirical construction is described in Section 5. Finally, Section 6 concludes the study, and the outcomes of the proposed model are discussed.

2. Literature review

ARIMA and lagged forward price modeling were used to forecast the price of lead and zinc prices. Dooley and Lenihan [7] indicated that both methods could be helpful in price prediction in mine industries. Due to zinc corresponding results, there was no significant difference that one model outperformed the others in forecasting ability. However, in terms of the lead price prediction, it was discovered that the ARIMA model works better. Ediger and Akar [8] used the Seasonal ARIMA (SARIMA) and regular ARIMA techniques to assess Turkey's future essential energy demand requests from 2005 to 2020. Khashei and Bijari [9] utilized a hybrid methodology that combines ANN and ARIMA for forecasting time series. They compared their model to different methods, such as ARIMA, ANN, and Zhang's hybrid model, and they achieved better performance. Their model's Mean Square Error (MSE) was considerably smaller than utilizing only the ANN or ARIMA models in prediction. Their forecast horizon was 1, 6, and 12 months, and the one-step-ahead forecast had the smallest MSE.

Khashei and Bijari [10] utilized ARIMA models' excellent capacity in linear displaying to distinguish and strengthen the current linear structure in the information. Afterward, a Multilayer Perceptron (MLP) was employed to decide on a model to catch the basic information, create a process, and forecast the future using preprocessed information. Experimental outcomes with three real-data collections demonstrated that the proposed model should pass a rough path to yield more capable results than Zhang's hybrid one. Therefore, both models utilized independently can be an elective determining tool for anticipating time series. Another research introduces an adaptable calculation dependent on ANN and Fuzzy Regression (FR) to adapt to long-time oil prices in uncertain, complex conditions [11]. Analysis Of Variance (ANOVA) and Duncan's Multiple Range Test (DMRT) was connected to investigate the results obtained from ANN and FR models. It was demonstrated that the chosen ANN models significantly outperformed the FR models considering Mean Absolute Percentage Error (MAPE) as the comparison benchmark. Besides, the Spearman relationship test was employed to check and approve the results [12].

Khandelwal et al. [13] developed a new approach using Discrete Wavelet Transform (DWT) to improve

the time series accuracy. DWT was employed to break down the in-test preparing time series dataset into two parts: linear and non-linear. Therefore, the ARIMA and ANN models independently perceived and individually predicted the remaining estimated parts. The experimental outcomes with four certifiable time series showed that the proposed technique resulted in prominently and preferred figures over ARIMA, ANN, and Zhang's hybrid model. Buncic and Moretto [14] collected data from the London Metal Exchange to estimate month-to-month copper prices using the proposed Dynamic Model Averaging and Selection (DMA/DMS) structure. They also developed eighteen predictor variables, including different financial data series. The data covered an out-of-sample period of ten years, engaging standard statistical measurement criteria and demonstrating that the R square of DMA can be as high as 18.5%.

In another study, the researcher proposed a method for predicting minimum and maximum day stock prices for Brazilian distribution companies based on feature selection and ANN [15]. They concluded that it is vital to determine what kinds of features should be in the database when working with time series for prediction. Therefore, considering different attributes, companies' stocks can be influenced differently. Fan et al. [16] established an MLP network model to predict short terms after recognizing coal cost's disorganized qualities. The model performance was evaluated using four different MAPE, Root Mean Square Error (RMSE), direction statistic, and Theil index. They indicated that the MLP acted better than the ARIMA model. Gangopadhyay et al. [17] presented a model to foresee gold prices in India, utilizing a vector error correction model. They recognized speculation choice and inflation support as prime stimuli of their information. The model showed that the gold price had a long-term relationship with the stock market index, oil price, exchange rate, and oil price.

Pradeepkumar and Ravi [18] developed a Particle Swarm Optimization trained Quantile Regression Neural Network (PSOQRNN) to anticipate time series. They contrasted the adequacy of PSOQRNN and the conventional instability anticipating models, i.e., GARCH and a few ANNs. They demonstrated that their proposed model overcame the others in terms of MSE on a larger part of the eight money-related time series. In another paper [19], Long-Short Term Memory (LSTM) networks were connected to a vast-scale monetary market forecast mission on the S&P 500 from December 1992 to October 2015. An LSTM network can successfully remove significant data from money-related time-series information. Contrasted with irregular timberlands, standard profound nets, and calculated relapse, it was the technique for a decision concerning predicational precision and regard-

ing day-by-day returned after exchange costs [19]. Sezer and Ozbayoglu [20] proposed an arrangement that analyzed financial time-series information and changed it into 2-D pictures. In this investigation, they endeavored to predict passage and leave purposes of the time series values as "Buy," "Sell," and "Hold" marks for beneficial exchanges. They used Dow Jones 30 stock costs and ETFs as money-related time-series information. The outcomes showed that their proposed method outperformed other similar models.

In another paper, the Economic Policy Uncertainty (EPU) forecast on a day-by-day scale was evaluated based on Bitcoin prices from 18th July 2010 to 15th November 2017 [21]. Utilizing the Bayesian Graphical Structural Vector Autoregressive (BGSVA) model just as the Ordinary Least Squares (OLS) and the Quantile-on-Quantile Regression (QQR) was evaluated to assess the results. At that point, it noticed that the EPU had a prescient power on the Bitcoin restores.

Chou and Tran [22] offered an extensive revision and examination of a few time-series anticipating procedures for foreseeing the energy utilization of structures one-day ahead of time. Numerous investigations were executed using five AI methods, i.e., ANN, Support Vector Regressor (SVR), Classification and Regression Tree (CART), and Linear Regression (LR), with data on energy consumption. A general analogy illustrated that hybrid models are more effective than single ones. Furthermore, Kim and Won [23] used LSTM with GARCH to forecast stock market volatility. Then they compared the results of their hybrid model with single models and achieved a better outcome.

Considering the price spike through Phillips et al. [24] approach, allowed researchers in [25] to capture the extreme price movements of cryptocurrencies. Moreover, it expanded limited knowledge of Bitcoin's and other cryptocurrencies' short- and long-term price behavior. Bouri et al. [25] stated that Bitcoin, in particular, was found to be more subject to long-lived explosivity. Furthermore, logistic regression analysis showed that the probability of an explosion in a cryptocurrency depends on an explosion in others.

Bouri et al. [26] examined herding behavior in the cryptocurrency market and the outcome of mass collaboration and imitation by investigating a window-rolling analysis on the data series illustrated a significant herding behavior that changed over time. Besides, logistic regression showed that herding behavior occurs with increasing uncertainty. Their results provided portfolio and risk managers and trading strategies with helpful insights.

A set of measures to examine connectedness via return and volatility spillovers across six prominent cryptocurrencies was addressed in [27]. The results showed that Bitcoin and Litecoin were at the center

of the connected returns network. The shocks arising from these two cryptocurrencies had the most significant effect on the others. Connectedness via positive returns was not as strong as positive ones, and as far as the volatility spillovers were concerned, Bitcoin was the most influential cryptocurrency. Further analyses illustrated that trading volumes, global financial, and uncertainty effects were determinants of net directional spillovers.

A proposed methodology based on feature selection and machine learning methods was applied to the Bitcoin exchange rates, providing better results than the previous studies. Mallqui and Fernandes [28] utilized feature selection methods to find the most relevant attributes, and then by analyzing the best attributes, they explained the behavior of the daily variation of the exchange rate.

Alameer et al. [29] did long-term gold price forecasting. They considered a new metaheuristic method named Whale Optimization Algorithm (WOA) for training the Neural Network (NN). The result was compared with the other metaheuristic optimization methods such as Genetic Algorithm (GA), PSO, and grey wolf; then they employed an ARIMA model. The results have shown that the hybrid of WOA-NN was superior to the other mixed methods. The proposed WOA-NN displays progress in the foreseeing precision acquired from the NN, PSO-NN, GA-NN, GWO-NN, and ARIMA models by 41.25%, 24.19%, 25.40%, 25.40%, and 85.84% was diminishing in MSE, individually.

Two-hybrid anticipating models were recommended to improve the exactness of the stock market costs foreseeing, incorporating the two sorts of Empirical Mode Decomposition (EMD) with the LSTM [30]. Atsalakis et al. [31] utilized a complex neuro-fuzzy technique to forecast the direction of the Bitcoin price changes and compared the result with the simple approach of neuro-fuzzy and ANN methods. They realized that the efficiency of their method was 71.21% higher than the other techniques applied in their paper.

Bouri et al. [32] focused on the cryptocurrency market that intrigued investment communities, not least because of its substantial price spikes and slippage. The results proposed that the event of a jump in one cryptocurrency increments with a jump in other cryptocurrencies. Moreover, the discoveries inferred needing to incorporate jumps and co-jumps when modeling cryptocurrencies' volatility dynamics within multivariate models.

