Avrupa Bilim ve Teknoloji Dergisi
Sayı 21, S. 444-454, Ocak 2021
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 21, pp. 444-454, January 2021
Copyright © 2021 EJOSAT
Research Article

Comparison of Statistical and Machine Learning Algorithms for Forecasting Daily Bitcoin Returns

Betül Aygün[1*], Eylül Günay Kabakçı[2]

[1*] İzmir Demokrasi Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, YBS, İzmir, Türkiye (ORCID: 0000-0001-9610-9235),betul.aygun@idu.edu.tr
[2] İzmir Demokrasi Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat, İzmir, Türkiye (ORCID: 0000-0001-5547-4316), eylul.kabakci@idu.edu.tr

(First received 5 November 2020 and in final form 23 January 2021)

(**DOI:** 10.31590/ejosat.822153)

ATIF/REFERENCE: Aygün, B., Günay Kabakçı, E. (2021). Comparison of Statistical and Machine Learning Algorithms for Forecasting Daily Bitcoin Returns. *European Journal of Science and Technology*, (21), 444-454.

Abstract

Increasing fluctuations in pricing and having great profit potential, utilization in advanced machine learning technologies to make robust predictions of cryptocurrencies especially bitcoin have attracted great attention in recent years. In this study, various statistical techniques; Moving Average Analysis and Autoregressive Integrated Moving Average and machine learning (ML) techniques; Artificial Neural Network, Recurrent Neural Network (RNN) and Convolutional Neural Network have been conducted and compared to predict the future value of Bitcoin cryptocurrency price. They have been applied for the univariate time series analysis with a window size of 32. To prove the usefulness of ML algorithms, and to show that the results of RNN is a better, mean squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) indicators have been applied. The study revealed that recurrent neural network yields better results than other methods in predicting daily Bitcoin price in terms of MSE, MAE and MAPE metrics. Besides, Wilcoxon-Mann-Whitney nonparametric statistic test is applied to test the performance between ARIMA and machine learning algorithms.

Keywords: Bitcoin, Statistical Analysis, Machine Learning, DNN, RNN, CNN, MVA, ARIMA.

Günlük Bitcoin Değerini Tahmin Etmek İçin İstatistiksel ve Makine Öğrenimi Algoritmalarının Karşılaştırılması

Öz

Fiyatlandırmada artan dalgalanmalar ve büyük kar potansiyeline sahip olan Bitcoin başta olmak üzere kripto para birimlerinin sağlam tahminini yapmak için gelişmiş makine öğrenimi teknolojilerinin kullanılması son yıllarda büyük ilgi gördü. Bu çalışmada çeşitli istatistiksel teknikler; Hareketli Ortalama Analizi ve Otoregresif Entegre Hareketli Ortalama ve makine öğrenimi (ML) teknikleri; Yapay Sinir Ağı, Tekrarlayan Sinir Ağı (RNN) ve Evrişimli Sinir Ağı, Bitcoin kripto para birimi fiyatının gelecekteki değerini tahmin etmek için uygulanmıştır ve bulunan sonuçlar karşılaştırılmıştır. Bu teknikler 35 pencere boyutu ile tek değişkenli zaman serisi analizi kapsamında uygulandı. Makine öğrenimi algoritmalarının yararlılığını kanıtlamak ve RNN sonuçlarının daha iyi olduğunu göstermek için ortalama hata karesi (MSE), ortalama mutlak hata (MAE) ve ortalama mutlak yüzde hata (MAPE) göstergeleri uygulanmıştır. Çalışma, tekrarlayan sinir ağının MSE, MAE ve MAPE ölçümleri açısından günlük Bitcoin fiyatını tahmin etmede diğer yöntemlerden daha iyi sonuçlar verdiğini ortaya koydu. Bununa birlikte, ARIMA ve makine öğrenme algoritmalarının performansını karşılaştırmak için Wilcoxon-Mann-Whitney (WMW) parametrik olmayan istatistik testi uygulanmıştır.

Anahtar Kelimeler: Bitcoin, Tahminleme, İstatiksel Analiz, Makine öğrenmesi, DNN, RNN, CNN, MVA, ARIMA

[*] Corresponding Author: betul.aygun@idu.edu.tr

1. Introduction

The digital transformation, which we frequently encounter in all areas, started with Satoshi Nakamoto putting forward the idea of Bitcoin currency about 10 years ago in finance, which is probably one of the most difficult areas (Nakamoto, 2008). After Bitcoin is introduced, other digital cryptocurrencies such as Ethereum, XRP or Stellar, have emerged remarkably fast. However, Bitcoin still totally dominates the digital economy and has the biggest portion among cryptocurrencies. Market capitalization is almost 250 billion USD as of now according to the https://www.coindesk.com/price/bitcoin and 300.000 transactions occur per day in November 2020. On the contrary to classic currency, Bitcoin transaction is carried out between two parties without any need for the third trusted financial institution.

As in all financial currencies, predicting the future value of Bitcoin is started to become a very important and interesting subject area that gets attention by both researchers and economist because it is highly unpredictable (Higbee, 2018). As stated in the study of Zhang et al., there are four approaches for prediction in which two of which are traditional time series forecasting and machine learning algorithms (Zhang & Wan, 2007). For predicting the future values of the currency in the literature, time series analysis is generally carried out (Jalles, 2009). On the other hand, Refenes et al. claim that traditional statistical forecasting is fade out of the side due to the nonlinearities in the finance data set (Refenes, Zapranis, & Francis, 1994) In addition to this, several studies are presented that machine learning algorithms, having capability to adaptation both linear and nonlinear models, has advantages over traditional time series analysis (Zhang & Wan, 2007) (Yao & Tan, 2000). Moreover, using machine learning algorithms for time series analysis becomes popular and attractive for other engineering subjects (Özhan, 2020) (Güleryüz & Özden, 2020).

