Bitcoin price forecasting by applying combination of stacking method and Differential Evolution Algorithm

Mohammad Hossein Asgari
Department of Financial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran,
Eng.m.h.asgari@gmail.com

Mehdi Madanchi Zaj
Department of Financial Management, Electronic Branch, Islamic Azad University, Tehran, Iran, madanchi@iauec.ac.ir.
(Corresponding author)

Amir Daneshvar
Department of Information Technology Management, Electronic Branch, Islamic Azad University, Tehran, Iran,
daneshvar.amir@gmail.com

Davood Manzoor
Department of Economics, Imam Sadegh University, Tehran, Iran, manzoor@isu.ac.ir.

Submit: 03/07/2022 **Accept:** 25/10/2022

ABSTRACT

Due to the high market share and the leadership of Bitcoin's price in the cryptocurrency market, this crypto asset has always been the leader of the trading trends of digital currencies and its movements have a significant impact on the price trends of other cryptocurrencies. In this article, by using a combination of meta-heuristic and machine learning methods, a model with the least error for estimating the price of Bitcoin is presented. This research uses the stacking approach of popular machine learning algorithms along with optimizing the parameters of the algorithms. In this study, for the first time in the field of cryptocurrencies, the differential evolution method has been used to find the most optimal stack combination and also to identify the most suitable parameters of each algorithm. This research in terms of type, data collection method and purpose is descriptive, modeling, and applied ,respectively.

In this study, the model results have been compared in three scenarios with inputs including OHLC features, technical indicators (more than 160 indicators), fundamental analysis indices (7 indicators) and their combination. Considering that most articles have applied technical indicators as input to machine learning models, the use of fundamental indices is one of the distinguishing features of this research. The results show that depending on the features used in each scenario, the type and order of placement of learning algorithms in the price prediction stack changes. All three scenarios investigated in this research have acceptable accuracy for price prediction. Daily Bitcoin prices in US Dollar during 2019/1/1 to 2021/9/29 are used in this study.

Keywords: Bitcoin Price forecasting, Cryptocurrency, Differential Evolution Algorithm (DEA), Fundamental analysis, Machine learning,Technical indicator.

1. Introduction

The financial crisis of 2008 led to a kind of pessimism about traditional financial systems. On October 31, 2008, an individual (or group of individuals) operating under the pseudonym Satoshi Nakamoto released the Bitcoin white paper and described this decentralized platform as:" A pure peer-to peer version of electronic cash that can be sent online for payment from one party to another without going through a counterparty, ie. a financial institution" (Nakamoto, 2009). This innovative concept with limited 21 million digital coin supply challenged the dominance of governments and central banks over the circulation and control of common currencies.

Bitcoin introduced the world to a new concept called cryptocurrencies or digital currency. Cryptocurrencies are decentralized exchange platforms which utilize cryptographic functions to conduct financial transactions. Cryptocurrencies apply innovative blockchain technology (a digital ledger of economic transactions) to achieve decentralization, transparency and immutability (Fang, et al, 2021)

In recent years, the widespread application of cryptocurrencies in various fields has accelerated, and huge investments have been made in blockchain projects and the creation of new cryptocurrencies.

On December 20, 2019, there were 4,950 cryptocurrencies and 20,325 cryptocurrency markets with a market value of approximately $ 190 billion. This number as of April 20, 2021 (14 months later), includes 9,342 cryptocurrencies, 367 exchanges and a market value of $ 1978 billion (a figure of more than ten percent of the global gold market value). Figure 1 shows the historical data of the total market value of cryptocurrencies in US dollars from April 20, 2013 to April 20, 2021. The total market value of cryptocurrencies is obtained by aggregating the market value (the product of the price multiplied by the number of tokens per cryptocurrency) of each cryptocurrency.

Interest in cryptocurrencies as a new financial asset has grown exponentially in recent years. Cryptocurrencies owe much of their fortunes to the release of news and reports of their unexpected returns, which are somewhat reminiscent of the period of the rush of people to discover gold in the United States (gold rush) over the centuries.

Due to the sharp fluctuations in the price of cryptocurrencies and the failure of many projects to create digital tokens and coins, as well as the exposure of cryptocurrencies to price bubbles, regulatory challenges and cybercrime, the need to predict price trends and returns of cryptocurrencies have considerable importance. Most cryptocurrency researches have focused mainly on examining the role of cryptocurrencies (especially Bitcoin) as a diversified portfolio asset (consisting of all asset classes), a risk hedging tool, or a safe haven. Therefore, less attention has been paid to the price forecasting of cryptocurrencies.

Figure 1: Historical data of total market value of cryptocurrencies in US dollars from April 20, 2013 to April 20, 2021 (Coinmarketcap.com).

Trading prices in different cryptocurrencies depend highly on the market value of the largest available cryptocurrency, Bitcoin, and their prices are somehow affected by the price of Bitcoin. In other words, in the short run, the price behavior of Bitcoin plays a major role in shaping the price of other cryptocurrencies. In particular, there is a strong correlation between Bitcoin and Ethereum (which has the highest market value after Bitcoin). Therefore, Bitcoin price forecasting assists to estimate the price of other cryptocurrencies and evaluate the general crypto market trend. The main purpose of this research is to find a model for predicting the price of Bitcoin with the lowest error rate using meta-heuristic methods.

Prior to price forecasting of cryptocurrencies, it is first necessary to identify and examine the factors and variables affecting their price. In general, several factors affect the pricing of cryptocurrencies. These factors include economic and non-economic parameters. Economic factors include stock indices in the world's major stock exchanges, gold prices, commodities, interest rates announced by central banks in the world's largest economies and other factors. Non-economic factors reflect the perceptions and interpretations of community-accepted economic agents of Bitcoin and cryptocurrencies as financial assets. Non-economic factors include, for instance, the number of tweets posted on Twitter, the uncertainty of economic policies, the global risks of economic and political geography, and the unusual and unpredictable negative events known as the "black swans." Research findings show that non-economic factors have a more active and more influential effect on the price formation of cryptocurrencies compared to economic factors (Ausan, et al., 2019).

It has been found that the most influential factor in price reduction on cryptocurrencies is the direct order to ban or restrict cryptocurrency transactions by the world's central banks, as well as published news about the possible use of cryptocurrencies for money laundering or terrorist financing purposes. In contrast, news on the development of special legal regimes and the initial public offering of tokens have a positive effect on the profitability of cryptocurrencies (Auer and Claessens, 2018).

On the other hand, given the Bitcoin market share of about 50%, in the presence of limited investment opportunities, Bitcoin has always been the first choice of investors in the cryptocurrency market. Therefore, due to the high market share and leadership of Bitcoin prices in the cryptocurrency market, as well as the correlation of the price of most cryptocurrencies with the Bitcoin price trend, the price of bitcoin alone has a noticeable impact on the price of other cryptocurrencies.

In addition to the economic and non-economic factors mentioned above, the price changes of Bitcoin also depend on some internal factors, and that is the lower price range, which is equivalent to the final costs of Bitcoin mining. This factor refutes the claim that Bitcoin has no fundamental value (Hayes ,2018).

The high price fluctuations of cryptocurrencies are largely due to the publication of news and posts on social networks. This effect is intensified by the efforts of investors to verify the accuracy of the published information. Twitter plays a special role in this field. Not only does provide Twitter live updates on developments and events in the field of cryptocurrencies, but also provides a rich source of measuring the emotional intelligence of traders by expressing the feelings and tendencies of investors from the context of this application. Behavioral finance implies that emotions can have a profound effect on individual behavior and decisions. The cryptocurrency market players show a strong herd behavior which exacerbates the volatility of this market (Kraaijeveld and De Smedt, 2020).

Generally, research results show that the cryptocurrency market is significantly independent of other financial markets. The most drastic impact of traditional financial asset prices on Bitcoin is seen during the fall of these markets, among which, commodity and gold markets, have the most impact.

The results of researches on the impact of the US Federal Reserve and the People's Bank of China monetary policy on the price dynamics of the four major cryptocurrencies (including Bitcoin, Ethereum, Litecoin and Ripple) are interesting. This study shows that only the contractionary policy of the People's Bank of China has had a positive impact on the cryptocurrency market. However, the monetary policies of the central banks of the United Kingdom, Japan and the European Union do not cause a significant reaction in the cryptocurrency market (Vladimirovnan et al., 2021).

At present, the cryptocurrency market, despite its high volatility, does not have a significant impact on the stability of global financial markets. This is mainly due to the relatively shallow depth of this market compared to the market of other financial assets and the use of low leverage in cryptocurrency transactions (only about 20% of cryptocurrency holders use leverage in transactions). The cryptocurrency market is also reacting strongly to shocks that undermine global financial stability. The cryptocurrency market, for example, reacted strongly to the global crisis of the corona virus.

The purpose of this study is to obtain a model that with available inputs including general price characteristics and technical and fundamental indicators to provide an estimator with the least error to predict the price of Bitcoin as the most important crypto asset. Considering that in most articles, technical indicators are mostly used as input to machine learning models, using fundamental indices and combining them with technical indicators is one of the distinguishing features of this research compared to other similar studies. Also, in terms of the number of technical analysis indicators used (more than 160 indicators), this study is different from similar studies. According to the done studies so far, the stacking technique and Differential Evolution Method (DEM) have not been used to optimize the problems related to the price of cryptocurrencies and this study is the first research conducted using this method.

In this section, the most important factors affecting the price of cryptocurrencies were examined. In the following sections, while reviewing the most prominent methods of predicting the price of cryptocurrencies, the proposed model is presented. The results of the research are also presented in the final part of this article.