Correlations were found between the volatility surprises of eight prominent cryptocurrencies in [33]. To address the communications, they were based on volatility surprises and examine the causal relationships between the volatility of leading cryptocurrencies through frequency-domain testing, distinguishing between transient and permanent causes. The results showed causality among volatility surprises that did not necessarily come from the most prominent cryptocurrency. The fact that the reasons are different between short-term and long-term for some cryptocurrencies indicates the importance of low- and high-frequency causes when considering integration. The study also suggested predicting the volatility surprise at specific frequencies. In other words, that crypto-traders should monitor other cryptocurrencies when trading a cryptocurrency.

Another study investigated the multi-scale dependence and market relationship with the five leading cryptocurrencies [34]. The top results of wavelet-based techniques covering time and frequency showed that the cryptocurrency market's unstable trends coincided with high frequencies. The period 2016–2018 shows the continuation of market contagions.

The comparison of the literature review is shown in Table 1, and according to the literature, the main contributions of this paper are as follows:

- Checking the cointegration of Bitcoin with three other popular coins to determining a coin that has the highest ability to be used as a predictor of Bitcoin price, especially in a situation where there is little information about Bitcoin;

- Forecasting the Bitcoin price based on other coins, using machine learning algorithms such as LR, Gradient Boosting Regressor (GBR), SVR, Random Forest Regressor (RFR);

- Utilizing time series models to recognize which coin fits the historical data better than others in the cointegration process.

3. Methodology

Fundamental techniques, along with four well-known cryptocurrencies, are described in this section. We then describe our proposed method to predict Bitcoin price based on the other digital currencies.

3.1. Data description

For the practical application, four-digital coins were collected from coinmarketcap.com from 1st April 2018 to 31st March 2019. These are as follows:

- **Bitcoin:** Bitcoin (฿) is a digital currency, a type of electronic money. It is decentralized advanced cash without a national bank or single chairman that can be sent from client to client on the shared Bitcoin arrangement without middle people's requirement.

    Exchanges are checked by system hubs through cryptography and recorded in an open disseminated record called a blockchain. Bitcoin was designed by an obscure individual or gathering of individuals

Table 1. The comparison table of literature review.

Authors	Methods														Data
	ARIMA	ANN	Fuzzy method	Zhang's hybrid	SARIMA	SVR	LSTM	CART	DMA	DMS	GARCH	BGSVA	EMD	Meta heuristic	
Dooley and Lenihan [7]	✓														Lead/zinc price
Ediger and Akar [8]	✓				✓										Energy demand
Khashei and Bijari [9]	✓	✓		✓											Sunspot
Khashei and Bijari [10]	✓	✓		✓											Sunspot
Azadeh et al. [12]		✓	✓												Oil price
Khandelwal et al. [13]	✓	✓		✓											Lynx price
Buncic and Moretto [14]									✓	✓					Copper price
Fan et al. [16]	✓	✓													Coal price
Pradeepkumar and Ravi [18]		✓									✓			✓	Different currency
Fischer and Krauss [19]							✓								S&P 500
Demir et al. [21]												✓			Bitcoin price
Chou and Tran [22]		✓			✓	✓		✓							Energy consumption
Kim and Won [23]							✓				✓				Stock market prices
Alameer et al. [29]	✓	✓												✓	Gold price
Cao et al. [30]							✓						✓		Stock market prices
Atsalakis et al. [31]		✓	✓												Bitcoin price

known as Satoshi Nakamoto and discharged as open-source programming in 2009. Bitcoins are made as a fee for an operation known as mining. They can be dealt with for various fiscal structures, things, and organizations. Research delivered by the University of Cambridge gauges that in 2017, there were 2.9 to 5.8 million remarkable clients using a digital currency wallet, the vast majority of them using Bitcoin.

- **Ethereum:** Another type of cryptocurrency is Ethereum, which was founded in 2015. It is based on a modified version of Nakamoto's theory. Ethereum was suggested in late 2013 by Vitalik Buterin, a digital cash specialist and programmer. Ethereum is the pioneer for blockchain-based smart contracts. Smart contracts consider code to be run as customized with no plausibility of downtime, control, deceit, or outsider impedance on the blockchain. The worth of the Ethereum token raised more than 13,000% in 2017 to over $1400. By September 2018, it had decreased to $200.

- **Litecoin:** Litecoin (LTC or L) is a distributed digital money and open-source programming project discharged under the MIT/X11 license. The cre-

ation and exchange of coins depend on an open-source cryptographic convention and are not over-seen by any focal expert. Litecoin was an early Bitcoin spinoff or altcoin, beginning in October 2011. In specialized subtleties, Litecoin is about indistinguishable from Bitcoin.

- **Zcash:** Zcash is a digital currency that uses cryptography to improve its clients' security, contrasted with other digital money, such as Bitcoin. Zcash has a total fixed supply of 21 million units. Exchanges can be "transparent" and similar to Bitcoin exchanges, in which case a t-addr constrains them, or can be a kind of zero-learning verification called zk-SNARKs; the exchanges are then said to be "shielded" and are constrained by a z-addr.

3.2. Statistical models and AI methods
This part explains several prevalent statistical models, such as AR, MA, ARIMA, and ANFIS, as intelligent methods that combine fuzzy and ANNs.

3.2.1. Autoregressive (AR)
An AR model illustrates a sort of random process; all things considered, it is used to define time-fluctuating procedures in nature and financial aspects. In machine

learning, an AR model achieves from a series of timed steps and accepts estimations from past activities as inputs for a regression model to predict the next time value. AR models expose how time-series information are consecutive – it is essential to preserve the order or sequence of the data to make exact forecasts. The forecasts for AR models are regressed on previous observations in the time series. These models use weighted sums of past values to predict future ones. The notation $AR(p)$ shows an AR model of order p. It is characterized as:

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t, \tag{1}$$

where φ_i is the *model's parameter*, c is a constant, and ε_t is white noise – A white noise process is one with a mean zero and no correlation between its values at different times.

Once the above equation's parameters have been estimated, the autoregression can be used to forecast an arbitrary number of periods into the future.

3.2.2. Moving Average (MA)

In time series techniques, the MA model, called the moving-average procedure, is a typical methodology for displaying univariate time series. The moving-average model determines that the upshot variable depends directly on the present and different past values of a stochastic term. A MA model of order q (i.e., $MA(q)$) is characterized as follows:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \tag{2}$$

where μ is the mean of the series, $\theta_1, \cdots, \theta_q$ are the parameters of the model and the $\varepsilon_t, \cdots, \varepsilon_{t-q}$ are white noise error terms.

3.2.3. Autoregressive Integrated Moving Average (ARIMA)

ARIMA, in principle, is the broadest class of anticipating models for a time series which may become "stationary" by subtracting, and be related to nonlinear changes, for example, logging or deflating. It also is a speculation of AR and MA models. Both models are fitted to time series information to predict future points in the series. The AR part of ARIMA specifies that the interest's evolving variable is regressed on its lagged values. The MA part denotes that the regression fault is a linear compound of fault terms whose values coincided at various times in the past. The "I" shows that the data values are altered with the diversity between their values and the olden values. Every one of these attributes aims to fit the model throughout the data beyond what many would consider possible.

ARIMA models are commonly $ARIMA(p, d, q)$ where parameters p, d, and q are positive integers, p is the order of the AR model, d is the level of subtracting,

and q is the order of the MA model. This model is as follows:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1 - L)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right)\varepsilon_t. \tag{3}$$

3.2.4. Machine Learning (ML)

One of AI applications is machine learning, which gives systems the capability to learn and make better from experience without any definitive program spontaneously. Computer programs have developed by accessing data and utilizing it for learning themselves in machine learning.

The way toward learning starts with perceptions or information to search for patterns in data and settle on better choices, then, based on the precedents. The essential point is to permit the PCs to adapt naturally without human mediation or help and change activities in like manner. Machine learning algorithms refer to the type of targets that can be divide into regression models, classification models, and clustering models.

3.3. Cointegration

Cointegration is a statistical measurement of a collection $(X_1, X_2, \cdots, X_k)$ of time series variables. If at least two series are exclusively coordinated. In any case, some linear compound has a lower order of integration; at that point, the series is said to be cointegrated. The three primary techniques for testing cointegration are:

- Engle-Granger two-step method;
- Johansen test;
- Phillips-Ouliaris cointegration test.

Several data types are cointegrated, such as economic substitutes such as heating oil and natural gas, platinum and palladium, corn and wheat, and corn and sugar.