Starting from this point, during this study, for the prediction of Bitcoin currency, univariate time series analysis is performed and instead of not only classical traditional statistical analysis; moving average analysis and autoregressive integrated moving average analysis; but also, machine learning, and neural networks are applied to predict the future value. The first aim is to propose an algorithm that has high accuracy for forecasting. As a feature for the independent variable, only the closing price (USD) of Bitcoin daily has been considered. Since currency change is taken as daily, the dataset contains 2713 values. Secondly, analysis is performed to conclude that whether the various neural network learning algorithms for predicting the future value of the Bitcoin currency is applicable or not. From this point of view, one-layer artificial neural network, three-layer artificial neural network (ANN), a convolutional neural network (CNN) and a recurrent neural network (RNN) was applied (Tsai, Zeng, & Chang, 2018). These algorithms were compared by using three main metrics; mean absolute error, mean squared error, and mean absolute percentage error. Besides, the Wilcoxon-Mann-Whitney test was carried out to decide whether the differences between the performances of the models are statistically significant or not. Also, algorithms were examined according to the execution times. Further, since many studies about forecasting Bitcoin do not cross-validate which causes overfitting the data (Albariqi & Winarko, 2020), cross-validation to prevent it is carried out.

In summary, during this research, we design different neural networks architecture to predict the future value of the Bitcoin cryptocurrency and discuss the suitability of these for the Bitcoin dataset. The main contribution stated during this study is summarized as below:

- In addition to statistical techniques, ANN, RNN and CNN were applied to predict the future Bitcoin price and the performance of these approaches over standard statistical approaches was discussed.
- The serious shortages caused by machine learning algorithms while designing neural networks were discussed and several recommendations to overcome them were given.

The study is structured as follows: In Section 2, the studies in the literature about Bitcoin forecasting are summarized. Section 3 details the theoretical background of used statistical methods and proposed artificial network designs with hyper-parameter optimization. Section 4 begins with data preprocessing and continues with the results of the techniques used. In Section 5, the results of the study are discussed and in Section 6, a conclusion is given, and several future works are proposed.

2. Literature Review

There exist two different approaches in the literature for the Bitcoin price prediction. In the first approach, researchers focus on the influences of some specific factors like trading volume, Bitcoin popularity and attractiveness. In this way, the explanatory concept is tried to be presented. Chen et al. and Liu et al. predicted the Bitcoin future value by using the technological parameters such as block size, mining profitability and economical parameters such as gold price or crude oil price with Bitcoin exchange rate. These types of studies are certainly impressive especially in selecting technological and economical determinants (Chen, Xu, Jia, & Gao, 2020) (Liu, Li, Li, Zhu, & Yao, 2020). From this approach, there exist enlightening studies in the literature. Moreover, instead of predicting the future value of Bitcoin, several studies focus on anomaly detection to observe currency fluctuations by using clustering algorithms like support vector machines, DBSCAN or k-means algorithms (Dokuz, Çelik, & Ecemiş, 2020).

In the second approach, researchers focus on just Bitcoin currency values such as opening, average, high or low values as time series analysis to predict the future value of Bitcoin. Since in this study, we predict the perspective of the second approach, only the studies about the second approach on literature are summarized in this section.

Since statistical techniques are more popular for time series analysis, there exist many studies that uses ARIMA model to forecast the future value of Bitcoin currency (Bakar & Rosbi, 2017) (Ayaz, Fiaidhi, Sabah, & Ansari, 2020) (Azari, 2019) (Alahmari, 2019) (Munim, Shakil, & Alon, 2019) (Yenidoğan, Çayir, Kozan, Dağ, & Arslan, 2018). Besides, in the study of Munim et al., it was claimed that forecast results of ARIMA is better than neural network autoregression (Munim, Shakil, & Alon, 2019). Greaves et al. and Atsalakis et al. proposed neuro-fuzzy models to predict values (Atsalakis, Atsalaki, Pasiouras, & Zopounidis, 2019) (Greaves & Au, 2015). Mudassir et al., Madan et al., and Lahmiri et al. all implemented support vector machine prediction algorithms (Mudassir, Bennbaia, Unal, & Hammoudeh, 2020) (Madan, Saluja, & Zhao, 2015) (Lahmiri & Bekiros, 2020). Besides, in these studies, random forest, kNN and neural network algorithms were also applied. In fact, there are huge number of researchers that used the different type of neural networks mostly RNN for Bitcoin forecasting in the literature (Phaladisailoed & Numnonda, 2018) (Adcock & Gradojevic, 2019) (Nakano, Takahashi, & Takahashi, 2018) (McNally, Roche, & Caton, 2018). In addition, several researchers applied the LSTM to predict the future value of Bitcoin in their studies (Deokar, Dandage, & Jawandhiya, 2020) (Ulumuddin, Sunardi, & Fadlil, 2020) (Chen, Xu, Jia, & Gao, 2020). On the other hand, Liu et al. used the SDAE method to predict Bitcoin (Liu, Li, Li, Zhu, & Yao, 2020).

When the literature is carefully reviewed, it was seen that convolutional networks are rarely used for the Bitcoin price prediction. In the study of Li and Dai, hybrid network model combining CNN and LSTM is proposed (Li & Dai, 2020). Their results show that the hybrid model improves the accuracy compared with a single model.

There are different approaches categorized as daily, weekly, monthly, minutely, and secondly in the literature in terms of time interval of Bitcoin dataset. Madan et al. constructed three datasets, first one is daily values, the second one is 10 minutes time interval and the last one is 10 seconds interval (Madan, Saluja, & Zhao, 2015). Besides, several studies take into consideration of predicting other popular cryptocurrencies. Valencia et al. and Alahmari did not only predict Bitcoin but also predict the values of Ethereum, Ripple and Litecoin and XRP in their studies (Alahmari, 2019) (Valencia, Gómez-Espinosa, & Valdés-Aguirre, 2019).

The studies about Bitcoin price prediction are summarized in Table 2 in terms of the in which time interval the dataset is taken, and which algorithms are used to. The studies conducted since 2018 are handled, and they are sorted according to the publishing date from newest to oldest in Table 1.

3. Methodology

3.1. Theoretical Background of the Algorithms

Simple Moving average (SMA) analysis is the simplest forecasting method to predict the future value of time series data. It calculates the average of the values in the defined period of the data. It is a form of lagging indicator and evaluate the average Bitcoin price over time. SMA is a technical indicator to set the future value of the Bitcoin price with the average values in the previous N window size. The mathematical formula for SMA is as follows:

$$\boldsymbol{F_{n+1}} = \frac{\boldsymbol{A_1 + A_2 + \cdots + A_N}}{\boldsymbol{N}} = \frac{\boldsymbol{\sum_{i=1}^{N} A_n}}{\boldsymbol{N}} \tag{1}$$

F_{n+1} is the predicted value of (n+1).day and A_n, n is in [1, N] is the actual price value of n^{th} day. N is the window size that represents the closing Bitcoin prices taken in previous N days. Due to its simplicity in the background of the SMA theory, ARIMA method is applied.