2.Theoretical Foundations and Literature Review

2.1 Predictive models of market conditions of cryptocurrencies

In financial markets, there are mainly two types of analysis to predict market conditions and determine the strategy of investors, which are technical analysis and fundamental analysis. Both methods are similar in that they rely on measurable information that can be tested on historical data to verify their performance (Sadeghi et al, 2020). In addition to the two technical and fundamental strategies, there are new strategies called emerging trading strategies. Emerging trading strategies for cryptocurrencies include econometric-based strategies and machine learning technologies (Fang et al., 2021). These methods or their combinations can be used to predict the price and determine the trading strategy in the cryptocurrency market. Econometric methods use a combination of statistical and economic theories to estimate economic variables and predict their values. Statistical models apply arithmetic equations to encode information extracted from data.

In cryptocurrency trading studies using econometric methods, researchers have used statistical models based on time series data such as Generalized autoregressive conditional heteroscedasticity (GARCH) and BEKK (derived from Baba, Engle, Kraft and Kroner) to evaluate the fluctuations of cryptocurrencies. The most common linear statistical model used in time series analysis is the Auto Regressive Moving Average (ARMA) model (Fang et.al, 2021).

Machine learning is an effective tool for developing trading strategies and predicting the price of Bitcoin and other cryptocurrencies, as this technology can infer data communications that are often not directly visible to humans. Machine learning is the scientific study of algorithms and statistical models used by computer systems that utilize patterns and inferences to perform tasks instead of explicit instructions. Machine learning is a science that allows computers to learn a particular subject without the need for an explicit program. As a subset of artificial intelligence, machine learning algorithms create a mathematical model based on sample data or training data for predicting or making decisions without explicit planning. Machine learning is used in many fields, including engineering, business, financial markets and medicine. Google search engines, for instance, apply machine learning because their machine learning software understands how to rank a web page. The rapid development of machine learning in recent years has increased their application in cryptocurrency transactions, especially in predicting the price and return of cryptocurrencies. In this section, after getting acquainted with the concepts and principles of machine learning methods, some of the most prominent studies conducted in the field of cryptocurrency trading are reviewed.

2.1.1 Market conditions forecasting strategies based on machine learning technology

Generally speaking, machine learning relies on the definition of two fundamental components: input features and objective function. Defining input features (data sources) is where the knowledge of fundamental and technical analysis comes into play. Inputs can be divided into several feature groups. To exemplify, data based on economic indicators (such as GDP, interest rates, etc.), social indicators (Google trends, Twitter, etc.), technical indicators (price, volume, etc.) and other seasonal indicators (time of day, day of the week, etc.). The objective function defines the fit criterion that is used to judge whether the machine learning model has learned its task or not. The process of using machine learning techniques to predict the price / return of cryptocurrencies is shown in Figure 2.

Depending on the formulation of the main learning loop, machine learning methods can be divided into three categories: supervised learning, unsupervised learning, and reinforcement learning. In recent years, several machine learning techniques have been used in cryptocurrency transactions. These technologies are divided into classification, clustering, regression, reinforcement learning and deep learning algorithms according to the target set. (Fang et al., 2021). Each of these algorithms is briefly described below.

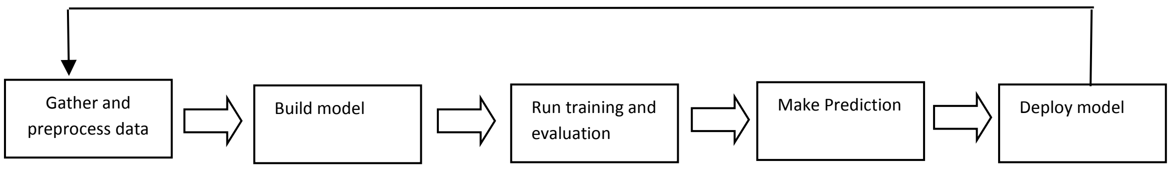

Figure 2: Process of machine learning in predicting cryptocurrency (Fang et.al, 2021).

2.2 Literature Review

In this section, some of the most significant researches on the application of machine learning models in the field of cryptocurrency transactions are mentioned. Typically, technical indicators are used as input features in the field of machine learning trading signals. (Nakano et al., 2018) examined daily technical Bitcoin transactions based on Artificial Neural Networks (ANN) to predict returns. In that study, information on the price and volume of 15-minute Bitcoin transactions in a cryptocurrency exchange was used. An ANN forecasts the price trends (up and down) in the upcoming period from the input data. Data is preprocessed to build a training dataset that includes a matrix of technical patterns containing Exponential Moving Average (EMA), Emerging Markets Small Cap (EMSD), relative strength index (RSI), etc.

The results have proved that the utilization of assorted technical indicators probably prevents over-fitting in the classification of non-stationary financial time-series data, which improves trading performance in comparision with the initial technical trading strategy. (Buy-and-Hold is the benchmark strategy in that study).

Some machine learning’s classification and regression models have been used to predict price trends in cryptocurrency transactions. Most scholars have concentrated on comparing different methods of machine learning’s classification and regression. (Sun et al., 2019) applied random forests (RFs) with factors in Alpha01 (capturing features from the history of the cryptocurrency market) to construct a forecasting model. The study collected data from API in cryptocurrency exchanges and chose 5-minute frequency data for backtesting. The results of the study display that the performance is proportional to the number of data (more data, more accuracy) and it seems that the factors used in the RF model are more important. (Vo et al., 2018) used RFs in High-Frequency cryptocurrency Trading (HFT) and compared it with deep learning models. Different time periods and RF trees have been tested in this study. The results demonstrated that RF is effective despite multicollinearity occurring in ML features, the lack of model identification also potentially leading to model identification issues. That study also tried to create an HFT strategy for Bitcoin using RF. (Maryna et al., 2018) investigated the profitability of an algorithmic trading strategy based on training an SVM model to identify cryptocurrencies with high or low predicted returns. The results of that research showed that the performance of SVM strategy is ranked fourth among the other 4 trading strategies. The authors of that paper observed that SVM requires a large number of parameters and is very prone to overfitting, which leads to poor performance of this model. (Barnwal et al., 2019) applied generative and discriminative classifiers to create a stacking model, specially 3 generative and 6 discriminative classifiers combined by a one-layer Neural Network, to forecast the direction of cryptocurrency price. A discriminative classifier directly models the relationship between unknown and known data, while generative classifiers model the prediction indirectly through the data generation distribution. In that study, technical indicators including trend line, momentum, volume and volatility were selected as model features. Finally, the authors of the paper discuss how different classifiers and features affect forecasting. (Attanasio et al., 2019) compared a variety of classification algorithms including SVM, NB and RF in forecasting next-day price trends of a given cryptocurrency. The results displayed that owing to the heterogeneity and volatility of cryptocurrencies' financial instruments, predicting models based on a series of forecasts seemed better than a single classification technology in trading cryptocurrencies. (Madan et al., 2015) modeled the Bitcoin price prediction problem into a binomial classification. In this study, they used daily, 10 minute and 10 second data. The results of this experiment showed that 10-minute data had better sensitivity and specificityratio compared to 10-second data (10-second prediction had an accuracy of about 10%). Similarly, (Virk, 207) compared RF, SVM, GB and LR to forecast the Bitcoin price. The results of this research showed that SVM with 62.31% accuracy had the highest accuracy among binomial classification machine learning algorithms.

Various deep learning models have also been utilized to find patterns of price movements in the cryptocurrency market. (Zhengy et al., 2019) implemented two machine learning models, fully-connected ANN and LSTM to predict cryptocur rency price dynamics. The results of this study showed that ANN performs better than LSTM, although in theory, LSTM is more appropriate than ANN in terms of time series modeling dynamics. The results demonstrates that the future state of a time series for cryptocurrencies highly depends on its historic evolution. (Kwon et al., 2019) applied an LSTM model, with a three-dimensional price tensor representing the past price variations of cryptocurrencies as input. This study displayed that LSTM is more suitable for classifying crypto asset data with high volatility. (Alessandretti et al., 2018) examined Gradient boosting decision trees (including single regression and XGBoost-augmented regression) and the LSTM model on predicting daily cryptocurrency prices. They showed that when predictions occur over time periods of 5 to 10 days, gradient-based methods perform better, while LSTM outperforms for predicting time periods based on 50 days of data. In that study, the relative importance of features in both models was compared and an optimized portfolio composition (based on geometric mean returns and Sharp ratios) was discussed. (Phaladisailoed et al., 2018) selected regression models and deep learning-based models to compare the performance predictions of the rise and fall of Bitcoin price.

Scholars have also conducted studies to compare classical statistical models and deep/machine learning models. (Rane et al., 2019) described classical time series forecast methods and machine learning algorithms used for Bitcoin price prediction. Statistical models such as Autoregressive Integrated Moving Average models (ARIMA), Binomial Generalized Linear Model and GARCH are compared with machine learning models such as SVM, LSTM and Non-linear Auto-Regressive with Exogenous Input Model (NARX). The findings of this study demonstrated that the NARX model had the best performance with close to 52% prediction accuracy based on 10-second intervals.

3 Methodology

This research in terms of type, method (data collection method) and purpose is descriptive, modeling, and applied ,respectively. Based on the nature of the data, this research is quantitative. The framework of the process in this study is shown in Figure 3. The conventional data of Bitcoin such as (Opening price, Highest price, Lowest price and Close price) OHLC, technical indicators as well as fundamental analysis indicators are used as input features to build an estimator model with the least error for Bitcoin in this study. The effectiveness of each of these datasets has been examined in three scenarios separately and the results have been compared with each other.

To build the model in this research, the stacking approach of the proposed algorithms along with the optimization of the algorithm parameters is used. In the study, differential evolution method has been used to stack regression algorithms (to find the optimal stack composition) as well as to identify the most appropriate parameters of each of these algorithms. According to the studies conducted so far, the method of differential evolution has not been used to optimize the issues related to the price of cryptocurrencies and this study is the first research conducted using this method. In this research, the daily trading data (price) of Bitcoin in US dollars for the period from January 1, 2019 to September 29, 2021 has been used. This period includes a complete cycle of recession and prosperity in the cryptocurrency market (ups and downs of prices in the cryptocurrency market). The data was extracted from “coinmarketcap.com” as one of the most reliable cryptocurrency trading information sites. Information about Bitcoin network data and other technical variables, including fundamental variables, has also been extracted from “coinmetrics.com”. To analyze the data in the study, Python software and other data mining software and statistical software have been used accordingly. One of the activities performed to prepare the data was to extract technical features from the data. For this purpose, libraries implemented in the Python programming language have been used. In this paper, more than 160 technical indicators have been used as input in the proposed model, which is more than 124 technical indicators used in the similar article in 2019 (Huang et al., 2019). In this study, the fundamental analysis features of Bitcoin have been used, which is the distinguishing point of this research compared to similar studies. Combination of fundamental and technical indicators are also applied in this study as input features of the proposed model. In the next section, the technical and fundamental features used in this research are briefly introduced.