3.4. Our proposal framework

In this part, we discuss the theoretical procedure of the experiment. Firstly, we analyze the changing price data of four digital currencies to realize the main trends and discover if they are stationary. As mentioned before, the first purpose of this paper is to forecast the price of Bitcoin based on other digital currencies. It is assumed that price fluctuations in each digital currency follow the random walk process. Therefore, it cannot be predicted as it is non-stationary. There are many ways to make a time series stationery. Several statistical tests determine whether a process is a random walk, such as Dickey-Fuller (DF) test, the Augmented Dickey-Fuller (ADF) test, the Ng-Perron test, the DF-GLS test, ERS-point optimal test, and Phillips-Perron test. This paper tries to make a process (daily price of Bitcoin) stationary by subtracting the other process. Cointegration was checked using different models.

Checking the cointegration of two processes has two steps. We regress the first variable (Y) on the second variable; therefore, the equation will be $Y - bX$, and we should run the ADF test or any other statistical test on it to check if it is still a random walk. Estimating the regression line for checking the cointegration can be executed through the following methods:

1. Ordinary Least Squares (OLS);

2. Fully Modified Ordinary Least Squares (FMOLS);

3. Dynamic Ordinary Least Squares (DOLS).

After checking the stationarity condition, statistical time series models such as AR, MA, ARIMA, and AI model such as neuro-fuzzy model can be applied. Also, we try to run a code to tune the parameters based on the performance criteria, such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Then we try to cover all of the performance criteria to evaluate each method, such as:

1. Mean Square Error (MSE);

2. Mean Absolute Percentage Error (MAPE);

3. Mean Absolute Error (MAE);

4. Akaike Information Criterion (AIC);

5. Bayesian Information Criterion (BIC).

4. Data analysis and models implementation

In this section, the behavior of cryptocurrencies' prices, time series plots, models' execution, and evaluation criteria are discussed in detail. Then, the cointegration between Bitcoin, as a fixed part, is checked with the other cryptocurrencies.

4.1. Time series plots

As mentioned, cryptocurrencies' datasets indicated daily price changes throughout the year. The most crucial application of analyzing a time series plot is to determine whether the process is stationary, along with this part. Therefore, the Auto Correlation Function (ACF) can be a valuable method to discover it. Time series are plotted in Figure 1 for this purpose.

As shown in Figure 1, the price peaked in May
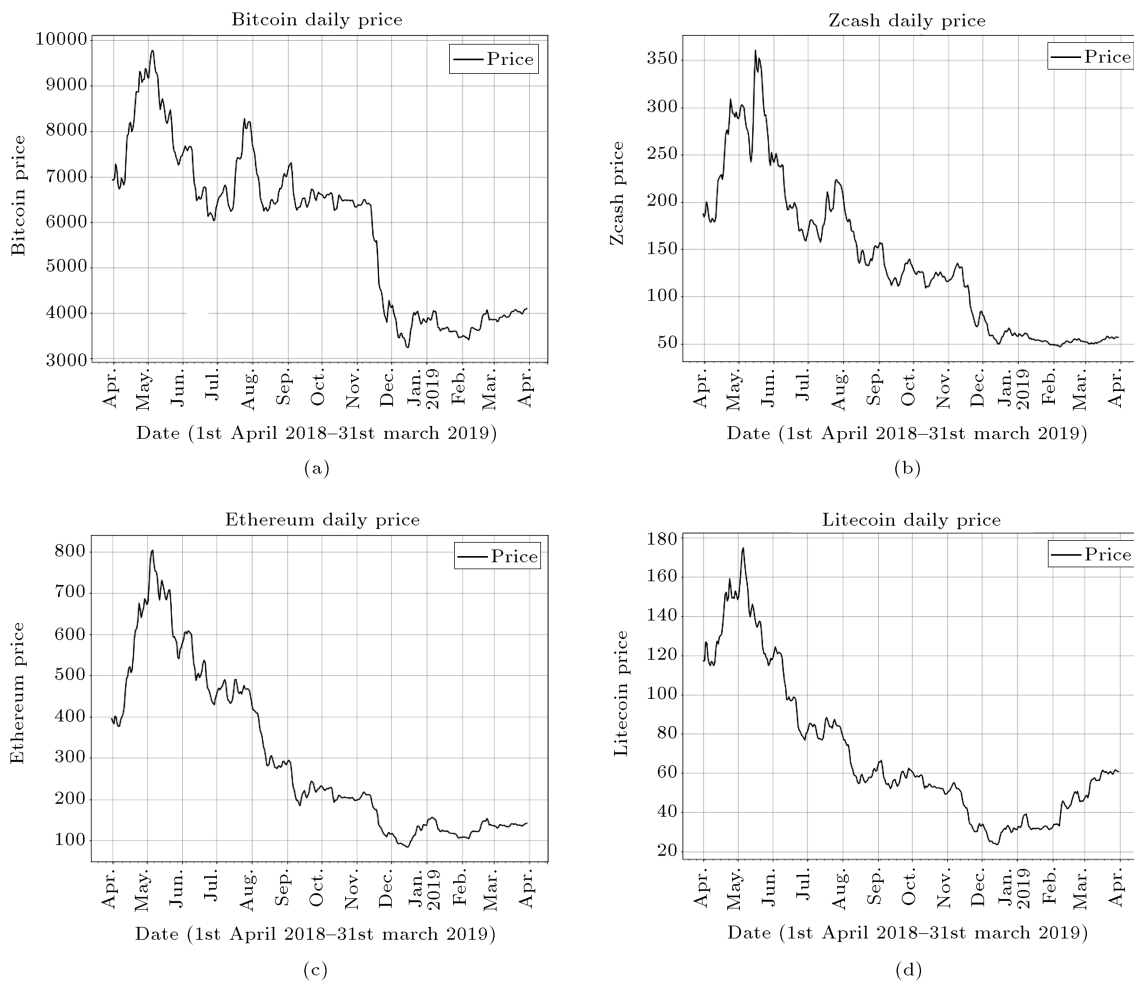
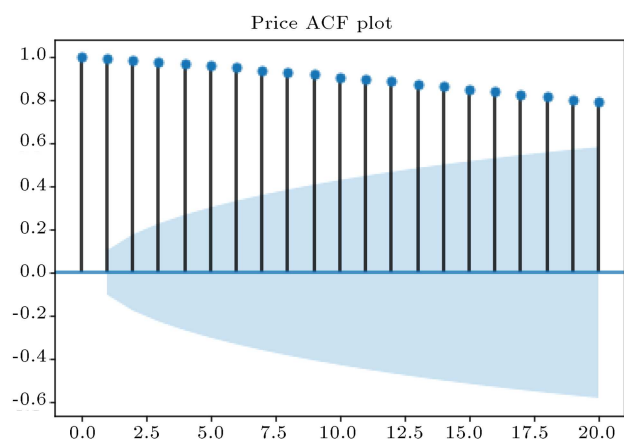**Figure 1.** Time series plots for (a) Bitcoin, (b) Zcash, (c) Ethereum, and (d) Litecoin.
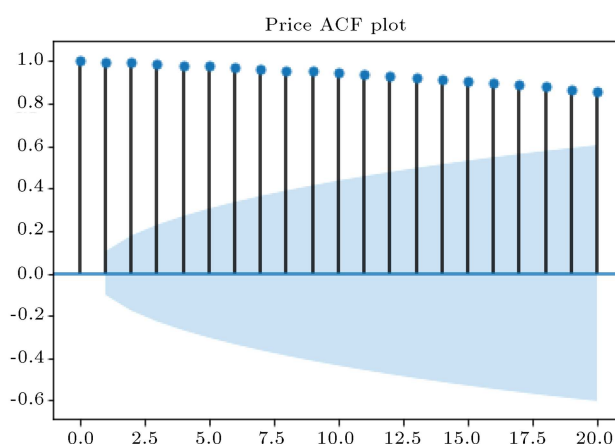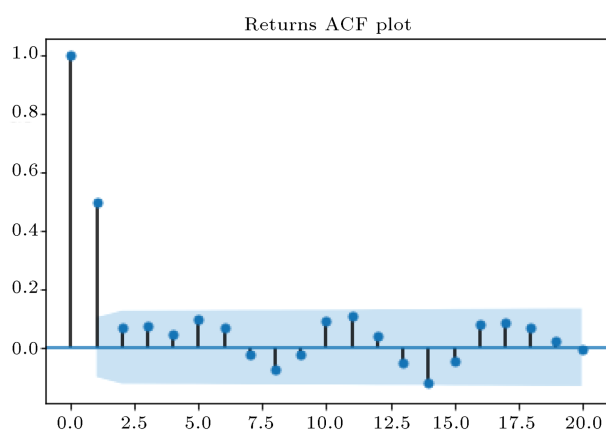
Figure 2. Bitcoin prices' ACF.

Figure 3. Bitcoin returns' ACF.

Figure 4. Bitcoin returns' change.

Figure 5. Ethereum prices' ACF.