ARIMA is a popular statistical analysis technique for forecasting time series data. After firstly introduced, it has widely used in many areas for prediction. It stands for Autoregressive integrated moving average and is represented as ARIMA(p,d,q). p parameter is used for autoregressive, which represents the relationship between an observation and number of lagged observations; and q is used for moving average part. d is the degree of difference that makes the time series stationary. In the ARIMA model, p is the number of previous observations in autoregression and formulated as:

$$\boldsymbol{x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \cdots + \phi_p x_{t-p} + \phi} \tag{2}$$

In this model, q is the number of previous number of errors used and formulated as the following.

$$\boldsymbol{x_t = \theta - \theta_1 x_{t-1} - \theta_2 x_{t-2} - \theta_3 x_{t-3} - \cdots - \theta_q x_{t-q}} \tag{3}$$

Lastly, d is the number of differencing to make time-series data stationary. If d=1 then

$$\boldsymbol{\nabla x_t = x_t - x_{t-1}} \tag{4}$$

By using the previous formula, the mathematical formulation to predict the x_t value is as the following in this study:

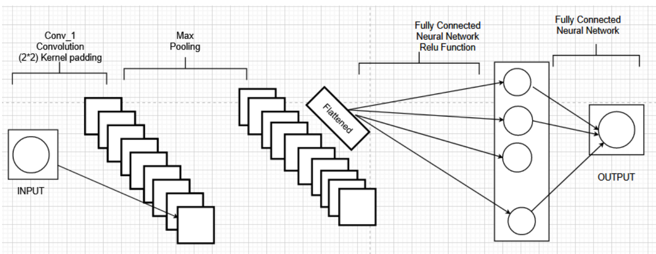
$$\begin{aligned}\boldsymbol{\nabla x_t = } & \boldsymbol{\phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \cdots + \phi_p x_{t-p} + \phi} \\ & \boldsymbol{+ \theta - \theta_1 x_{t-1} - \theta_2 x_{t-2} - \theta_3 x_{t-3} - \cdots} \\ & \boldsymbol{- \theta_q x_{t-q}}\end{aligned} \tag{5}$$

On the other hand, to apply ARIMA statistical technique, there are many constraints such as making data stationary or linearizing by taking the exponential of data. In addition, this method is generally poor at predicting turning points (Meyler, Kenny, & Quinn, 1998) (Wadi, 2018) which happens in our case.

Although the concept for artificial neural networks dates back to the 1950s, it is still used in many areas such as engineering, medical, finance and for several various data analyses like clustering, regression, classification, or prediction. During this study, we apply artificial neural network, convolutional neural network and recurrent neural network which are generally different in terms of theoretical background.

ANN, the simplest neural network, consisting of many layers in which it contains several neurons. It is a known as feed forward network in which the information passes in one direction through various nodes. A convolutional neural network contains convolutional layers that are connected or pooled entirely. In the convolutional layers, the input is transformed by using filters. The CNN network architecture used in this study is given in Figure 1.

Figure 1. CNN Network Architecture



On the other hand, the recurrent neural network is more complex and pass information in two directions and contains cycles and loops. The activation function used in all neural networks during this study is rectified linear unit (relu) function shown in the following formula:

Table 1. Summaries of Related Works

Studies	Time Interval/ Period	Algorithms
Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants (Chen, Xu, Jia, & Gao, 2020)	Period I: August 2011 to December 2013 Period II: August 2011 to December 2014 Period III: July 2014 to December 2017 Period IV: July 2015 to July 2018	ANN, SVM, RF, LSTM, ARIMA
Forecasting the price of Bitcoin using deep learning (Liu, Li, Li, Zhu, & Yao, 2020)	July 2013 - December 2019	SDAE, BPNN, SVR
Design & Implementation of Crypto Currency Prediction Using Machine Learning Approach (Deokar, Dandage, & Jawandhiya, 2020)	Bitcoin: January 2012 to March 2018 (minutely	RNN and LSTM
Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model (Li & Dai, 2020)	December 2016 – August 2018	Hybrid of CNN with LSTM
Forecasting the movements of Bitcoin prices: an application of machine learning algorithms (Pabuçcu, Ongan, & Ongan, 2020)	2008-2019	SVM, ANN, NB and RF
Prediction of Bitcoin Price Change using Neural Networks (Albariqi & Winarko, 2020)	August 2010 to October 2017 / 2-days period	Multilayer Perceptron, RNN
Bitcoin Price Prediction using ARIMA model (Ayaz, Fiaidhi, Sabah, & Ansari, 2020)	August 2019 to January 2020 /daily	ARIMA
Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach (Mudassir, Bennbaia, Unal, & Hammoudeh, 2020)	Interval I April 2013 to July 2016 Interval II April 2013 to April 2017 Interval III April 2013 to December, 2019	ANN, SANN, SVM and LSTM
Intelligent forecasting with machine learning trading systems in chaotic intraday Bitcoin market (Lahmiri & Bekiros, 2020)	January 2016 to March 16, 2018 / 5 min interval	SVR, GRP, RT, kNN, BPNN, BRNN, and RBFNN
Bitcoin Price Prediction: An ARIMA Approach (Azari, 2019)	September 2015 to September 2018 / daily	ARIMA
Using Machine Learning ARIMA to Predict the Price of Cryptocurrencies (Alahmari, 2019)	Bitcoin April 28, 2013 to December 15, 2018 XRP April 8, 2013 to December 18, 2018 Ethereum August 7, 2015 to December 18, 2018 /daily, weekly, and monthly	ARIMA
Next-Day Bitcoin Price Forecast (Munim, Shakil, & Alon, 2019)	January 1, 2012 to October 4, 2018	ARIMA and neural network autoregression
Bitcoin price forecasting with neuro-fuzzy techniques (Atsalakis, Atsalaki, Pasiouras, & Zopounidis, 2019)	September 13, 2011 to October 12, 2017 / daily	Fuzzy Logics, Neural Networks, Adaptive Neuro Fuzzy Inference System
Non-fundamental, non-parametric Bitcoin forecasting (Adcock & Gradojevic, 2019)	July 19, 2010 to March 5, 2018 /daily	ARIMAX, GARCH; Linear Regression, Quantile Regression ANN, Recurrent ANN, Kernel regression
Bitcoin technical trading with artificial neural network (Nakano, Takahashi, & Takahashi, 2018)	July 31, 2016 15:00 (GMT) to January 24, 2018 07:30 (GMT) / 15 min interval	ANN with different layers, activation function and inputs
Machine Learning Models Comparison for Bitcoin Price Prediction (Phaladisailoed & Numnonda, 2018)	January 1, 2012 - January 18, 2018 / 1-minute interval	Theil Sen Regression, Huber Regression, LSTM, Gated Recurrent Unit
Bitcoin Forecasting Using ARIMA and PROPHET (Yenidoğan, Çayir, Kozan, Dağ, & Arslan, 2018)	May 2016 to March 2018 / daily	ARIMA and PROPHET
Predicting the Price of Bitcoin Using Machine Learning (McNally, Roche, & Caton, 2018)	August 2013 to July 2016 / daily	ARIMA, RNN and LSTM
Forecasting of Bitcoin Daily Returns with	July 2010 to January 17, 2018 / daily	