3.1 Features based on technical analysis

Technical analysis is the study of market behaviors using charts with the aim of predicting the future of price trends. The cornerstone of modern technical analysis is attributed to the writings of Charles Dow in the late nineteenth century. Dow's three famous theories, on which the foundations of technical analysis are based, are:

- Everything is included in the price.
- Prices move according to trends.
- History repeats itself (Murphy, 1999)

In summary, technical analysis is the art of studying the history of an asset's price movements to predict the future price of that asset. The reason that technical analysis often works depends on several factors, including the following two:

Investor Behavior: Behavioral finance researches show that investors make their decisions based on certain psychological biases that repeat themselves.

Collective Psychology: Many market players use similar technical analysis methods, therefore, resistance and support price levels are strengthened and validated (Kiana, 2019).

Technical indicators are functions that are designed with special calculations, and by comparing the variations of these functions with the price chart, special analytical results can be obtained that lead to predicting the next movement of prices. Indicators are divided into different groups. Researchers are constantly designing new indicators and regularly introduce new indicators. The three most important types of indicators include oscillators, trend indicators, and volume indicators (Murphy, 1999).

In most papers, indicators and oscillators are used to examine trading strategies and determine suitable points for entering and exiting the desired cryptocurrencies. In this section, some of the most well-known and widely used indicators used in recent years in the field of cryptocurrencies are mentioned (Sadeghi et.al, 2019).

In the paper, more than 160 technical indicators and oscillators available in the TA-Lib library have been used as input to the proposed model (Benediktsson et al., 2022). A list of applied indicators in this study is provided in Appendix 1. It should be noted that (Huang et al., 2019) used 124 technical indicators to predict the price of Bitcoin.

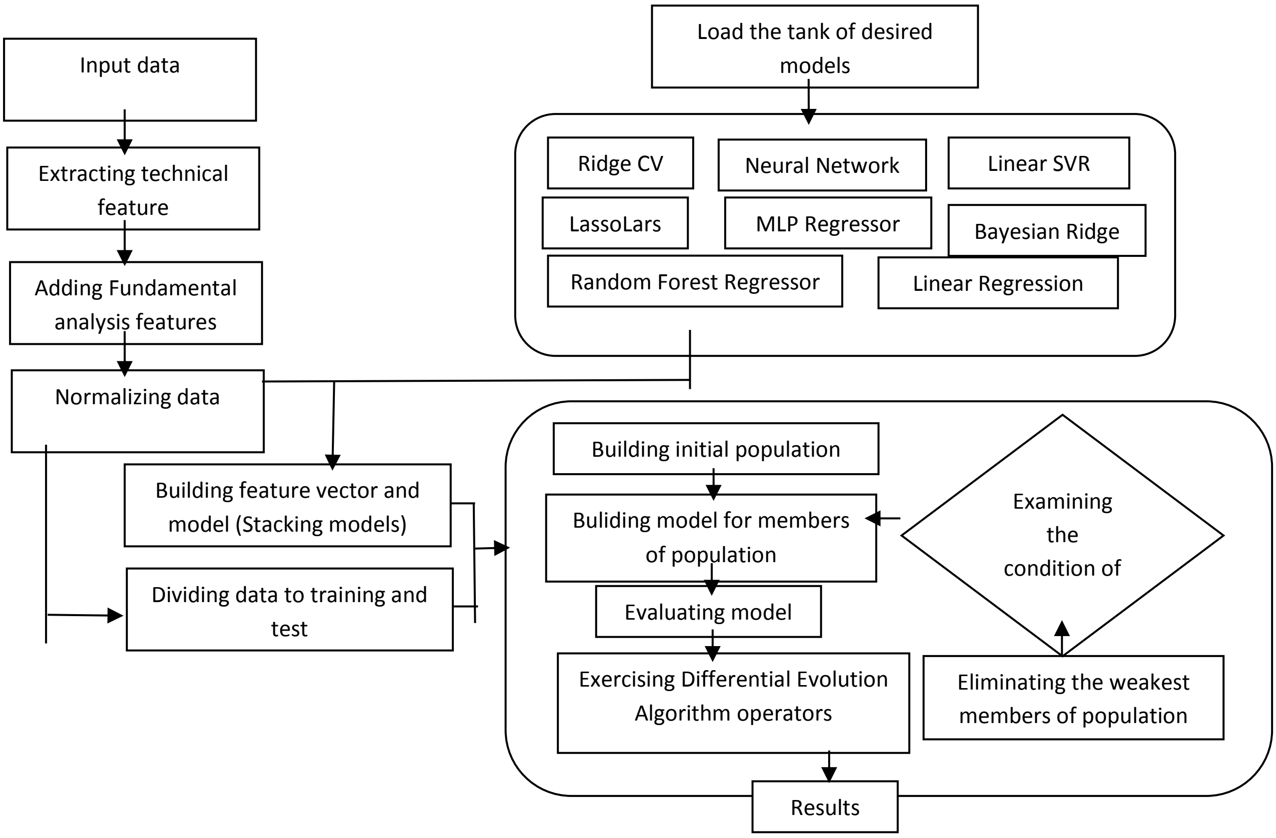

Figure 3: Proposed general framework for Bitcoin price forecasting in this study.

Table 1: Some of the most well-known technical indicators used in cryptocurrency trading studies in recent years.

Name	Formulation	Type
RSI	$100 \times (1 - \frac{1}{RS})$	Oscillator
PD_RSI	$\frac{RSI_t - EMARSI_t}{EMARSI_t}$	Oscillator
PD_EMA	$\frac{Current\ Price - EMA}{EMA}\ \%$	Indicator (trend)
MACD	EMA(Close,12)-EMA(Close,26)	Oscillator
STOC	K%=(Close-Min(Low(K%))/Max(High(K%))-Min(LOW(K%)))*100 %D = SMA (%K, N)	Oscillator

3.2 Features based on fundamental analysis

Fundamental analysis requires examining the drivers of the intrinsic value of an asset. For instance, in the case of stocks, fundamental analysis involves assessing a company's operating health by carefully examining a company's balance sheet, profit and loss statement, and cash flow statement, taking into account the company's long-term macroeconomic outlook. Indicators such as price-to-earnings, price-to-sales, book value, and stock returns, which are used to determine a company's value and compare it to similar companies, are derived from fundamental analysis.

One of the most common questions about cryptocurrencies is what gives them value, given that they have no physical manifestation? Because cryptocurrencies are born of computer software, their value stems from the community and marketplace around which cryptocurrencies are formed. In general, two value components for cryptocurrencies are defined as the value of utility and the value of speculation (Burniske, C, Tartar, 2018).

The value of utility refers to the type, quality and differentiation of services provided by the cryptocurrency, which is determined by the supply and demand of utility among users. In the case of Bitcoin, for example, the utility is that the cryptocurrency can safely, quickly and efficiently transfer money from one person to anywhere in the world. All you have to do is type in someone's Bitcoin address and click the submit button. The utility of Bitcoin in transferring money using the Internet is similar to the utility of Skype in fast, high security and efficient transfer of audio and video to people around the world.

The process of conducting fundamental analysis in cryptocurrencies is different from that of corporate stocks, because cryptocurrency assets are not an enterprise or corporation. Corporate assets may be created by a company or group of individuals that it is necessary to understand that company or individuals, while cryptocurrencies themselves should be priced as commodities with markets based on supply and demand balance. Fundamental analysis of cryptocurrencies includes the study and research on the basic features and characteristics of cryptocurrencies. These include the following:

- Cryptocurrency white paper
- The community and developers of the cryptocurrency
- Connections with its digital siblings
- Public offering model.

Moreover, a number of other technical factors are involved in valuing cryptocurrencies in fundamental analysis, including the number of miners, hash rate, corporates' support, and users' acceptance which are measured by the number of users and the number of transactions (Burniske, C, Tartar, 2018). In the following section, a brief description about aforementioned items are presented.

One of the ways to determine the relative security of cryptocurrencies is hash rate. The hash rate of a cryptocurrency indicates the combined computing power of mining computers connected to the network. Hash rate shows the speed of mining performance. The unit of hash rate is hash per second, which indicates the number of times the machine can execute commands in one second. A higher hash rate means that more computers have been added to support the network, which in turn means more security. This usually happens when the value of the cryptocurrency and the transactions associated with it are increasing. Sometimes the price can be a function of the hash rate. This happens when miners have good expectations for the future of cryptocurrencies, so they will actively seek to connect new machines to help secure the network. This conveys approprate signals to the market, which itself leads to an increase in prices.

One of the most basic ratios to understand the fundamental valuation of cryptocurrencies is the ratio of the value of the network to the value of daily transactions (Network Value to Transactions (NVT)). This index is equivalent to the price-to-earnings (P/E) index in the stock market. To calculate this index, instead of the share price, the value of the network, which is equivalent to the market value of that cryptocurrency, is used. Also, instead of the profit per share, the amount of dollars transferred by the currency per day is used. If a cryptocurrency has a high network value (market value) but its network activity is low, the NVT ratio for that cryptocurrency will be a high number. If the cryptocurrency has an average market value and a high network activity value, this index will get a lower number for it. These information for some cryptocurrencies is available on "coinmetrics.io". Empirically, if the value of this index is less than 15, it means the price of crypto asset is undervalued, and if the value of the said index is greater than 25, it means the price is ovevalued.