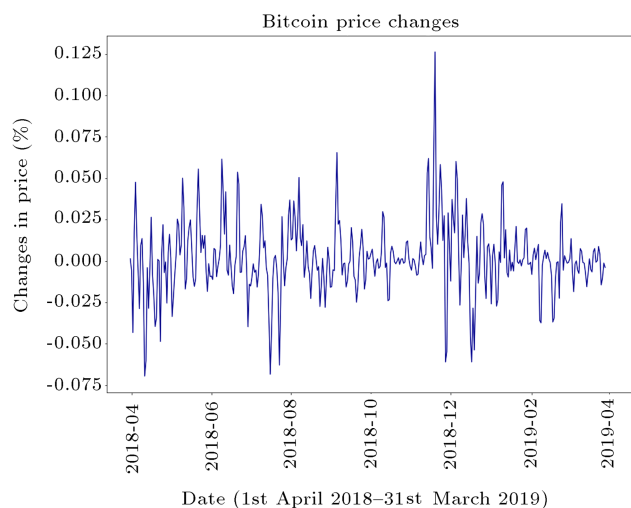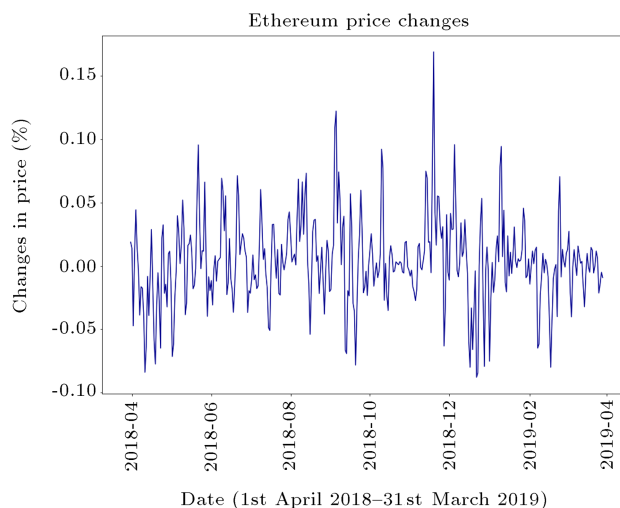**Figure 6.** Ethereum returns' ACF.

Figure 7. Ethereum returns' change.

and gradually fell until December. It showed that they are full of changes, making forecasting complex because the price of cryptocurrencies, such as stocks, follows a random walk. Hence, they are not predictable, which makes us use another attribute to forecast them. This attribute is recognized as "return," a difference that makes the process stationary. In Figures 2–13, firstly,

the prices and returns' ACF are shown. Secondly, their returns are plotted to ensure differentiating works as expected.

Figures 2–4 show the difference between the autocorrelation in Bitcoin price and Bitcoin price returns. As shown in Bitcoin prices' ACF, the Bitcoin price's
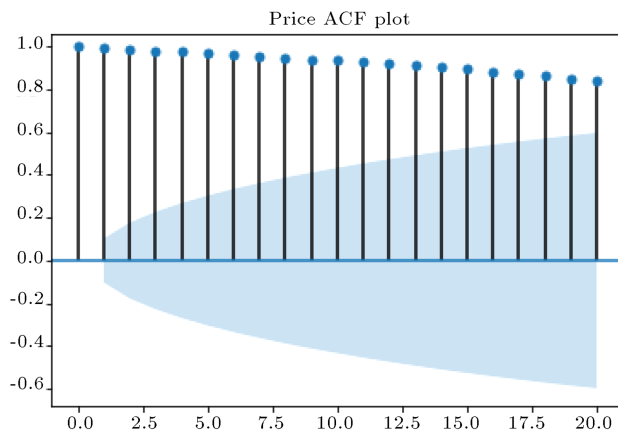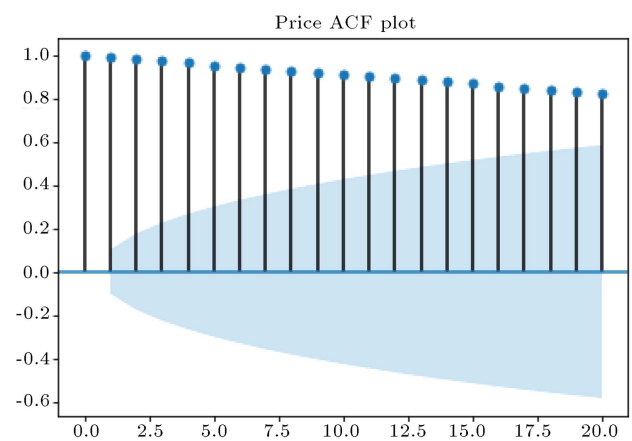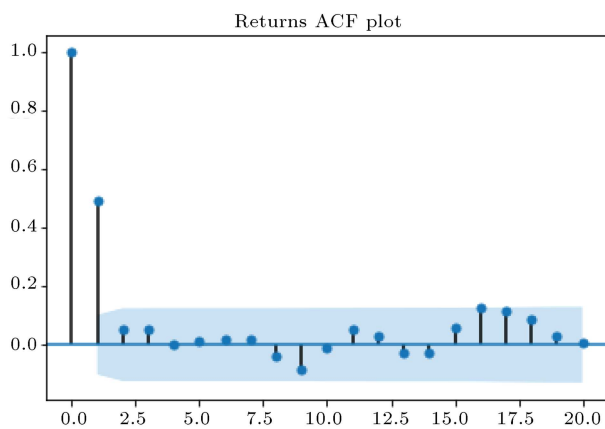
Figure 8. Litecoin prices' ACF.

Figure 9. Litecoin returns' ACF.

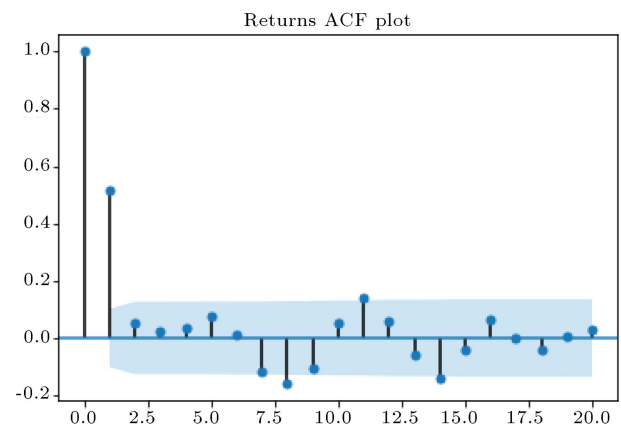

Figure 10. Litecoin returns' change.
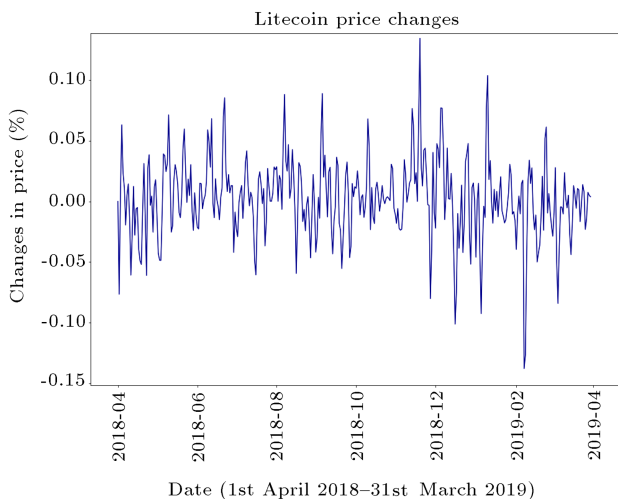

Figure 11. Zcash prices' ACF.

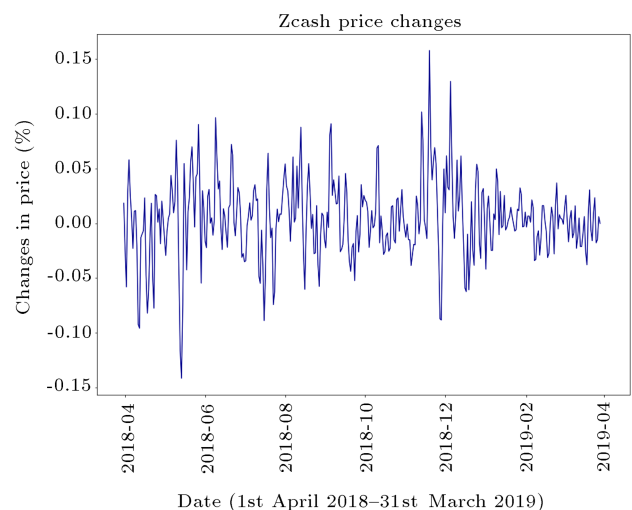**Figure 12.** Zcash returns' ACF.