EEMD-ELMAN based Model (Khaldi, Afia, Chiheb, & Faizi, 2018)		EEMD-ELMAN based Model
Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction (Bakar & Rosbi, 2017)	January 2013 to October 2017 / monthly	ARIMA
Automated Bitcoin Trading via Machine Learning Algorithms (Madan, Saluja, & Zhao, 2015)	/daily, 10 min interval, 10 sec intervals	Binomial GLM, Random Forest and SVM

$$\text{ReLU} = \boldsymbol{f(x)} = \begin{cases} \mathbf{0}, & \boldsymbol{x} < \mathbf{0} \\ \boldsymbol{x}, & \boldsymbol{x} \geq \mathbf{0} \end{cases} \quad [0,\infty] \tag{6}$$

3.2. Hyper-Parameter Settings

For the design of each network of machine learning algorithms and each statistical method, there are diverse parameters such as number of hidden layers, nodes, learning parameter and window size that needs to be adjusted to reach satisfactory results.

For the SMA analysis, there is only one parameter; window length and it is selected empirically with trial-and-error. For the ARIMA model, parameter values p, d and q are very important for the effectiveness of the algorithms. The optimum solution is achieved by using Akaike Information Criteria (AIC) introduced by (Bozdogan, 1987).

During this study, trial and error approach is used to set the window size of the neural network. For the robust search, the algorithms were tried from 20 days to the lag of 100 days by increasing 5 and 35 is found as the most effective window length.

For the learning rate of the neural network layers, the loss vs learning rate function is used to set up it. This function is drawn by using the Learning rate scheduler function of Keras library. The graphs of the loss vs learning rate and comments about it are detailed in each of the applied network algorithms. On the other hand, number of hidden layers, number of nodes, type of activation function must be set up cautiously for the reason that the performance of the networks varies hugely depending on the parameters. To keep the size of parameters in a minimum, the models were created containing the minimum number of layers and parameters were selected empirically with trial and error. Furthermore, parameters are determined by the existing studies. Since Relu function is used in several studies in the literature, it is used as an activation function for training especially recurrent neural networks throughout the study. In addition, other activation functions (sigmoid, tanh) were also used during the study, and the results were analyzed.

3.3. Performance Measures

For the forecasting performance measure, mean squared error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE) are used. MAE and MSE estimate the average of absolute Euclidean distance and squared Euclidean distance respectively between predicted and actual values. The formulas for these metrics are given in the Table 2.

Table 2. Metrics and Formulations

Metrics	Formulas
Mean Squared Error	$\frac{1}{N}\sum_{t=1}^{N} {e_t}^2$
Mean Absolute Error	$\frac{1}{N}\sum_{t=1}^{N} \lvert e_t \rvert$
Mean Absolute Percentage Error (MAPE)	$\frac{100\%}{N}\sum_{t=1}^{N} \left\lvert \frac{e_t}{y_t} \right\rvert$

3.4. Used Technology

Because several different algorithms were applied during this study, Python was used for creating all models. For the statistical analysis such as ARIMA, statsmodels library was used. Nolds, pyEntropy and pyrem packages were used for the entropy analysis. For the machine learning algorithms and evaluation metrics, keras deep learning library was used.

4. Results

During this study, Simple Moving Average and ARIMA statistical models and Artificial Neural Network, Convolutional Neural Network and Recurrent Neural Network machine learning algorithms are applied to predict the future price of the Bitcoin financial asset and compared to conclude which algorithm produces better results.

Bitcoin daily prices are considered as time series data and these techniques were applied in the concept of time series analysis. All approaches used in this study make forecasting future values based on the previous ones.

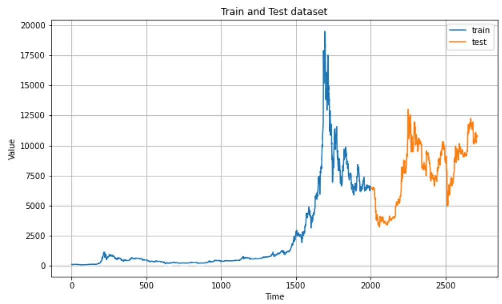
4.1. Data Preprocessing and Feature Extraction

The Bitcoin data set was collected by https://www.coindesk.com. It is investigated at one day frequency ranges from 10th of September 2013 to 2nd of October 2020 and contains closing prices. The reason is that the study of McNally et al. presents that the close price is one of the most important variables among open, high, low, and closing prices (McNally, Roche, & Caton, 2018). Besides, the study of Atsalakis et al. also uses the close prices of Bitcoin in their research (Atsalakis, Atsalaki, Pasiouras, & Zopounidis, 2019).

Feature selecting is one of the most important parts for the deep learning algorithms to reach better results in forecasting. The dataset has totally 2713 elements and divided into two subsets:

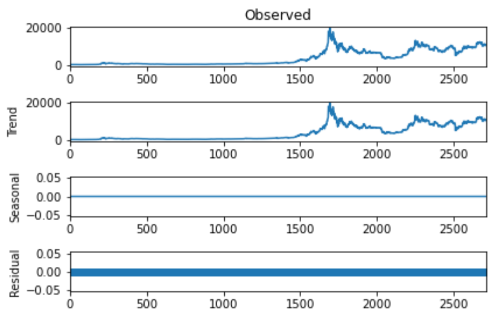
training and testing data. In the literature, there are different approaches for the division of samples into training and testing data such as 50:50, 70:30 and 80:20. In this study, first 2000 samples were used as training data and the remaining approximately %25 of samples were used for testing the model proposed. The training and testing data are illustrated in Figure 2. The data used for training the model was drawn in blue line and the data draw in orange line was used for testing the model.