Another important indicator that is used in the fundamental analysis of cryptocurrencies is the Price to Mining Cost Breakeven (P/BE) Ratio. This index compares the price of a cryptocurrency with its mining cost. This index is only calculated for mineable coins and does not work for tokens or coins that have already been mined. Most of the time, this ratio is calculated and compared to one of the major cryptocurrencies, such as Bitcoin or Ethereum. For instance, if the price of Bitcoin is \$6400 and the mining price is \$7000, this ratio will be 0.91. Empirically, if this number is less than 1.2, it indicates an upward trend in the market, and if it is greater than 3.2, it indicates a downward trend in the market.

In the table below, the list of fundamental variables that are used in this study as input in the proposed prediction model is provided. Used data for Bitcoin in the study are extracted from "Coinmetrics.io" website. Considering that in most papers, technical indicators are mostly used as input to machine learning models, the use of fundamental indices can be a distinguishing aspect of this research compared to other similar studies.

After familiarizing with the used technical and fundamental analysis features in this research, in the next section the details of the proposed method for Bitcoin's price forecasting will be explained.

Table 2: The list of fundamental indices used in this research as input in the proposed prediction model.

Fundamental index	Description
AdrActCnt	The sum count of unique addresses that were active in the network (either as a recipient or originator of a ledger change) that interval. All parties in a ledger change action (recipients and originators) are counted. Individual addresses are not double-counted if previously active.
CapMVRVCur	The ratio of the sum USD value of the current supply to the sum realized USD value of the current supply.
CapMrktCurUSD	The sum USD value of the current supply. Also referred to as network value or market capitalization.
NVTAdj90	The ratio of the network value (or market capitalization, current supply) to the 90-day moving average of the adjusted transfer value. Also referred to as NVT.
ROI30d	The return on investment for the asset assuming a purchase 30 days prior.
TxCntSec	The sum count of transactions divided by the number of seconds that interval.
HashRate30d	The mean rate at which miners are solving hashes over the last 30 days. Hash rate is the speed at which computations are being completed across all miners in the network. The unit of measurement varies depending on the protocol

3.3 Stacking of models

Data mining methods can be used to help researchers to develop the forecasting process. In order to achieve better results, researchers have proposed combining methods to get a more accurate and appropriate model. One of the most effective and innovative aggregation methods is the application of stacking technique, in which ensemble multiple models and helps to increase the accuracy of forecasting models and enhances the output.

Stacking is one of the regular and accurate methods in forecasting modeling. Stacking is a method based on meta-learning that combines multiple classifications or regression models. This approach, also known as Stack Generalization, was first introduced by Wolpert (1992). The classification based on this model consists of two stages where the output of the first level of the classifier is used as an input for the second level. Namely, it can be said that the prediction of the classifier at one level is used as a feature for another classifier (Shabankareh et al, 2021).

3.3.1 Machine learning models used in the study

In this research, several popular machine learning models have been used, which will be described and briefly explained in the following.

MLP Regressor

MLP Regressor is the corresponding class for numeric prediction tasks based on training multilayer perceptrons with one hidden layer. The former has as many output units as there are classes, the latter only one output unit. Both of these methods minimize a penalised squared error with a quadratic penalty on the (non-bias) weights,. both classes apply BFGS optimization by default to find parameters that correspond to a local minimum of the error function. Logistic functions are used as the activation functions for all units regardless of the output unit in MLPRegressor, MLPRegressor also rescales the target attribute using standardization. All network parameters are initialized with small normally distributed random values (Gabralla & Abraham, 2014).

Linear Regression

In the linear regression model, having certain values of the independent variable (Xi), the values of Yi or the dependent variable can be estimated linearly. The general model of the regression function is the following equation:

$$Y_i=B_1+B_2x_i+\ldots+B_jx_i \quad (1)$$

Where Y_i is the estimation of the dependent values, B_1 is the constant coefficient or y-intercept, B_2 is the coefficient of the variable X_i and X_i is the independent variable. The estimation criterion in this model is the adjusted coefficient of determination, which ranges from 0 to 1, and values close to 1 indicate good fit.

Bayesian Ridge Regression (BRR)

BRR utilizes the regularized parameters λ (accuracy of weight) and α (accuracy of noise) to fit the model. BRR is designated based on Bayesian linear regression. BRR solves the problem that Bayesian linear regression is hard to determine the model in Maximum likelihood estimation(MLE) by introducing the penalty parameters and the Bayesian method of ridge regression (Yang and Yang, 2020).

Random Forest Regressor (RFR)

RFR algorithm (Breiman, 2001) is a conceptually simple supervised learning algorithm with extensive application, that consists of an ensemble (or forest) of decision trees. Each tree includes a tree-like sequence of decision nodes, based on which the tree splits into its assorted branches until the end of the tree ("the leaf") is reached. This leaf is the forecasting of the decision tree. Each decision node is based on whether one of the input features is above a certain value. A significant aspect of the RF is that each tree of the forest is trained on a subset of the full training data, thus providing a slightly different approximation of the model. A prediction is then made by averaging the forecasts of the individual trees. Random forest regression is less susceptible to overfitting and the predictions it produces are more stable compared to results from a single decision tree (Breiman, 2001). Random forest is widely used because it is simple to implement, suitable for both regression and classification problems, does not need data transformation, and is less prone to irrelevant or highly correlated input features. Moreover, random forest allows a simple evaluation of the factors controlling the predictions, the decision structure and the relative significance of each input variable. Analyzing these features can provide valuable insight into the factors controlling the underlying mechanisms (Keller & Evans, 2019).

Lasso Lars (LASSO)

The LASSO is a constrained version of ordinary least squares that has been extensively used in regression analysis for large models (Tibshirani, 1996). It minimizes the residual sum of squares while constraining the sum of absolute values of the regression coefficients. Consider the model for phenotypes of some quantitative trait as:

$$\min\left\{\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{m}x_{ij}\beta_j\right)^2\right\} \text{ subject to } \sum_{j=1}^{m}|\beta_j|\leqslant t, \quad \text{for } t\geqslant 0. \tag{2}$$

Where y_i is the phenotype of the i_{th} individual; x_{ij} is the genotype of the i_{th} individual at the j_{th} marker of 1 to m markers, with x_{ij}=0 for genotype 11; x_{ij}=1 for genotype 12 and x_{ij}=2 for genotype 22; b_j is the allelic substitution effect for the j_{th} marker and e_i is the random residual of the ith individual. Then the LASSO solution is the set of b_j that satisfy:

$$\sum_{i=1}^{n}y_i=0,\ \sum_{i=1}^{n}x_{ij}=0 \quad\text{and}\quad \sum_{i=1}^{n}x_{ij}^2=1, \quad \text{for } j=1,2,3,\ldots,m. \tag{3}$$

The constraint *t* allows some estimated SNP effects to be exactly zero (Usai et al, 2009).

Linear Support Vector Regression (SVR)

SVR is based on principles similar to those of Support Vector Machine (SVM). SVM is usually used for classification, while SVR is used for regression and hence, its output variable is continuous in nature. This characteristic makes forecasting more difficult because there are an infinite number of possibilities for the output variable. SVR applies margin of tolerance E and tolerates the error up to this tolerance. The major principle of minimizing error and maximizing margin while tolerating part of the error remains same (Thakkar, 2015).

Differential Evolution Algorithm (DE)

DE algorithm is an optimization algorithm that was first introduced in 1995 by Rainer Stern and Kenneth Price. These scholars showed in an article that this algorithm has a good ability to optimize non-differentiable nonlinear functions, which is introduced as a powerful and fast method for optimization problems in continuous spaces. The DE algorithm is presented to overcome the main disadvantage of the genetic algorithm, i.e. the lack of local search in this algorithm. The main difference between genetic algorithms and DE algorithm is in the selection operator. Like other population-based optimization algorithms, DE also includes two steps: initialization and evolution. In the initial phase, if there is no information about the problem, the DE population is randomly generated. In the evolution phase, the individuals of the population are repeatedly improved through mutation, recombination and selection process until the termination criterion of the algorithm is satisfied. In general, DE is a minimization problem of the function f(x):

$$Min\, f(x),\ X = [x_1, x_2, \ldots\ldots, x_d] \quad (4)$$
$$s.t.\ x_i \in [a_i, b_i]$$

Where f(x) is the objective function and X is the decision vector consisting of D variables. a_i and b_i represent the upper and lower limits of variable x_i, respectively. In the standard DE model, each individual represents a candidate solution for f(x) in the search space. In each generation, the objective function is assessed for each individual and the obtained value is used to evaluate the quality of individuals during the optimization process in order to find the best member (Bo eshagh and Eftekhari, 2014).

Initialization

Individuals in DE are as follows:

$$X_i^g = [x_{i1}^g, x_{i2}^g, \ldots\ldots, x_{iD}^g], i \in \{1, 2, \ldots\ldots, NP\} \quad (5)$$

Where X^g_{ij} represents the candidate parameter for the j_{th} variable, the i_{th} individual in the g_{th} generation for the objective function f(x). The parameter NP is the population size and D is the dimension of the objective function. In the initial phase, all individuals from the entire population are randomly assigned with a uniform probability distribution in the search space.

Mutation

The main difference between DE and genetic algorithm is mutation. There are several methods of mutation in DE. In this section, the simplest method is explained. For each vector X^g_{ij}, DE generates a mutation vector V^i_g, based on the difference between vectors randomly selected from the population.

$$V_i^g = X_{r3}^g + F^k.(X_{r2}^{gk} - X_{r1}^{gk}) \quad (6)$$
$$r_1, r_2, r_3 \in \{1, 2, \ldots\ldots, NP\}, k \in \{1, 2\}, F^k \in [0, 1]$$

Where F is the mutation factor that has a close relationship with the convergence speed and usually takes a value between 0 and 2. K represents the number of differential vectors that participate in the mutation operation. In practice, when there are two difference vectors, usually the value of F1 is equal to F2. The parameters X^g_{r1}, X^g_{r2} and X^g_{r3} are individuals from the population that are randomly selected from the population and are different from the individual X^g_i that is running.