Figure 13. Zcash returns' change.

ACF plot is not stationary, but we can see that the returns are entirely stationary.

As shown in Figures 5–13, Ethereum, Litecoin, and Zcash also have the same condition as Bitcoin has. Their price fluctuations are not stationary, but we notice that they are stationary if we examine their price returns.

4.2. MA, AR, and ARIMA models for price returns

First, we apply the ARIMA models to each cryptocurrency's price returns to predict one step forecast. We use ARIMA because it includes MA and AR models as we apply ARIMA $(0, 0, q)$ and ARIMA $(p, 0, 0)$, respectively. In the previous subsection, we show

all the returns were stationary; thus, we can apply the foreseeing models. The ACF plots have shown that returns are not so simple, and we cannot execute a pure AR or MA model on them. Figures 14-17 simultaneously contain actual values, fitted values, and the forecasting part. As usual, the confidence interval is about 95% percent. These results are shown in Figures 14–17.

Different parameter values are tested on Bitcoin's returns, and the best outcome was $p = 1$ and $q = 1$ with one difference.

These figures are based on the ARIMA models, and the best values for parameters were obtained using different programs. After the tuning process, we plotted the final results to show the function of the best models. Table 2 shows the quick results of the programs. The same processes as Bitcoin have been executed for the rest of the cryptocurrencies, and the best model for each one is ARIMA(2, 1, 1), ARIMA(2, 1, 1), and ARIMA(1, 1, 1) for Ethereum, Litecoin, and Zcash, respectively.

4.3. Machine learning prediction

Machine learning algorithms are employed to predict the Bitcoin price and realize which cryptocurrencies (Ethereum, Zcash, and Litecoin) impact Bitcoin's price changes more. To use different ML algorithms, firstly, we must set a dataset that contains Ethereum's, Zcash's, and Litecoin's prices as independent variables and Bitcoin's price as the dependent variable. Then, ML algorithms are applied to find out which cryptocurrencies can forecast Bitcoin's price changes without any data on the trend of Bitcoin's price.

Regression methods are one of the forecasting algorithms in machine learning. Moreover, to achieve referred goals, we use Linear Regression (LR), GBR, SVR, Lasso Regression (LR), and RFR algorithms.

In each method, 80 percent of data is assigned to the training phase, and the rest is devoted to the testing phase; furthermore, this study considers MSE as the comparison and performance criterion. The MSE formula is as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2. \tag{4}$$

4.3.1. SVR algorithm
Support Vector Machines algorithm (SVM) as an ML method is divided into: (1) classification and (2) regression. According to our goal, SVR is operated with a "Linear" kernel, MSE is calculated, and a comparison figure between real Bitcoin's price values and predicted ones by SVR is shown in Figure 18.

$$MSE = 516804.798.$$

4.3.2. LR algorithm
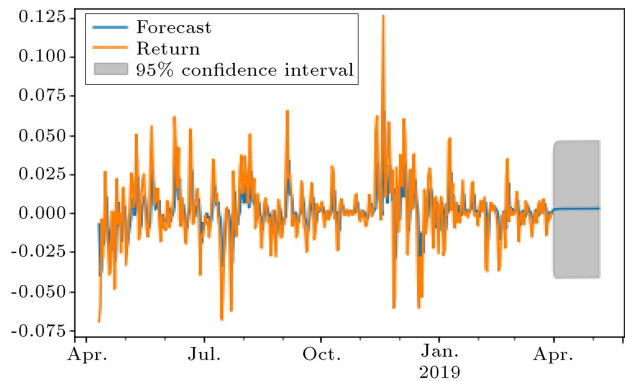The LR algorithm is one of the simplest methods
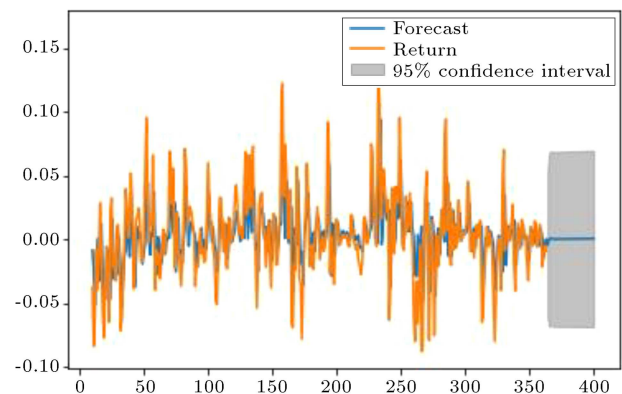

Figure 14. ARIMA model for Bitcoin return.

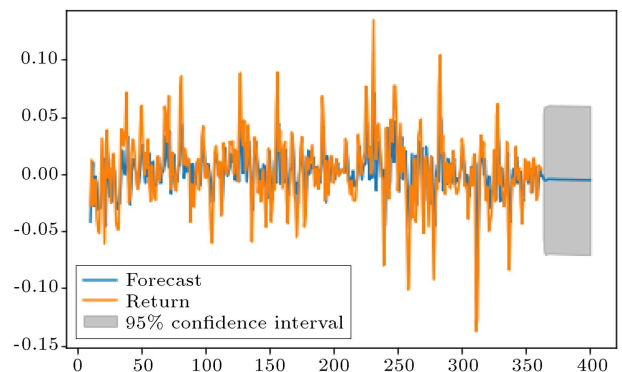**Figure 15.** ARIMA model for Ethereum return.

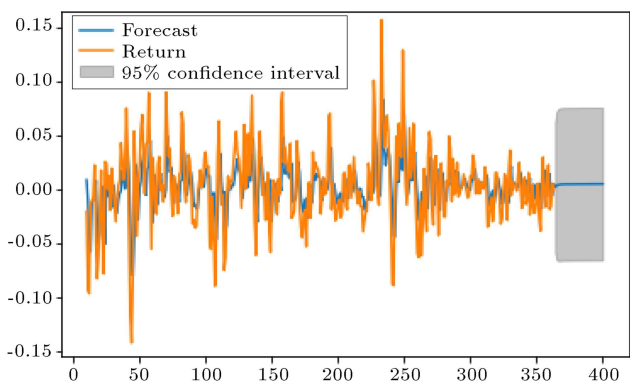**Figure 16.** ARIMA model for Litecoin return.

Figure 17. ARIMA model for Zcash.

Table 2. Parameters tuning.

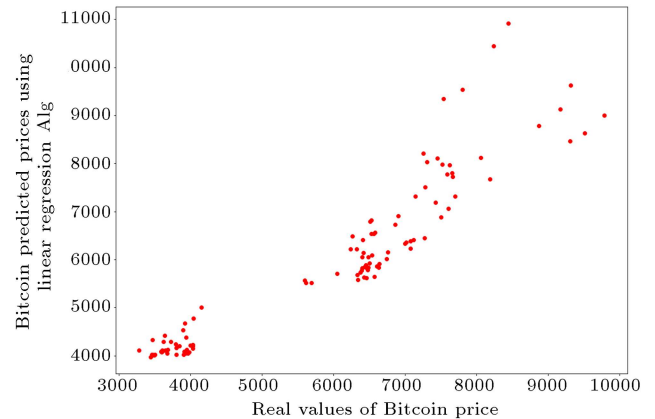| Bitcoin						Litecoin					
ARIMA(p, d, q)	**p**	**q**	**d**	**AIC**	**BIC**	**ARIMA(p, d, q)**	**p**	**q**	**d**	**AIC**	**BIC**
Model 1	1	0	0	2104.20	2101.30	Model 1	1	0	0	2010.27	2008.10
Model 2	1	1	0	2056.20	2040.31	Model 2	1	1	0	1960.42	1954.00
Model 3	1	1	1	1821.86	1806.29	Model 3	1	1	1	1921.01	1916.89
Model 4	0	1	1	2423.25	2418.32	Model 4	0	1	1	2003.21	1982.24
Model 5	0	0	1	2843.21	2820.45	Model 5	0	0	1	2413.51	2401.21
Model 6	2	1	0	2024.32	2010.02	Model 6	2	1	0	1650.31	1623.74
Model 7	2	0	1	2112.02	2118.32	Model 7	2	0	1	1580.27	1560.12
Model 8	2	1	1	2090.12	2095.74	Model 8	2	1	1	1557.62	1538.12
Model 9	1	2	1	2312.45	2308.42	Model 9	1	2	1	1904.15	1890.42
Zcash						Ethereum					
ARIMA(p, d, q)	**p**	**q**	**d**	**AIC**	**BIC**	**ARIMA(p, d, q)**	**p**	**q**	**d**	**AIC**	**BIC**
Model 1	1	0	0	1510.48	1506.14	Model 1	1	0	0	1913.18	1903.37
Model 2	1	1	0	1492.12	1488.25	Model 2	1	1	0	1991.51	1983.41
Model 3	1	1	1	1480.31	1464.73	Model 3	1	1	1	1921.01	1916.89
Model 4	0	1	1	1743.34	1742.17	Model 4	0	1	1	1942.72	1934.00
Model 5	0	0	1	1800.56	1794.20	Model 5	0	0	1	1973.97	1964.49
Model 6	2	1	0	1624.54	1616.74	Model 6	2	1	0	1705.47	1697.33
Model 7	2	0	1	1647.45	1635.74	Model 7	2	0	1	1682.46	1674.78
Model 8	2	1	1	1552.64	1545.10	Model 8	2	1	1	1645.84	1620.70
Model 9	1	2	1	1734.24	1719.46	Model 9	1	2	1	1812.46	1802.75
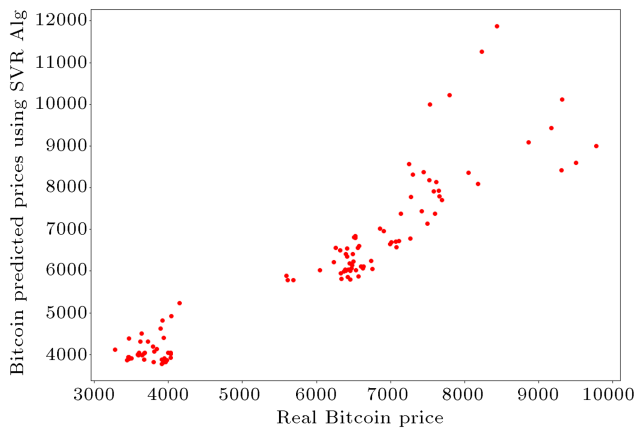
Figure 18. Regression line estimation by SVR.