Figure 2. Train and Test Dataset



Time series for Bitcoin data is decomposed into four main part; random, seasonal, trend and observed. The time series plot of the data is given in Figure 3. As shown in the Figure 3, there is no clear seasonality and residual in Bitcoin dataset but, there is a trend. It is demonstrated that prices are non-stationary time series with ADF testing statistics of -1.386372 in Figure 5. However, to apply ARIMA time series analysis, the dataset has been transformed to stationary data series.

Figure 3. Decomposition of additive Bitcoin dataset timeseries

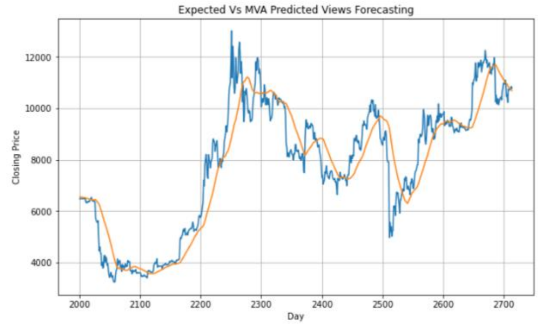


Dataset is a univariate time series containing the closing prices of Bitcoin as a sequence of observations. This sequence data must be transformed into input and output features to apply supervised learning techniques. It is done by correlating Bitcoin data into window size of 35 days. Thus, to forecast the daily Bitcoin value, previous 35 days are used as input for each network design.

4.2. Forecasting Results

Firstly, SMA is applied to forecast Bitcoin price. The predictive price is the average value of the previous n time series data. The n value is called as window size and it is taken as 30 in this study. In other words, the average value of 30 previous Bitcoin close prices gives us the Bitcoin price of 31st day. In this algorithm, each price is equally weighted. The graph of the results for the prediction of the test data set is given in Figure 4. In all figures showing the predictive analysis of the algorithms, the blue line shows the exact data and the orange line represents the predicted value.

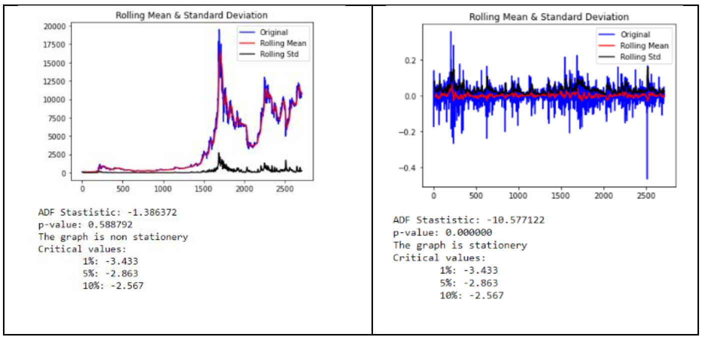
Figure 4. Moving Average Analysis Forecasting



SMA performs usually insufficiently in ranging markets that have no obvious trend in movement. Besides, SMA does not have the capability of prediction market movements (Ellis & Parbery, 2005) and it is the simplest method that just evaluates the averages. Due to these deficiencies, Autoregressive Integrated Moving Average (ARIMA) statistical method, the most robust and prominent forecasting technique, was implemented to make an accurate prediction (Sato, 2013).

To apply the ARIMA method, the time series data must be stationary. An augmented Dickey–Fuller test (ADF) is done to test whether the graph is stationary or not. Since p-value (0.588792) > 0.5, the test confirms that the data set is non-stationary. Before implementation, differencing in which seasonality and trend are eliminated is applied to make data stationary. The original dataset and differenced dataset stationary results are given in Figure 5.

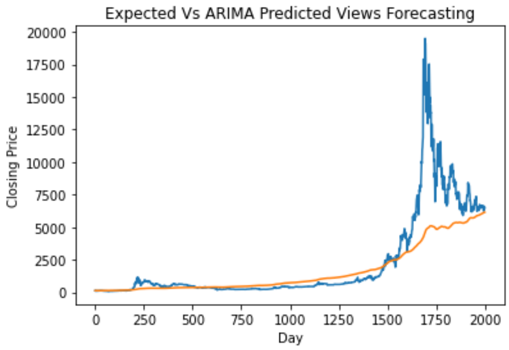
Figure 5. ADF Statistics Results



ARIMA model is applied to the differenced dataset and the following results have obtained. Model is created by using 25th order auto regressive model with 3rd moving average. The predicted values of ARIMA (25,1,3) model and expected values for testing data are detailed in Figure 6. In each prediction, the ARIMA model is compiled again and again with new added data makes this model cumbersome. Further, it has difficulty in forecasting turning points.

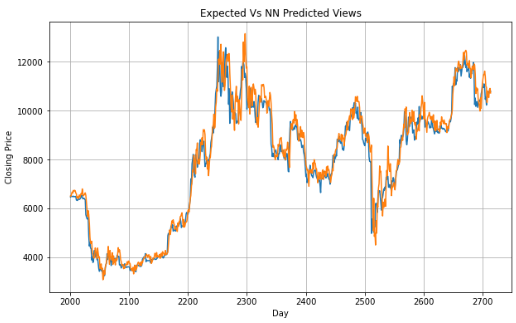
For one-layer, three-layers, and recurrent neural networks, as an optimizer, a stochastic gradient descent algorithm with $5e^{-10}$ learning rate is used. The learning rate is selected by using the learning rate scheduler function. The graphic of loss vs different learning rates showed that the learning rate of $5e^{-10}$ seems to yield better results. To increase the prediction performance and yield better results, 3-layers ANN are added to the design of the network. As an activation function, rectified linear unit, the most popular activation function, for neural network algorithms is used. Output array of the first layer has shape of (None, 30) and second layer has a shape of (None, 40) and third and the last layer has a shape of (None, 1) which is the forecasted value of Bitcoin price. Under 200 epochs, the results of the predicted and actual values of one-layer neural network and three-layers neural network are shown in Figure 7 and 8, respectively. These networks are feed forward networks that have not cycles, loops, or filters.