Recombination

In this algorithm, the candidate child vector is recombined according to the following equation.

$$u_{ij}^g = \begin{cases} v_{ij}^g & \mathrm{rand}(j) \le CR \\ x_{ij}^g & \mathrm{rand}(j) > CR \end{cases} \quad j \in \{1,2, \ldots, D\} \quad (7)$$

Where u^g_{ij} is a parameter of the candidate child vector u^g_j and v^g_{ij} is a parameter of the vector v^g_j. CR is the recombination rate bounded between [0, 1].

Selection

Selection has the role of competitive mechanism. Candidate child u^g_j and old person x^g_j compete based on their merit. The winner will have the chance to survive and be passed on to the next generation.

$$x_i^{g+1} = \begin{cases} U_i^g, & \text{if } f(x_i^g) > f(U_i^g) \\ x_i^g, & \text{otherwise} \end{cases} \quad (8)$$

The processes of mutation, recombination and selection are repeated until the termination conditions of the algorithm are satisfied and the output is a final candidate solution for f(x) (Bo eshagh and Eftekhari, 2014).

4. Results

In this section, the results of the implementation of proposed model are presented. The results of model have been investigated in three different scenarios. In the first scenario, merely the results obtained with OHLC input variables have been examined. In the second scenario, the results are examined with OHLC and technical indicators as input variables, and in the third scenario, the results are examined with OHLC, technical indicators and fundamental indices as input variables.

The results of using OHLC input variables are shown in Table 3. “Open” variables are known as the best subsets for building the model in this scenario. The best model built for this scenario consists of a stack of 5 models (Random Forest, Random Forest, Random Forest, Random Forest, and Linear Regression, respectively). The estimator model which using OHLC featurse as input is presented in Figure 4.

It should be noted that regarding tables No. 3, 4 and 5, the parameters used for each of these models are presented in Appendix No. 2. Also, the explanations related to each of the selected features can be seen in Appendix No. 1 (related to technical indicators) and the “coinmetrics.io” website (other indicators). The explanation of parameters can be found in the (www. scikit-learn.org.) website.

Table 3: The results of using OHLC featurse as input to the estimator model.

Selected feature set	Open
The best built model	RandomForestRegressor(Params_1) RandomForestRegressor(Params_2) RandomForestRegressor(Params_3) RandomForestRegressor(Params_4) MLPRegressor(Params_5)
Training time (Seconds)	123.2943141
Mean Square Error	1.4568612229103932

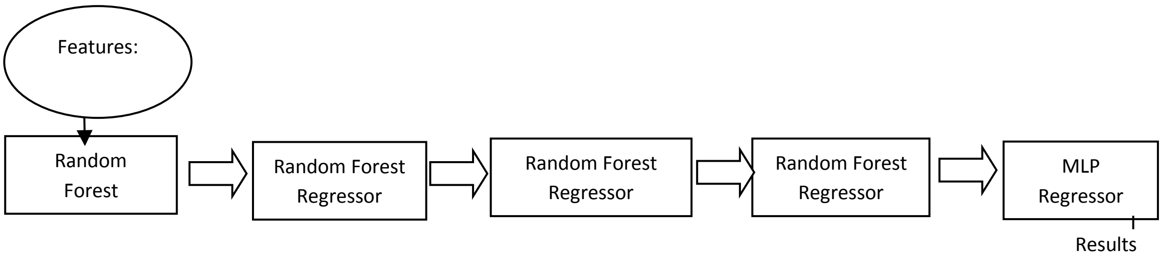

Figure 4: Proposed model for Bitcoin’s price forecasting in the first scenario (using OHLC featurse as input to the estimator model).

Figure 5 shows the chart of Bitcoin price changes related to the training and test data as well as the prediction made by the best introduced model. As can be seen, these graphs have high compliance, which means high accuracy and negligible error of the trained model.

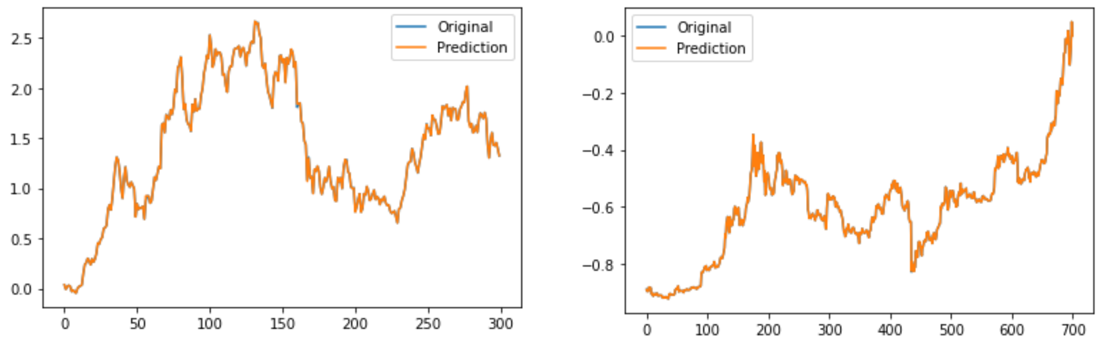

Figure 5: Bitcoin price changes related to the training and prediction data made by the best model introduced in the first scenario (right) and Bitcoin price changes related to the test and prediction data made by the best introduction model done in the first scenario (left).

Due to the fact that the differential evolution method works on both the selection of the most suitable subset of data and the way of stacking different algorithms and optimizing their parameters, it faces a wide range of results and moves towards the most optimal solution. An example of the worst generated solutions is also shown in Figure 6. As can be seen, in case of improper feature selection, improper stacking, and also setting inaccurate parameters for each algorithm, we will face very bad answers.

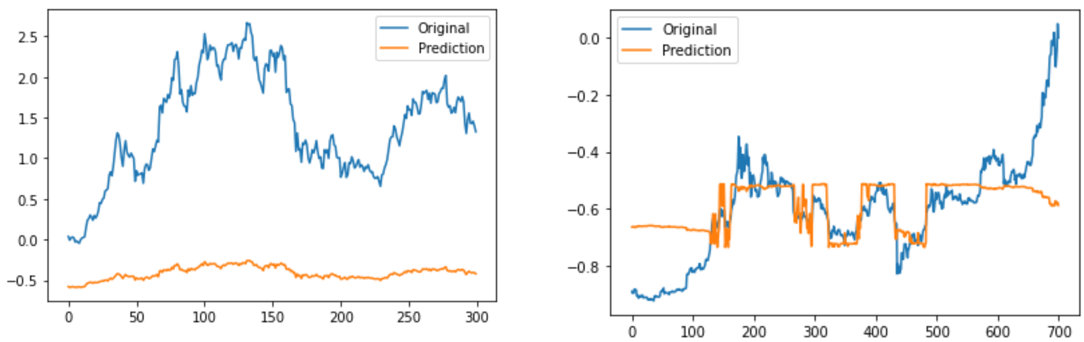

Figure 6: An example of the answer of an inappropriate model on the training data (right) and an example of the answer of an inappropriate model on the test data. (Left side).

The results of using OHLC features and technical indicators as input for estimator model are shown in Table 4. The best model constructed for this scenario consists of a stack of 5 different models as follows (Lasso Lars, Random Forest, Random Forest, Random Forest and Linear Regression, respectively). The estimator model which using OHLC and technical indicators featurse as input is presented in Figure 7. The stacking process in this model is consisted of four stages of Random Forest. The process ends with MLP regressor.

Table 4: The results of using OHLC features and technical indicators as input to the estimator model.

Selected feature set	Close, volatility_bbp, volatility_bbli, volatility_kcl, volatility_kcli, volatility_dcl, volatility_ui, trend_macd, trend_ema_fast, trend_adx, trend_vortex_ind_diff, trend_kst, trend_kst_sig, trend_stc, momentum_rsi, momentum_stoch_rsi, momentum_tsi, momentum_wr, momentum_ao, momentum_kama, momentum_ppo_hist, others_dlr,
The best built model	Lasso Lars (Params_1) Random ForestRegressor(Params_2) Random ForestRegressor(Params_3) Random ForestRegressor(Params_4) Linear Regression (Params_5))
Training time (Seconds)	1991.781
Mean Square Error	1.4477005817242348

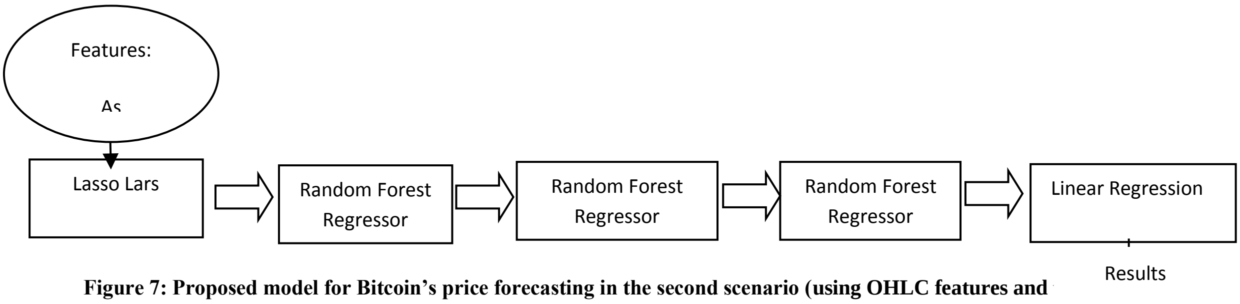

Figure 7: Proposed model for Bitcoin's price forecasting in the second scenario (using OHLC features and indicators as input to the estimator model).

As it can be seen, this estimator model uses 23 features amongst OHLC and technical variables as input and leads to less estimation error in comparison with the first scenario. Although the training time is considerably higher than the first scenario. The stacking process in this model is consisted of three consecutive stages of Random Forest.