Figure 19. Regression line estimation by LR.

among machine learning algorithms to produce appropriate answers. The MSE value shows that this method performs better than SVR. Figure 19 illustrates the actual Bitcoin price values and predicted ones by LR.

$$MSE = 421034.06468415906.$$

4.3.3. RFR algorithm
Another ML algorithm is RFR, an ensemble method formed by a few Decision Tree methods. RFR can be used in both classifiers and regression forms. The optimal RFR's parameters are achieved experimentally and 207 trees was the best. The MSE value showed that this method acted better than SVR and LR. The results are represented in Figure 20.

$$MSE = 23365.60049452004.$$

4.3.4. GBR algorithm
Another ensemble method as a famous ML algorithm is GBR, which contains some parameters. The best of this method is achieved when the number of estimators is equal to 100, the learning rate equals 0.12, and the subsample equals 0.6 (based on try and error approach). It can be inferred from comparing the MSE result of this method with others that GBR is the most
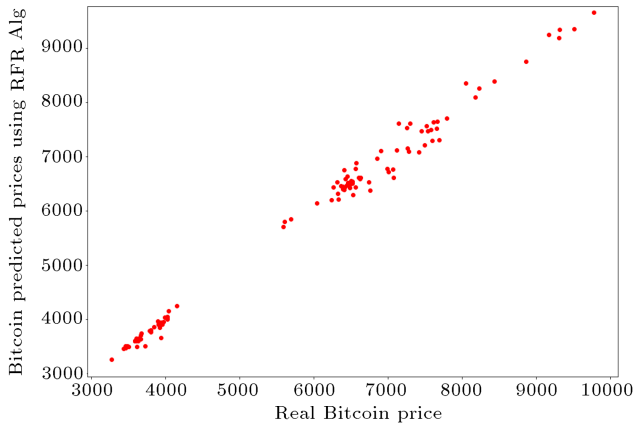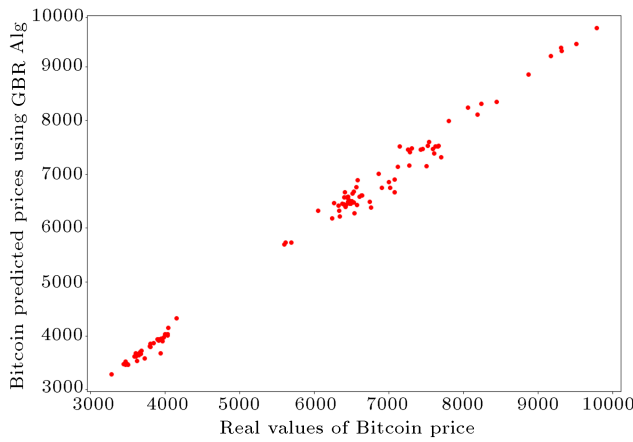
Figure 20. Regression line estimation by RFR.

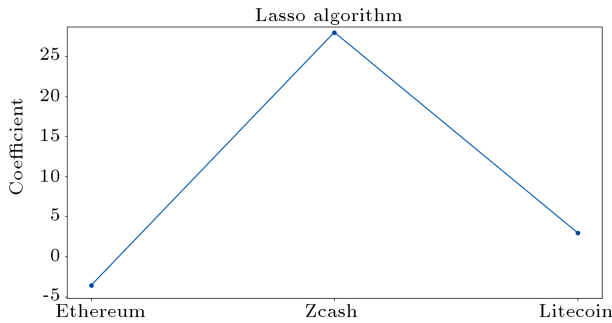

Figure 21. Regression line estimation by GBR.

Figure 22. The best cryptocurrency predictor.

appropriate approach for estimating Bitcoin's price. It is also understandable from the estimated line in Figure 21.

$$MSE = 18359.190867996942.$$

4.3.5. Lasso regression algorithm

The Lasso Regression algorithm is applied to pursue the study's second purpose, determining the most desirable cryptocurrency for predicting Bitcoin's price without any information. The fundamental of this method is based on the data's variance changing. According to Figure 22, we can realize that Zcash's price trend has the most capability to predict Bitcoin's

Table 3. The comparison table of applied algorithms.

Algorithms name	MSE	R^2	R^2 adjusted
SVR	516804.797833455	0.8345	0.8391
LR	421034.06468415906	0.8652	0.8689
RFR	23365.60049452004	0.9925	0.9927
GBR	18359.190867996942	0.994	0.9942

price. As a summary of all applied algorithms, Table 3 represents the used method with other evaluation criteria. The results show that the Gradient Boosting Algorithm has minimum MSE.

4.4. Cointegration, our proposed method

This section investigates the cointegration between Bitcoin and other cryptocurrencies through two steps. Adding this part can help to predict Bitcoin, as a well-known crypto, one step ahead without any information about its previous trend. We recognized the relation between Bitcoin and other cryptos to use the best crypto for foreseeing Bitcoin price.

As mentioned before, the first step is to regress Bitcoin's prices on other cryptos. Then we try to address the augmented Dickey-Fuller test as the second part to discover whether they are cointegrated or not. As the residuals are not stationary, first-order differencing is done to be compatible with ARIMA models. We coded the whole process and algorithms in Python programming language as shown in Figure 23.

After applying time series models on residuals, we evaluate each model and identify the appropriate one. Different measurements such as BIC and AIC are employed to assess the models:

$$AIC = \ln\left(\frac{\sum\limits_{t=1}^{T} e_t^2}{T}\right) + \frac{2p}{T}, \tag{5}$$

$$BIC = \ln\left(\frac{\sum\limits_{t=1}^{T} e_t^2}{T}\right) + \frac{p\mathrm{Ln}(T)}{T}, \tag{6}$$

```
start
{
Add constant value to a crypto [1];
    Result = Ordinary Least Square model (Bitcoin, crypto(x))
    Bitcoin = regression_coeff * crypto[1] + residuals
    residuals = (Bitcoin - regression_coeff * crypto(x))
Dickey fuller test on residuals
    test result
}
end
```