Figure 6. ARIMA Predicted and Expected Values



Besides, we apply CNN to predict Bitcoin price. The idea behind this network is filtering data by extracting the right and relevant features in the input data. 1 dimensional convolution layer with 88 filter and 2 kernel size is used. As in the previous networks, relu function is used for the activation function. After the convolution layer, feed forward network layer is added. The summary of CNN design is given in Figure 9.

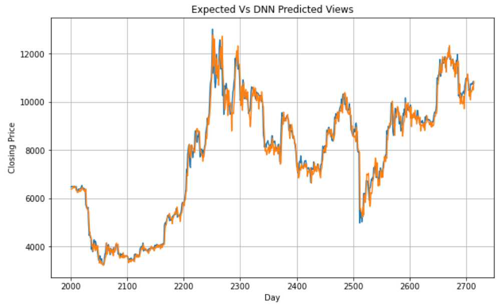
Figure 7. ANN with one layer Predicted and Expected Values



One-layer Artificial Neural Network Model Summary

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_3 (Dense)              (None, 1)                 36
=================================================================
Total params: 36
Trainable params: 36
Non-trainable params: 0
_________________________________________________________________
```

Figure 8. ANN with 3 layers Predicted and Expected Values



Three-layer Neural Network Model Summary

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 30)                1080
_________________________________________________________________
dense_2 (Dense)              (None, 40)                1240
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 41
=================================================================
Total params: 2,361
Trainable params: 2,361
Non-trainable params: 0
_________________________________________________________________
```

In the literature, there have been many variants of ANN depending on data flows. To improve the performance, we applied networks with cyclic data flows called Recurrent neural networks. These networks have memory and have two inputs: current data and past data. So, RNN evaluates the output based on the previous computations. In practice, they are limited to look back in a defined number of steps. The parameters used for RNN is given in Table 3. The algorithms run under 30 tests with 200 epochs. Considering the execution time even out of scope in this study, RNN algorithms execute in 430.35 seconds which is almost nine times the execution time of ANN algorithms which lasts 40.034 seconds. In summary, although this approach produces more accurate predictions than other neural networks, execution time of the algorithm is more than the others. The summary of RNN design is given in Figure 10. It represents that the predicted and expected values are highly close each other. Two simple RNN layer with 40 units are created. RNN layers have the full sequence as input.

Table 3. Parameter Values for Neural Networks

Parameters/Artificial Neural Networks	**1-Layer Artificial Neural Network**	**3-Layers Artificial Neural network**	**Recurrent neural network**	**Convolutional Neural Network**
Batch Size	32	32	32	32
# of Inputs	35	35	35	35
# of Outputs	1	1	1	1
# of Hidden Layer/# of nodes of each layer	1(1)	3(40,40,1)	3(40,40,1)	3(88,200,1)
Activation Function	Rectified Linear Unit	Rectified Linear Unit	Rectified Linear Unit	Rectified Linear Unit
Loss Function	Mean Squared Error	Mean Squared Error	Mean Squared Error	Mean Squared Error
Optimization Function	Stochastic Gradient Descent	Stochastic Gradient Descent	Stochastic Gradient Descent	Adaptive Moment Estimation
Learning Rate	5e-10	5e-10	5e-10	0.01

Figure 9. CNN Predicted and Expected Values

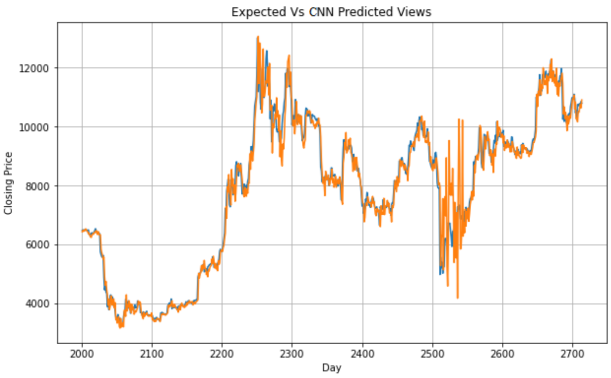
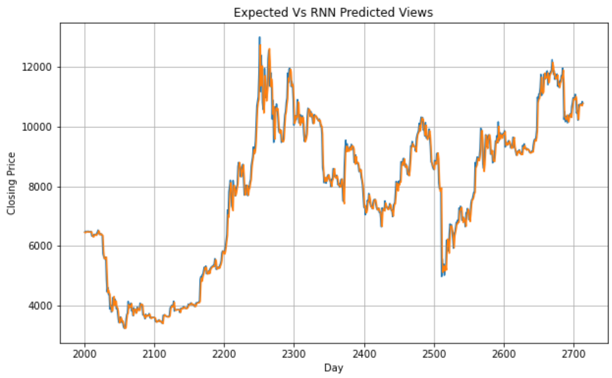


Figure 10.1 RNN Predicted and Expected Values



Convolutional Neural Network

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lambda (Lambda)              (None, 35, 1)             0
_________________________________________________________________
conv1d (Conv1D)              (None, 34, 88)            264
_________________________________________________________________
max_pooling1d (MaxPooling1D) (None, 34, 88)            0
_________________________________________________________________
flatten (Flatten)            (None, 2992)              0
_________________________________________________________________
dense (Dense)                (None, 200)               598600
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 201
_________________________________________________________________
lambda_1 (Lambda)            (None, 1)                 0
=================================================================
Total params: 599,065
Trainable params: 599,065
Non-trainable params: 0
```

Recurrent Neural Network Model Summary

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lambda (Lambda)              (None, 35, 1)             0
_________________________________________________________________
simple_rnn (SimpleRNN)       (None, 35, 40)            1680
_________________________________________________________________
simple_rnn_1 (SimpleRNN)     (None, 40)                3240
_________________________________________________________________
dense (Dense)                (None, 1)                 41
=================================================================
Total params: 4,961
Trainable params: 4,961
Non-trainable params: 0
```

4.3. Evaluations of Predictive Performance

In this section, a comparative discussion of the results for proposed networks are provided. While training and testing all neural networks, the same datasets are used as inputs. Table 4 summarizes the results of statistical and neural network algorithms by using the MSE, MAE and MAPE metrics detailed mathematically in the previous sections. To get more information about the results, mean, standard deviation, maximum and minimum statistical indicators were evaluated under the average of 30 runs independently. The best results for each metric are stated as bold in Table 4.