Figure 8 shows the low error for predicting Bitcoin price in training and testing data in the case where OHLC variables and technical indicators are used as inputs.

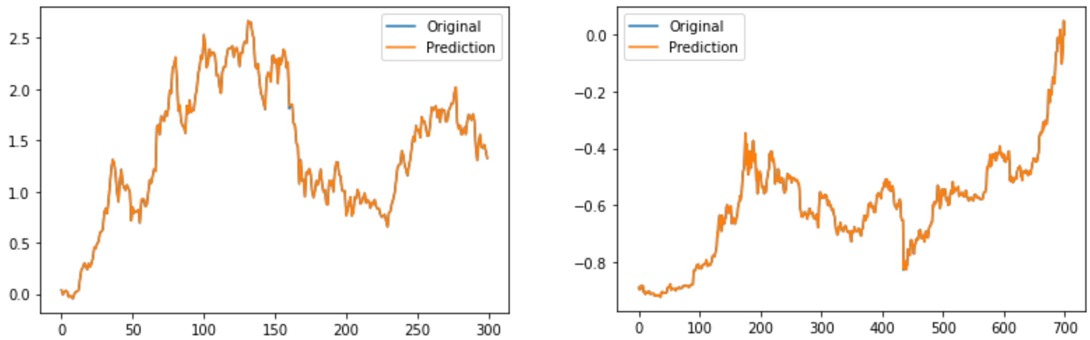

Figure 8: Bitcoin price changes related to training and prediction data made by the best model introduced in the second scenario. (right side) and Bitcoin price changes related to the test and prediction data made by the best method introduced in the second scenario (left side).

The results of using OHLC variables, technical indicators and fundamental indices as inputs of estimator model are shown in Table 5. The best model built for this scenario consists of a stack of 5 different models as follows (Random Forest, Random Forest, Random Forest, Random Forest, and Linear Regression, respectively). The estimator model which using OHLC, technical indicators and fundamental indices featurse as input is presented in Figure 9. Likewise scenario 1, the stacking process in this model is consisted of four stages of Random Forest and process ends with MLP regressor.

Table 5: The results of using OHLC variables, technical indicators and fundamental indices as input to the estimator model.

Selected feature set	volatility_atr, volatility_bbhi, volatility_kch, volatility_kcp, volatility_kcli, volatility_dcp, trend_macd_diff, trend_sma_fast, trend_ema_fast, trend_vortex_ind_pos, trend_trix, trend_visual_ichimoku_a, trend_psar_up_indicator, trend_psar_down_indicator, trend_stc, momentum_rsi, momentum_stoch_rsi, momentum_stoch_signal, momentum_ppo_hist, others_dlr, AdrActCnt, CapMrktCurUSD, ROI1yr
The best built model	RandomForestRegressor(Params_1) RandomForestRegressor(Params_2) RandomForestRegressor(Params_3) RandomForestRegressor(Params_4) MLPRegressor(Params_5)
Training time (Seconds)	1730.719
Mean Square Error	1.4537249534629806

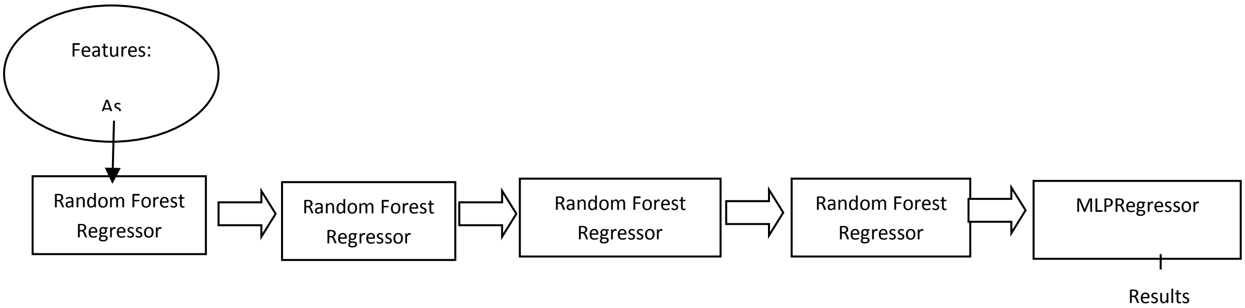

Figure 9: Proposed model for Bitcoin's price forecasting in the third scenario (using OHLC variables, technical in… and fundamental indices as input to the estimator model).

As it can be seen, this estimator model uses 24 features amongst OHLC, technical variables and fundamental analysis indices as input and leads to more estimation error in comparison with the second scenario. The training time in this scenario is less than that of the second scenario. The stacking process in this model is consisted of four consecutive stages of Random Forest.

Figure 10 shows that if OHLC, technical indicators and fundamental indices are used in predicting the price of Bitcoin in the training and test data, the prediction error is very low.

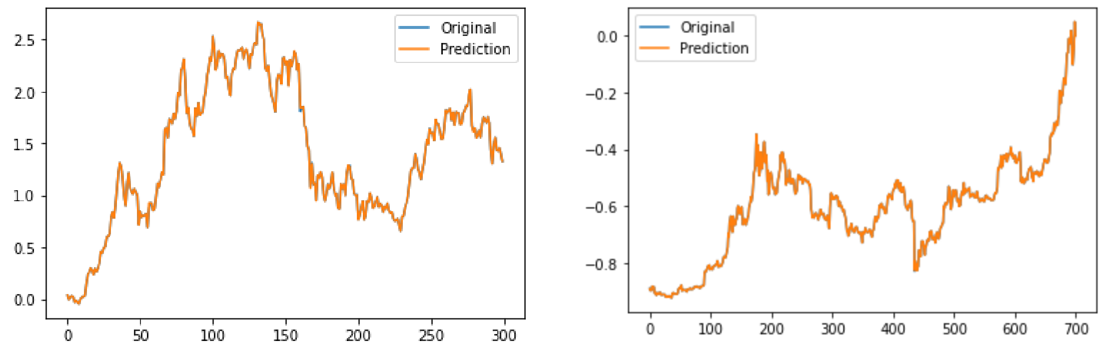

Figure 10: Bitcoin price changes related to training and prediction data made by the best model introduced in the third scenario (right chart) and Bitcoin price changes related to test and prediction data made by the best introduced model in the third scenario (left diagram).

In Table 6, the results related to the amount of training time and mean square error in three scenarios are compared with each other.

As can be seen, all three scenarios examined in this research have acceptable accuracy for predicting the price of Bitcoin. According to the investigations, the lowest prediction error is in the case where OHLC variables and technical indicators are used as input in the prediction model. The lowest training time is related to the model that uses the OHLC feature set. The first scenario just uses a single selected feature for price forecasting while the second and third scenarios utlize 22 and 23 features, respectively. Another noteworthy point is the insignificant effect of adding fundamental indices as input features on the amount of prediction error compared to the basic model using only OHLC variables. Another important point is the greater weight of the feauters of technical analysis indicators to increase the accuracy of price prediction.

In the above table, the best models for the price prediction stack in all three scenarios are also introduced. As can be seen, depending on the set of features examined in all three scenarios, the type and order of placement of learning models in the price prediction stack also changes. By comparing the proposed models in all three scenarios, it can be seen that the "random forest" are common in all three models (although the order of their placement in each scenario is different). This may imply that “Random Forest” is the best estimator of Bitcoin Price. The other interesting point is that the estimator model in the first and third scenarios are the same.

Depending on the type of available data, any of the mentioned models can be used to predict the price of Bitcoin. Another important point is that according to the model introduced in the first scenario, even with little data (just the daily opening price of Bitcoin), it is possible to predict the price of Bitcoin with relatively high accuracy.

Type of features	Training time (Seconds)	Mean Square error	The best model for price forecast stack	Number of selected features
OHLC	123.2943141	1.4568612229103932	RandomForestRegressor(Params_1) RandomForestRegressor(Params_2) RandomForestRegressor(Params_3) RandomForestRegressor(Params_4) MLPRegressor(Params_5)	1
OHLC variables and technical indicators	1991.781	1.4477005817242348	Lasso Lars (Params_1) Random ForestRegressor(Params_2) Random ForestRegressor(Params_3) Random ForestRegressor(Params_4) Linear Regression (Params_5)	22
OHLC variables, technical indicators and fundamental indices	1730.719	1.4537249534629806	RandomForestRegressor(Params_1) RandomForestRegressor(Params_2) RandomForestRegressor(Params_3) RandomForestRegressor(Params_4) MLPRegressor(Params_5)	23

5.Conclusions

By acknowledging cryptocurrencies as an investment opportunity and a new financial asset, the need to address issues such as determining optimal trading strategies, predicting the future trend of the price and returns of cryptocurrencies according to past trends and their fundamental value, and using risk management tools appears to be too important. Considering Bitcoin's market share of about 50% of the market value of cryptocurrencies, if there is a limitation in the investment opportunity, Bitcoin has always been the first choice of investors in the digital currency market (due to its high market share and the leadership of Bitcoin's price in cryptocurrency market).

In most studies, technical analysis, fundamental analysis, econometric methods, and machine learning-based methods have been used to determine trading strategies and predict cryptocurrency market conditions. The rapid development of machine learning in recent years has increased their use in cryptocurrency trading, especially in price prediction and cryptocurrency returns. The rapid development of machine learning techniques in recent years has increased their application in cryptocurrency trading, especially in price and returns' prediction. In this paper, by using meta-heuristic and ensemble learning methods (combination of machine learning methods), an estimation model with the least error for Bitcoin's price is presented. In this research, the stacking approach of popular machine learning algorithms along with the optimization of algorithm parameters has been used. In this study, the method of stacking regression algorithms (to find the most optimal combination of stacks) as well as identifying the most suitable parameters of each of these algorithms has been done by using the differential evolution method. To achieve the best result, the combination used in the model includes a stack of 5 machine learning algorithms. According to the investigations carried out so far, the differential evolution method has not been used to optimize issues related to the price of cryptocurrencies, and this study is the first research conducted using this method. Considering that in most articles, technical indicators are mostly used as input to machine learning models, the use of fundamental indices is one of the distinguishing aspects of this research compared to other similar studies. In this article, more than 160 technical indicators have been used as input in the proposed model, which is more than the 124 technical indicators used in a similar article in 2019.