Figure 23. Pseudo code for checking the cointegration of Bitcoin with the rest of the cryptos.
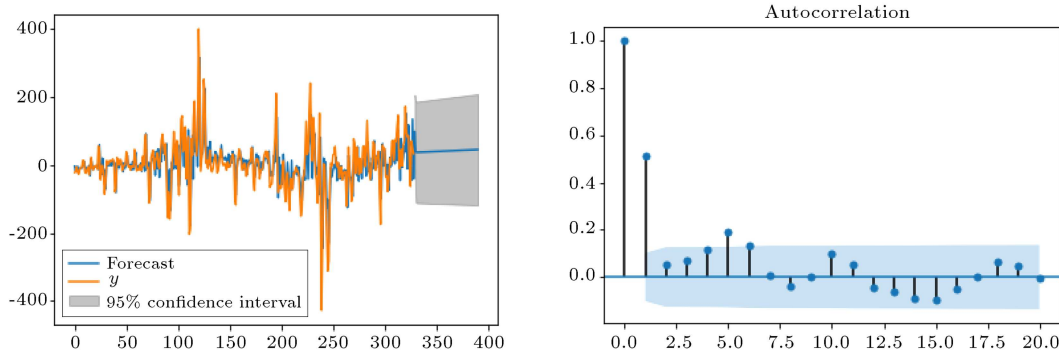
Figure 24. Model execution on residuals Bitcoin-Ethereum (left), and their residuals ACF (right).
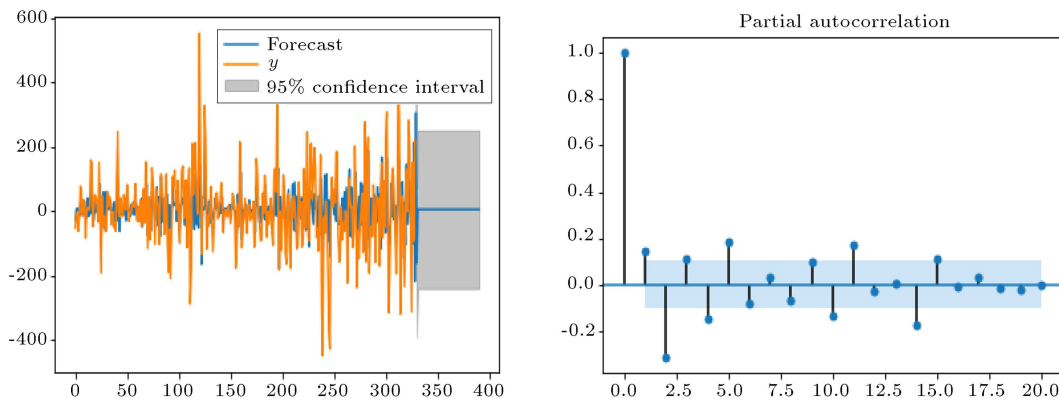

Figure 25. Model execution on residuals Bitcoin-Litecoin (left), and their residuals ACF (right).

where T is the number of periods; p the number of estimated parameters; and e_t the residuals.

Figure 24 shows the forms and order of models, the first condition assigned to cointegration between Bitcoin and Ethereum and ARIMA execution with the order of $p = 1$, $d = 1$, and $q = 2$.

The regression line is (Eq. (7)):

$$Bitcoin\_Price = 3510.527428 + 7.9343$$
$$\times \ Ethereum\_Price. \qquad (7)$$

P_value of the ADF-test:

3.018049e-05

AIC and BIC of their cointegration:

$$AIC = 4087.4245922612673,$$

$$BIC = 4192.7285924245916.$$

The second condition assigned to cointegration between Bitcoin and Litecoin and ARIMA execution with the order of $p = 0$, $d = 0$, and $q = 2$ (it is $MA(2)$) (Figure 25):

The regression line is (Eq. (8)):

$$Bitcoin\_Price = 3120.255075 + 39.973767$$
$$\times \ Litecoin\_Price. \qquad (8)$$

P_value of the ADF-test:

9.4400204e-05.

AIC and BIC of their cointegration:

$$AIC = 4151.424592183152,$$

$$BIC = 4166.728592612673.$$

The third condition assigned to cointegration between Bitcoin and Zcash and ARIMA execution with the order of $p = 1$, $d = 1$, and $q = 2$ (Figure 26):

The regression line is (Eq. (9)):

$$Bitcoin\_Price = 3053.563308 + 20.661674$$
$$\times \ Zcash\_Price. \qquad (9)$$

P_value of the ADF-test:

1.1552e-08.
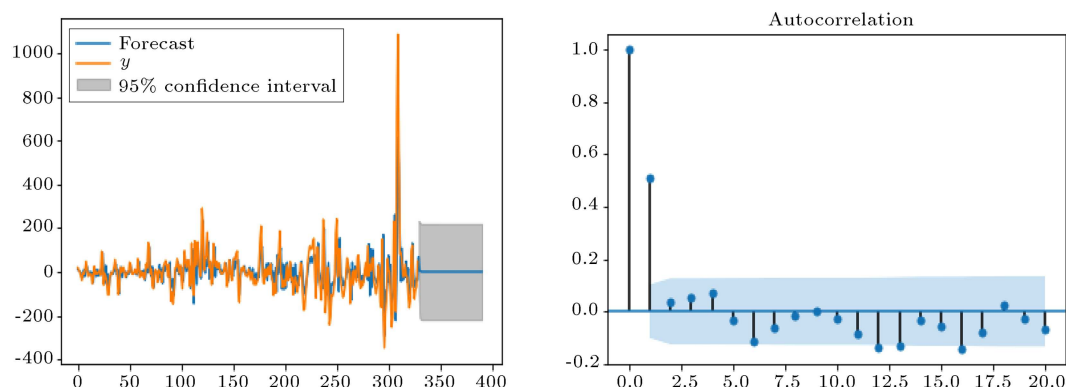
AIC and BIC of their cointegration:

Figure 26. Model execution on residuals Bitcoin-Litecoin (left), and their residuals ACF (right).
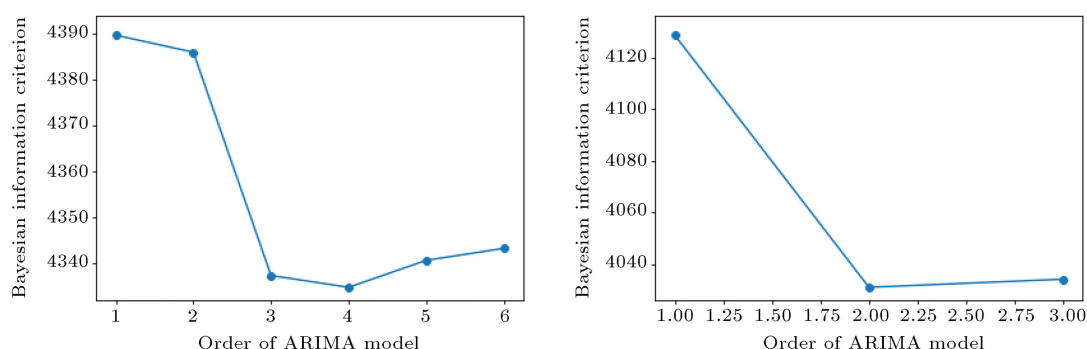

Figure 27. The best ARIMA's models order.

Table 4. Cointegration evaluation criteria.

Cryptocurrencies cointegration	AIC	BIC
Bitcoin-Ethereum	4087.4245922612673	4192.7285924245916
Bitcoin-Litecoin	4151.424592183152	4166.728592612673
Bitcoin-Zcash	3978.591052327704	3997.706281805119

$$AIC = 3978.591052327704,$$

$$BIC = 3997.706281805119.$$

The experimental results show that Zcash price can be the best predictor of the Bitcoin price, as it has the minimum value in AIC and BIC (shown in Table 4).

5. Sensitivity analysis

There are many different ways to determine the best values of models' parameters. Sensitivity analysis in time series models refers to changing models' order. For instance, we can change the number of periods that give the trend's smooth shape in MA models.

To achieve the best order in these models, we developed an algorithm that iterates different rankings and results based on the models' evaluation criteria.

Figure 27 shows related to some of the time series models.

As shown in the regression line estimation by the RFR figure, the evaluation criteria, especially BIC, were calculated using different orders to reach the minimum value. In the left graph, d and q were constant $(= 1)$, but p iterates and obtains the optimal order. The same processes have been done for each parameter. We have repeated these procedures for the whole model throughout the paper.

6. Discussions and conclusions

Nowadays, many people mine and trade different kinds of cryptocurrencies. As their popularity increased daily, their demands grew simultaneously. Price prediction and asset speculation are high-risk ventures. Therefore, the importance of finding an appropriate

way to predict their future values is undeniable. Since cryptos' price are volatile, applying a stable forecasting model is complicated. Thus, we decided to explore some features that can help to predict their price. To this end, we used the price returns to find that these features and address cryptos' volatility. As an open market determines a cryptocurrency's value, this presents unique challenges around volatility that other assets and currencies do not face. While cryptocurrency price prediction is an ever-moving target, market literacy is essential for getting the most value out of their participation in the cryptocurrency market.

In this paper, Bitcoin's cointegration is examined with three other popular coins to determine a currency that has the highest ability to be used as a predictor of Bitcoin price, especially when there is little information about Bitcoin. We forecasted the Bitcoin price based on other coins using machine learning algorithms such as Linear Regression (LR), Gradient Boosting Regressor (GBR), Support Vector Regressor (SVR), and Random Forest Regressor (RFR). Time series analysis models were applied to recognize which coin fits the historical data better than others in the cointegration process.