Table 4. Descriptive Statistics for the Performance of Different Models

	Analysis	Mean	Standard Deviation	Minimum	Maximum
MSE	*One Layer - ANN*	184628.33	**3096.51**	174625.75	186763.47
	Three Layer - ANN	170703.52	10381.69	155929.40	191888.40
	CNN	174822.84	23627.49	149403.69	238071.50
	RNN	**114554.01**	14418.32	**98488.79**	**152228.23**
MAE	*One Layer - ANN*	313.29	**4.99**	298.34	316.72
	Three Layer - ANN	299.06	17.84	272.23	330.62
	CNN	284.04	29.61	247.36	359.55
	RNN	**210.56**	18.22	**192.10**	**257.24**
MAPE	*One Layer - ANN*	4.05	**0.07**	3.85	4.10
	Three Layer - ANN	3.84	0.22	3.52	4.22
	CNN	3.75	0.40	3.27	4.79
	RNN	**2.70**	0.25	**2.46**	**3.34**

Table 4 clearly demonstrate that recurrent neural network approach could yield better predictive performance. RNN achieve the lowest mean of three indicators: MSE, MAE and MAPE of 114554.01, 210.56 and 2.70, respectively. CNN is the second approach having high performance in predictive ability. From the

Table, one-layer network has obviously the worst one in predicting Bitcoin price accurately.

To further examine the robustness of the methods, Wilcoxon-Mann-Whitney (WMW) test is conducted to compare the differences between the performance of the prediction methods statistically significant. WMW is a nonparametric and pairwise comparison statistical test. In addition to this, it does not expect that the distribution of the sample should be normal which makes it safer. The aim is to detect significant differences between the behavior of two algorithms (Derrac, García, Molina, & Herrera, 2011).

We test the following statistical hypothesis for three different indicators. p-values of the WMW test are shown in Table 5 for three evaluation indicators: MSE, MAR and MAPE. It is common method to compare the machine learning algorithms in terms of performance. If the p-value is less than 0,05, 0,01 or 0,001, the null hypothesis H_0 is rejected.

H_0: The difference in the performance of the algorithms is not statistically significant.

As stated in Table 5, p-statistic values between RNN and other algorithms except CNN are significant at level $p < 0.05$. Both accuracy metrics and WMW test show that RNN model is more accurate and robust at predicting Bitcoin prices than other algorithms except CNN in this study. However, the differences between RNN and CNN is not statistically significant for MSE and MAPE indicators. Since the p-value = ***0.5205049139315545*** > 0.05. However, p-value in MAE indicator between RNN and CNN represents that differences between algorithms are significant.

5. Discussion

Through this paper, we have shown that machine learning algorithms yield better results with low error rates than traditional statistical techniques. Two statistical and three machine learning algorithms with different natures were used. Moving average analysis which predicts the future value by taking the average of the previous n values, having the simplest computation, produced the highest error. On the other hand, for time series analysis, ARIMA, a popular technique used by several researchers yields better results than the simplest Neural network algorithm. However, making data stationary is a prerequisite condition to apply this technique.

On the other hand, during this research, neural network algorithms were applied on non-stationary time series of Bitcoin prices. Furthermore, even on non-stationary data, the recurrent neural network performed best among the neural network models and statistical models used during this study. Since RNN has a recurrent connection, it captured the sequence exist and temporal dependence in the input data which is the natural result of the type of time series data. For machine learning algorithms, since the amount of data affects the success of the forecasting, when Bitcoin prices are taken hourly rather than a day like in this study, the measured performance of the RNN could become more remarkable.

Table 5. Results of the t-test under three evaluation indicators

Indicators	Models	ARIMA	One-layer ANN	Three-layer ANN	CNN	RNN
MSE	**ARIMA**	-	8.671988141602892e-07	8.671988141602892e-07(***)	8.671988141602892e-07 (***)	8.671988141602892e-07 (***)
	One-layer ANN	-	-	3.4461451484174774e-05(***)	0.02723125198594554(*)	8.671988141602892e-07 (***)
	Three-layer ANN	-	-	-	**0.5205049139315545**	**0.5205049139315545**
	CNN	-	-	-	-	**0.5205049139315545**
	RNN	-	-	-	-	-
MAE	**ARIMA**	-	8.671988141602892e-07(***)	8.671988141602892e-07(***)	8.671988141602892e-07(***)	8.671988141602892e-07(***)
	One-layer ANN	-	-	0.0011275619453935454	4.459363712229841e-05(***)	8.671988141602892e-07(***)
	Three-layers ANN	-	-	-	0.01091336079801136(*)	0.01091336079801136(*)
	CNN	-	-	-	-	0.01091336079801136(*)
	RNN	-	-	-	-	-
MAPE	**ARIMA**	-	8.671988141602892e-07(***)	8.671988141602892e-07(***)	8.671988141602892e-07(***)	8.671988141602892e-07(***)
	One-layer ANN	-	-	0.00033196064213223927(***)	0.0012922796032760046(**)	8.671988141602892e-07(***)
	Three-layers ANN	-	-	-	**0.099305104972773**	**0.099305104972773**
	CNN	-	-	-	-	**0.099305104972773**
	RNN	-	-	-	-	-

$p<0.05$(*), $p<0.01$(**) and $p<0.001$(***)

Besides, even several analysis is achieved while setting parameters, still these do not guarantee to select the best ones. So, with applying more promising approaches for parameter selection from the literature, it is expected that nonlinear neural networks achieve better results than ARIMA and other statistical methods that push the data linearization.

6. Conclusions and Future Works

To sum up, during this study, different statistical techniques, and different neural networks in terms of mathematical background were implemented to predict the future value of the Bitcoin price. To measure the performance of the statistical techniques and networks, three performance indicators; MSE, MAE and MAPE were evaluated and for the robustness of the methods, Wilcoxon-Mann-Whitney test was applied. Performance indicators of MAE, MSE and MAPE showed that RNN yields better results than all other algorithms. However, WMW test claimed that the accuracy difference between RNN and CNN is not statistically significant. We can conclude that the performance of RNN is better than ARIMA and ANN and statistically significant. In addition to this, among the neural networks, more sophisticated networks gave better results as expected. However, designing these networks and selecting appropriate parameters require a great deal of effort. Even though in this study, we use the trial and error for window size parameter selection and the graphic of loss vs learning rate for the selection of learning rate, there are still many diverse parameters and the different number of layers that must be adjusted. Designing the model as sequential or functional, selecting suitable loss and optimizer function for the model and/or activation function for the layer, deciding the size of output nodes and layers are all parameters that must be managed carefully in the feature work. Along these lines, for the parameter's selection, heuristic search methods could be implemented. Otherwise, the trial and error approach has a limited capability, time-consuming and does not guarantee desirable parameters to predict satisfactory results. On the other hand, adding daily opening, mean prices of Bitcoin, market cap, volatility, or volume as other features that make the problem data multivariate time series might produce better prediction results. In addition to these macro-economic factors, adding technological properties with social parameters like twitter numbers might improve the predictive capability of the models.