In this research, the results of the mentioned model have been compared in three scenarios with inputs including OHLC features, technical indicators, fundamental analysis indicators and their combination. In the research, Bitcoin trading statistics between 1/1/2019 and 29/9/2021 have been used to conduct studies. The results of this research show that depending on the set of features examined in all three scenarios, the type and order of placement of learning models in the price prediction stack change. All three scenarios examined in this research have acceptable accuracy for predicting the price of Bitcoin. According to the investigations, the lowest prediction error is in the case where OHLC variables and technical indicators are used as inputs in the prediction model. Another important point is the greater weight of the characteristics of technical analysis indicators to increase the accuracy of price prediction.

By comparing the results, it can be seen that depending on the type of available data, any of the mentioned models can be used to predict the price of Bitcoin. Another important point is that according to the model introduced in the first scenario, even with little data (just opening price of Bitcoin), it is possible to predict the daily price of Bitcoin with high accuracy.

References

1) Alessandretti, L., ElBahrawy, A., Aiello, L.M., Baronchelli, A., 2018. Anticipating cryptocurrency prices using machine learning. Complexity 2018.
2) Attanasio, G., Cagliero, L., Garza, P., Baralis, E., 2019. Quantitative cryptocurrency trading: exploring the use of machine learning techniques, in: Proceedings of the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets, ACM. p. 1.
3) Auer R., Claessens S. (2018). Regulating cryptocurrencies: Assessing market reactions. BIS Working Paper.
4) Ausan A., Demir E., Gozgor G., Lau C. (2019). Effects of the geopolitical risks on Bitcoin returns and volatility. Research in International Business and Finance, V. 47, P. 511-518.
5) Barnwal, A., Bharti, H., Ali, A., Singh, V., 2019. Stacking with neural network for cryptocurrency investment. arXiv preprint arXiv:1902.07855

6) Benediktsson J, Other Contributors. TA-lib Indicators; 2017.
7) Bheemaiah, Kary, Collomb, Alexis. 2018, CRYPTOASSET VALUATION: Identifying the variables of analysis, Working Report v1.0.
8) Bo eshagh M., Eftekhari, M., “, "A new memetic approach in numerical optimization: Hierarchical differential evolution algorithm”, 9th symposium of Scince and technology advancements, Mashhad, 2014.
9) Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.
10) Burniske, C, Tartar, J., 2018, Cryptoassets, the innovative investor’s guide to Bitcoin and beyond, McGrae Hill Education.
11) Cheng, J., Dong, L., Lapata, M., 2016. Long short-term memorynetworks for machine reading. arXiv preprint arXiv:1601.06733 .
12) Fan Fang & Carmine Ventre & Michail Basios & Leslie Kanthan & Lingbo Li & David Martinez-Regoband & Fan Wu, 2020. "Cryptocurrency Trading: A Comprehensive Survey," Papers 2003.11352, arXiv.org, revised Jan 2022.
13) Gabralla, Lubna A., and Ajith Abraham. "Prediction of oil prices using bagging and random subspace." In Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014, pp. 343-354. Springer, Cham, 2014.
14) Hayes A. ,2018. Bitcoin price and its marginal cost of production: Support for a fundamental value. Applied Economics Letters, V. 26, No. 7, P. 554-560.
15) Huang, J. ,Huang W., Ni J., Predicting bitcoin returns using high-dimensional technical indicators, The Journal of Finance and Data Science, 2019. Vol 5, Issue 3, Pp. 140-155.
16) Jianliang, M., Haikun, S., Ling, B., 2009. The application on intrusion detection based on k-means cluster algorithm, in: 2009 International Forum on Information Technology and Applications, IEEE. pp. 150–152.
17) Kalchbrenner, N., Grefenstette, E., Blunsom, P., 2014. A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188
18) Karam A., Turani M. “Landslide susceptibility zoning using linear regression methods and hierarchical analysis process, case study: Haraz axis from Rodhen to Rineh. Applied Research of Geographical Sciences (Geographical Scienc).
19) Keller, Christoph A., and Mat J. Evans. "Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10." Geoscientific Model Development 12, no. 3 (2019): 1209-1225.
20) Kiana, Danial. 2019. Cryptocurrency Investing For Dummies, John Wiley & Sons, Inc.
21) Kraaijeveld, Olivier, De Smedt, Johannes. 2020, The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. Journal of International Financial Markets, Institutions & Money, V. 65, P 1-22.
22) Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., et al., 2005. Applied linear statistical models. volume 5. McGraw-Hill Irwin New York.
23) Kwon, D.H., Kim, J.B., Heo, J.S., Kim, C.M., Han, Y.H., 2019. Time series classification of cryptocurrency price trend based on a recurrent lstm neural network. Journal of Information Processing Systems 15.
24) Madan, I., Saluja, S., Zhao, A., 2015. Automated bitcoin trading via machine learning algorithms. URL: http://cs229. stanford.edu/proj2014/Isaac%20Madan 20.
25) Mikolov, T., Kombrink, S., Burget, L., Cˇ ernocky`, J., Khudanpur, S., 2011. Extensions of recurrent neural network language model, in: 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE. pp. 5528–5531.
26) Murphy, J. “ Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications., New York Institute of Finance, Fishkill, N.Y., 1999.
27) Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf
28) Nakano, M., Takahashi, A., Takahashi, S., 2018. Bitcoin technical trading with artificial neural network. Physica A: Statistical Mechanics and its Applications 510, 587–609.
29) Persson, S., Slottje, A., Shaw, I., 2018. Hybrid Autoregressive-Recurrent Neural Network Architecture for Algorithmic Trading of Cryptocurrencies. Cs230 deep learning thesis. Stanford University.
30) Phaladisailoed, T., Numnonda, T., 2018. Machine learning models comparison for bitcoin price prediction, in: 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), IEEE
31) Price K., RM Storn, JA Lampinen, DE a Practical Approach to Global Optimization, Springer, 2005.
32) Rane, P.V., Dhage, S.N., 2019. Systematic erudition of bitcoin price prediction using machine learning techniques, in: 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), IEEE. pp. 594–598.

33) Rebane, J., Karlsson, I., Denic, S., Papapetrou, P., 2018. Seq2seq rnns and arima models for cryptocurrency prediction: A comparative study. SIGKDD Fintech 18.
34) Sadeghi, A., Daneshvar, A., Madanchi Zaj, M., Combined ensemble multi-class SVM and fuzzy NSGA-II for trend forecasting and trading in Forex markets. Journal of Expert Systems with Applications (2021), Vol 185.
35) Sadeghi A., Daneshvar A., Madanchi Zaj M., “Development of an intelligent method based on fuzzy technical indicators for forecasting and trading the Euro-Dollar parity rate”, Financial Engineering and Securities Management Quarterly, No. 45, December 2019.
36) Shabankareh, Mohammad Javad, Mohammad Ali Shabankareh, Alireza Nazarian, Alireza Ranjbaran, and Nader Seyyedamiri. "A Stacking-Based Data Mining Solution to Customer Churn Prediction." Journal of Relationship Marketing (2021): 1-24.
37) Slepaczuk, R., Zenkova, M., 2018. Robustness of support vector machines in algorithmic trading on cryptocurrency market. Central European Economic Journal 5, 186 – 205.
38) Sriram, A., Jun, H., Satheesh, S., Coates, A., 2017. Cold fusion: Training seq2seq models together with language models. arXiv preprint arXiv:1708.06426 .
39) Stuerner, P., 2019. Algorithmic Cryptocurrency Trading. Ph.D. thesis. Ulm University.
40) Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267–288.
41) Thakkar, Priyank. "Comparing Multiple Linear Regression and Linear Support Vector Regression for Predicting in Stock Market”, 2015.
42) Usai, M. Graziano, Mike E. Goddard, and Ben J. Hayes. "LASSO with cross-validation for genomic selection." Genetics research 91, no. 6 (2009): 427-436.
43) Virk, D.S., 2017. Prediction of Bitcoin Price using Data Mining. Master’s thesis. National College of Ireland.
44) Vladimirovna, Narimanova Olga, Fariman ogly, Narimanov Nariman, 2021, CRYPTOCURRENCY MARKET: CURRENT STATE AND DEVELOPMENT TRENDS, International Conference on Process Management and Scientific Developments, pp 30-35.
45) Vo, A., Yost-Bremm, C., 2018. A high-frequency algorithmic trading strategy for cryptocurrency. Journal of Computer Information Systems , 1–14
46) Wolpert, D. (1992). Stacked generalization (stacking). Neural Networks, 5(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1
47) Yang, Yujun, and Yimei Yang. "Hybrid prediction method for wind speed combining ensemble empirical mode decomposition and bayesian ridge regression." IEEE Access 8 (2020): 71206-71218.
48) Zhengyang, W., Xingzhou, L., Jinjin, R., Jiaqing, K., 2019. Prediction of cryptocurrency price dynamics with multiple machine learning techniques, in: Proceedings of the 2019 4th International Conference on Machine Learning Technologies, ACM. pp. 15–19.