In the first phase, the relations among cryptocurrencies were investigated fundamentally by machine learning algorithms to ensure they had any meaningful connections. Then, each cryptocurrency's price trend was analyzed and plotted daily for a year to discover how they were acting by different kinds of time series models. After all, a finding of which two cryptocurrencies, Bitcoin as the stable base, had the most similarity, coming to the main contribution. The importance of finding the most similarity among three well-known cryptos showed itself when there is no/rare information on crypto (in this paper, worked on Bitcoin). Using the other crypto can predict what may happen one step forward. The cointegration between cryptocurrencies is suggested for the first time in the literature. According to the paper's goal, the cointegration between Bitcoin and three other cryptocurrencies was studied in four following stages:

1. Bitcoin's price regression on three other cryptos' prices;
2. Execution of Augmented Dickey-Fuller test, ADF-test;
3. Autoregressive Integrated Moving Average (ARIMA) model implementation on their residuals;
4. Criteria measurement's evaluation.

As a result, among the three cryptocurrencies (Ethereum, Zcash, and Litecoin), Zcash had the most trend's similarity to Bitcoin. We also inferred this result by testing the Lasso Regression algorithm as an

ML method. In other words, as the information on Bitcoin price was unavailable, we can use the Zcash price information to forecast the Bitcoin price one step forward.

In future work, we intend to work on more cryptocurrencies in an extended period to boost exactness. Moreover, we plan to apply variate models to the cointegration's estimated equation. Another common way to predict the price of cryptocurrencies is to utilize deep learning and its algorithms, which, due to their complexity, are expected to predict prices more accurately.

Nomenclature

X_t	Model variable in time t
c	Constant number of the model
φ_i	Parameter of the model
ε_t	White noise error
μ	Mean of the series
θ_i	Parameter of the model
p	The order of the AR model
d	The level of subtracting of ARIMA model
q	The order of the MA model
ϕ_i	Parameter of the model
L	Lag operator for period i
y_i	Real value of independent variable
$\widehat{y_i}$	Estimated value of independent variable
T	Number of periods
p	Number of estimated parameters
e_t	Residuals error
n	Number of observations

Superscript

i	Lag index

Subscript

t	Time index

References

1. Bhattacharya, D., Mukhoti, J., and Konar, A. "Learning regularity in an economic time-series for structure prediction", *Applied Soft Computing Journal*, **76**, pp. 31–44 (2019).

2. Corbet, S., Lucey, B., Urquhart, A., et al. "Cryptocurrencies as a financial asset: A systematic analysis", *International Review of Financial Analysis*, **62**, pp. 182–199 (2019).

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

4. Fryzlewicz, P., Bellegem, S., and Sachs, R. "Forecasting non-stationary time series by wavelet process modelling", *Annals of the Institute of Statistical Mathematics*, **55**, pp. 737–764 (2003).

5. Hossain, Z., Rahman, A., Hossain, et al. "Over-differencing and forecasting with non-stationary time series data", *Dhaka University Journal of Science*, **67**, pp. 21–26 (2019).

6. Arlinghaus, S.L., *PHB Practical Handbook of Curve Fitting*, In CRC Press (1994).

7. Dooley, G. and Lenihan, H. "An assessment of time series methods in metal price forecasting", *Resources Policy*, **30**(3), pp. 208–217 (2005).

8. Ediger, V. and Akar, S. "ARIMA forecasting of primary energy demand by fuel in Turkey", *Energy Policy*, **35**(3), pp. 1701–1708 (2007).

9. Khashei, M. and Bijari, M. "An artificial neural network (p, d, q) model for timeseries forecasting", *Expert Systems with Applications*, **37**(1), pp. 479–489 (2010).

10. Khashei, M. and Bijari, M. "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting", *Applied Soft Computing*, **11**(2), pp. 2664–2675 (2011).

11. Martiínez, F., Friías, M., Peérez-Godoy, M., et al. "Dealing with seasonality by narrowing the training set in time series forecasting with kNN", *Expert Systems with Applications*, **103**, pp. 38–48 (2018).

12. Azadeh, A., Moghaddam, M., Khakzad, M. et al. "A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting", *Computers & Industrial Engineering*, **62**(2), pp. 421–430 (2012).

13. Khandelwal, I., Adhikari, R., and Verma, G. "Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition", *Procedia Computer Science*, **48**(1), pp. 173–179 (2015).

14. Buncic, D. and Moretto, C. "Forecasting copper prices with dynamic averaging and selection models", *North American Journal of Economics and Finance*, **33**, pp. 1–38 (2015).

15. Laboissiere, L., Fernandes, R., and Lage, G. "Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks", *Applied Soft Computing Journal*, **35**, pp. 66–74 (2015).

16. Fan, X., Wang, L., and Li, S. "Predicting chaotic coal prices using a multi-layer perceptron network model", *Resources Policy*, **50**, pp. 86–92 (2016).

17. Gangopadhyay, K., Jangir, A., and Sensarma, R. "Forecasting the price of gold: An error correction approach", *IIMB Management Review*, **28**(1), pp. 6–12 (2016).

18. Pradeepkumar, D. and Ravi, V. "Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network", *Applied Soft Computing*, **58**, pp. 35–52 (2017).

19. Fischer, T. and Krauss, C. "Deep learning with long short-term memory networks for financial market predictions", *European Journal of Operational Research*, **270**(2), pp. 654–669 (2018).

20. Sezer, O.B. and Ozbayoglu, A.M. "Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach", *Applied Soft Computing*, **70**, pp. 525–538 (2018).

21. Demir, D., Gozgor, G., Lau, C.K.M., et al. "Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation", *Finance Research Letters*, **26**, pp. 145–149 (2018).

22. Chou, J.S. and Tran, D.S. "Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders", *Energy*, **165**, pp. 709–726 (2018).

23. Kim, H.Y. and Won, C.H. "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models", *Expert Systems with Applications*, **103**, pp. 25–37 (2018).

24. Phillips, P., Shi, S., and Yu, J. "Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500", *International Economic Review*, **56**(4), pp. 1043–1078 (2013).

25. Bouri, E., Shahzad, J., and Roubaud, D. "Co-Explosivity in the cryptocurrency market", *Finance Research Letters*, **29**, pp. 178–183 (2018).

26. Bouri, E., Gupta, R., and Roubaud, D. "Herding behaviour in cryptocurrencies", *Finance Research Letters*, **29**, pp. 216–221 (2018).

27. Ji, Q., Bouri, E., Lau, C.K., et al. "Dynamic connectedness and integration in cryptocurrency markets", *International Review of Financial Analysis*, **63**, pp. 257–272 (2018).

28. Mallqui, D.C. and Fernandes, R.A. "Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques", *Applied Soft Computing*, **75**, pp. 596–606 (2019).

29. Alameer, Z., Elaziz, M., Ewees, A., et al. "Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm", *Resources Policy*, **61**, pp. 250–260 (2019).

30. Cao, J., Li, Z., and Li, J. "Financial time series forecasting model based on CEEMDAN and LSTM", *Physica A: Statistical Mechanics and its Applications*, **519**, pp. 127–139 (2019).

31. Atsalakis, G.S., Atsalaki, I.G., Pasiouras, F., et al. "Bitcoin price forecasting with neuro-fuzzy techniques", *European Journal of Operational Research*, **276**, pp. 770–780 (2019).

32. Bouri, E., Roubaud, D., and Shahzad, J. "Do Bitcoin and other cryptocurrencies jump together?", *The Quarterly Review of Economics and Finance*, **76**, pp. 396–409 (2019).

33. Bouri, E., Lucey, B., and Roubaud, D. "The volatility surprise of leading cryptocurrencies: Transitory and permanent linkages", *Finance Research Letters*, **33**, 101188 (2020).

34. Qureshi, S., Aftab, M., Bouri, E., et al. "Dynamic interdependence of cryptocurrency markets: An analysis across time and frequency", *Physica A: Statistical Mechanics and its Applications*, **559**, 125077 (2020).

Biographies

Negar Maleki is a PhD student in Big Data Analytics program at Muma College of Business in University of South Florida, United States of America. She has her master's degree in Industrial Engineering from the University of Tehran. Her main research interests are Machine Learning, Data Analysis, Statistical Inferences, Optimization, and Biomedical Data.

Alireza Nikoubin has a master's degree in industrial engineering and was a research and teaching assistant at the University of Tehran. His main research interests are in financial markets, heuristic and meta-heuristic optimization algorithms, supply chain management, and data analysis.

Masoud Rabbani is Professor of Industrial Engineering in the School of Industrial and Systems Engineering at the University of Tehran. He has published more than 250 papers in international journals. His research interests include production planning, Maintenance planning, Mixed model assembly lines planning, inventory management systems, and applied graph theory.

Yasser Zeinali is a PhD student at Alberta School of Business, University of Alberta, Canada. His research interests are data analytics, machine learning algorithms and its applications in business and healthcare.