Appendix 1 (List of technical indicators used in the research)

Overlap Studies

BBANDS	Bollinger Bands
DEMA	Double Exponential Moving Average
EMA	Exponential Moving Average
HT_TRENDLINE	Hilbert Transform - Instantaneous Trendline
KAMA	Kaufman Adaptive Moving Average
MA	Moving average
MAMA	MESA Adaptive Moving Average
MAVP	Moving average with variable period
MIDPOINT	MidPoint over period
MIDPRICE	Midpoint Price over period
SAR	Parabolic SAR
SAREXT	Parabolic SAR - Extended
SMA	Simple Moving Average
T3	Triple Exponential Moving Average (T3)
TEMA	Triple Exponential Moving Average
TRIMA	Triangular Moving Average
WMA	Weighted Moving Average

Momentum Indicators

ADX	Average Directional Movement Index
ADXR	Average Directional Movement Index Rating
APO	Absolute Price Oscillator
AROON	Aroon
AROONOSC	Aroon Oscillator
BOP	Balance Of Power
CCI	Commodity Channel Index
CMO	Chande Momentum Oscillator
DX	Directional Movement Index
MACD	Moving Average Convergence/Divergence
MACDEXT	MACD with controllable MA type
MACDFIX	Moving Average Convergence/Divergence Fix 12/26
MFI	Money Flow Index
MINUS_DI	Minus Directional Indicator
MINUS_DM	Minus Directional Movement
MOM	Momentum
PLUS_DI	Plus Directional Indicator
PLUS_DM	Plus Directional Movement
PPO	Percentage Price Oscillator
ROC	Rate of change : ((price/prevPrice)-1)*100
ROCP	Rate of change Percentage: (price-prevPrice)/prevPrice
ROCR	Rate of change ratio: (price/prevPrice)
ROCR100	Rate of change ratio 100 scale: (price/prevPrice)*100
RSI	Relative Strength Index
STOCH	Stochastic
STOCHF	Stochastic Fast
STOCHRSI	Stochastic Relative Strength Index
TRIX	1-day Rate-Of-Change (ROC) of a Triple Smooth EMA
ULTOSC	Ultimate Oscillator
WILLR	Williams' %R

Volume Indicators

AD	Chaikin A/D Line
ADOSC	Chaikin A/D Oscillator
OBV	On Balance Volume

Cycle Indicators

HT_DCPERIOD	Hilbert Transform - Dominant Cycle Period
HT_DCPHASE	Hilbert Transform - Dominant Cycle Phase
HT_PHASOR	Hilbert Transform - Phasor Components
HT_SINE	Hilbert Transform - SineWave
HT_TRENDMODE	Hilbert Transform - Trend vs Cycle Mode

Volatility Indicators

AVGPRICE	Average Price
MEDPRICE	Median Price
TYPPRICE	Typical Price
WCLPRICE	Weighted Close Price

Pattern Recognition

CDL2CROWS	Two Crows
CDL3BLACKCROWS	Three Black Crows
CDL3INSIDE	Three Inside Up/Down
CDL3LINESTRIKE	Three-Line Strike
CDL3OUTSIDE	Three Outside Up/Down
CDL3STARSINSOUTH	Three Stars In The South
CDL3WHITESOLDIERS	Three Advancing White Soldiers
CDLABANDONEDBABY	Abandoned Baby
CDLADVANCEBLOCK	Advance Block
CDLBELTHOLD	Belt-hold
CDLBREAKAWAY	Breakaway
CDLCLOSINGMARUBOZU	Closing Marubozu
CDLCONCEALBABYSWALL	Concealing Baby Swallow
CDLCOUNTERATTACK	Counterattack
CDLDARKCLOUDCOVER	Dark Cloud Cover
CDLDOJI	Doji
CDLDOJISTAR	Doji Star
CDLDRAGONFLYDOJI	Dragonfly Doji
CDLENGULFING	Engulfing Pattern
CDLEVENINGDOJISTAR	Evening Doji Star
CDLEVENINGSTAR	Evening Star
CDLGAPSIDESIDEWHITE	Up/Down-gap side-by-side white lines
CDLGRAVESTONEDOJI	Gravestone Doji
CDLHAMMER	Hammer
CDLHANGINGMAN	Hanging Man
CDLHARAMI	Harami Pattern
CDLHARAMICROSS	Harami Cross Pattern
CDLHIGHWAVE	High-Wave Candle
CDLHIKKAKE	Hikkake Pattern
CDLHIKKAKEMOD	Modified Hikkake Pattern
CDLHOMINGPIGEON	Homing Pigeon
CDLIDENTICAL3CROWS	Identical Three Crows
CDLINNECK	In-Neck Pattern
CDLINVERTEDHAMMER	Inverted Hammer
CDLKICKING	Kicking
CDLKICKINGBYLENGTH	Kicking - bull/bear determined by the longer marubozu
CDLLADDERBOTTOM	Ladder Bottom
CDLLONGLEGGEDDOJI	Long Legged Doji
CDLLONGLINE	Long Line Candle
CDLMARUBOZU	Marubozu
CDLMATCHINGLOW	Matching Low
CDLMATHOLD	Mat Hold
CDLMORNINGDOJISTAR	Morning Doji Star
CDLMORNINGSTAR	Morning Star
CDLONNECK	On-Neck Pattern
CDLPIERCING	Piercing Pattern
CDLRICKSHAWMAN	Rickshaw Man
CDLRISEFALL3METHODS	Rising/Falling Three Methods
CDLSEPARATINGLINES	Separating Lines
CDLSHOOTINGSTAR	Shooting Star
CDLSHORTLINE	Short Line Candle
CDLSPINNINGTOP	Spinning Top
CDLSTALLEDPATTERN	Stalled Pattern
CDLSTICKSANDWICH	Stick Sandwich
CDLTAKURI	Takuri (Dragonfly Doji with very long lower shadow)

CDLTASUKIGAP Tasuki Gap
CDLTHRUSTING Thrusting Pattern
CDLTRISTAR Tristar Pattern
CDLUNIQUE3RIVER Unique 3 River
CDLUPSIDEGAP2CROWS Upside Gap Two Crows
CDLXSIDEGAP3METHODS Upside/Downside Gap Three Methods

Statisitic Function
BETA Beta
CORREL Pearson's Correlation Coefficient (r)
LINEARREG Linear Regression
LINEARREG_ANGLE Linear Regression Angle
LINEARREG_INTERCEPT Linear Regression Intercept
LINEARREG_SLOPE Linear Regression Slope
STDDEV Standard Deviation
TSF Time Series Forecast
VAR Variance

Math Transform
ACOS Vector Trigonometric ACos
ASIN Vector Trigonometric ASin
ATAN Vector Trigonometric ATan
CEIL Vector Ceil
COS Vector Trigonometric Cos
COSH Vector Trigonometric Cosh
EXP Vector Arithmetic Exp
FLOOR Vector Floor
LN Vector Log Natural
LOG10 Vector Log10
SIN Vector Trigonometric Sin
SINH Vector Trigonometric Sinh
SQRT Vector Square Root
TAN Vector Trigonometric Tan
TANH Vector Trigonometric Tanh

Math Operators
ADD Vector Arithmetic Add
DIV Vector Arithmetic Div
MAX Highest value over a specified period
MAXINDEX Index of highest value over a specified period
MIN Lowest value over a specified period
MININDEX Index of lowest value over a specified period
MINMAX Lowest and highest values over a specified period
MINMAXINDEX Indexes of lowest and highest values over a specified period
MULT Vector Arithmetic Mult
SUB Vector Arithmetic Substraction
SUM Summation

Appendix 2 (Parameters used to test variables in the three scenarios examined in the research)

Parameters used to test OHLC variables.				
Params_5	Params_4	Params_3	Params_2	Params_1
activation='identity' alpha=0.0001 batch_size='auto' beta_1=0.9 beta_2=0.999 early_stopping=True epsilon=1e-08 hidden_layer_sizes=(100) learning_rate='constant' learning_rate_init=0.001 max_fun=15000 max_iter=200 momentum=0.9 n_iter_no_change=10 nesterovs_momentum=False power_t=0.5 random_state=None shuffle=True solver='lbfgs' tol=0.0001 validation_fraction=0.0 verbose=False warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=2 max_features='sqrt' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=2 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mse' max_depth=2 max_features='sqrt' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=150 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=None max_features='log2' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=150 n_jobs=None oob_score=False random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=2 max_features='log2' max_leaf_nodes=2 max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=4 min_weight_fraction_leaf=0.0 n_estimators=150 n_jobs=None oob_score=False random_state=None verbose=0 warm_start=False

Parameters used to test OHLC variables and technical indicators.				
Params_5	Params_4	Params_3	Params_2	Params_1
copy_X=True fit_intercept=False n_jobs=None normalize=True positive=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=2 max_features='sqrt' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=2 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mse' max_depth=2 max_features='log2' max_leaf_nodes=2 max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=4 min_weight_fraction_leaf=0.0 n_estimators=100 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mse' max_depth=None max_features='auto' max_leaf_nodes=2 max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=4 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	alpha=0.0 copy_X=True eps=2.220446049250313e-16 fit_intercept=True fit_path=True jitter=None max_iter=100 normalize=True positive=True precompute='auto' random_state=None verbose=False

Parameters used to test OHLC variables, technical indicators and fundamental indices.				
Params_5	Params_4	Params_3	Params_2	Params_1
activation='logistic' alpha=0.0001 batch_size='auto' beta_1=0.9 beta_2=0.999 early_stopping=False epsilon=1e-08 hidden_layer_sizes=(100) learning_rate='constant' learning_rate_init=0.001 max_fun=15000 max_iter=200 momentum=0.9 n_iter_no_change=10 nesterovs_momentum=True power_t=0.5 random_state=None shuffle=True solver='lbfgs' tol=0.0001 validation_fraction=0.0 verbose=False warm_start=False	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=2 max_features='sqrt' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=2 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=2 max_features='log2' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=2 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=True	bootstrap=True ccp_alpha=0.0 criterion='mae' max_depth=None max_features='auto' max_leaf_nodes=None max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=2 min_weight_fraction_leaf=0.0 n_estimators=100 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=False	bootstrap=True ccp_alpha=0.0 criterion='mse' max_depth=None max_features='sqrt' max_leaf_nodes=2 max_samples=None min_impurity_decrease=0.0 min_impurity_split=None min_samples_leaf=4 min_samples_split=4 min_weight_fraction_leaf=0.0 n_estimators=50 n_jobs=None oob_score=True random_state=None verbose=0 warm_start=False

The detailed explanation of parameters can be found in the (www. scikit-learn.org.) website.