References

Adcock, R., & Gradojevic, N. (2019). Non-fundamental, non-parametric Bitcoin forecasting. Physica A: Statistical Mechanics and its Applications.

Alahmari, S. A. (2019). Using Machine Learning ARIMA to Predict the Price of Cryptocurrencies. ISeCure-The ISC International Journal of Information Security, 139-144.

Albariqi, R., & Winarko, E. (2020). Prediction of Bitcoin Price Change using Neural Networks. International Conference on Smart Technology and Applications (ICoSTA) (s. 1-4). IEEE.

Atsalakis, G. S., Atsalaki, I. G., Pasiouras, F., & Zopounidis, C. (2019). Bitcoin price forecasting with neuro-fuzzy techniques. European Journal of Operational Research, 770-780.

Ayaz, Z., Fiaidhi, J., Sabah, A., & Ansari, M. A. (2020). Bitcoin Price Prediction using ARIMA Model.

Azari, A. (2019). Bitcoin price prediction: An ARIMA approach. arXiv preprint arXiv:1904.05315.

Bakar, N. A., & Rosbi, S. (2017). Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction. . International Journal of Advanced Engineering Research and Science.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 345-370.

Chen, W., Xu, H., Jia, L., & Gao, Y. (2020). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. International Journal of Forecasting, 28-43.

Deokar, R. S., Dandage, S. M., & Jawandhiya, P. M. (2020). Design & Implementation of CryptoCurrency Prediction Using Machine Learning Approach. International Journal of Advent Research in Computer and Electronics (IJARCE), 8-12.

Derrac, J., García, S., Molina, D., & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm and Evolutionary Computation, 3-18.

Dokuz, A. Ş., Çelik, M., & Ecemiş, A. (2020). Anomaly Detection in Bitcoin Prices using DBSCAN Algorithm. European Journal of Science and Technology, 436-443.

Ellis, C. A., & Parbery, S. A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. Research in International Business and Finance, 399-411.

Greaves, A., & Au, B. (2015). Using the Bitcoin Transaction graph to predict the price of Bitcoin. http://snap.stanford.edu/class/cs224w-2015/projects _ 2015/ adresinden alındı

Güleryüz, D., & Özden, E. (2020). The Prediction of Brent Crude Oil Trend Using LSTM and Facebook Prophet. Avrupa Bilim ve Teknoloji Dergisi, 1-9.

Higbee, A. (2018). The role of crypto-currency in cybercrime. Computer Fraud & Security, 13-15.

Jalles, J. T. (2009). Structural Time Series Models and the Kalman Filter: A Concise Review.

Khaldi, R., Afia, A. E., Chiheb, R., & Faizi, R. (2018). Forecasting of Bitcoin daily returns with EEMD-ELMAN based model. International Conference on Learning and Optimization Algorithms: Theory and Applications, (s. 1-6).

Lahmiri, S., & Bekiros, S. (2020). Intelligent forecasting with machine learning trading systems in chaotic intraday Bitcoin market. Chaos, Solitons & Fractals.

Li, Y., & Dai, W. (2020). Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. The Journal of Engineering, 344-347.

Liu, M., Li, G., Li, J., Zhu, X., & Yao, Y. h. (2020). Forecasting the price of Bitcoin using deep learning. Finance Research Letters.

Madan, I., Saluja, S., & Zhao, A. (2015). Automated bitcoin trading via machine learning algorithms. http://cs229. stanford. edu/proj2014/Isaac% 20Madan adresinden alındı

McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of bitcoin using machine learning. 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (s. 339-343). IEEE.

Meyler, A., Kenny, G., & Quinn, T. (1998). Forecasting irish inflation using ARIMA models . Central Bank and Financial Services Authority of Ireland Technical Paper Series.

Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural Computing and Applications, 1-15.

Munim, Z. H., Shakil, M. H., & Alon, I. (2019). Next-day bitcoin price forecast. Journal of Risk and Financial Management.

Nakamoto, S. (2008). Bitcoin whitepaper. https://bitcoin. org/bitcoin. pdf-(Дата обращения: 17.07. 2019).: https://bitcoin. org/bitcoin. pdf-(Дата обращения: 17.07. 2019). adresinden alındı

Nakano, M., Takahashi, A., & Takahashi, S. (2018). Bitcoin technical trading with artificial neural network. Physica A: Statistical Mechanics and its Applications, 587-609.

Özhan, E. (2020). Estimation Of CO2 Equivalent Greenhouse Gas Emissions In Turkey By Artificial Neural Networks And Exponential Smoothing Method. European Journal of Science and Technology, 282-289.

Pabuçcu, H., Ongan, S., & Ongan, A. (2020). Forecasting the movements of Bitcoin prices: an application of machine learning algorithms. Quantitative Finance and Economics , 679-691.

Phaladisailoed, T., & Numnonda, T. (2018). Machine learning models comparison for bitcoin price prediction. 10th International Conference on Information Technology and Electrical Engineering (ICITEE) (s. 506-511). IEEE.

Refenes, A. N., Zapranis, A., & Francis, G. (1994). Stock performance modeling using neural networks: a comparative study with regression models. Neural networks, 375-388.

Sato, R. C. (2013). Disease Management with ARIMA model in Time Series. Einstein.

Tsai, Y. T., Zeng, Y. R., & Chang, Y. S. (2018). Air pollution forecasting using RNN with LSTM. 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (s. 1074-1079). IEEE.

Ulumuddin, I., Sunardi, & Fadlil, A. (2020). Bitcoin Price Prediction Using Long Short Term Memory (LSTM). JurnalMatik, 1090-1095.

Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. Entropy.

Wadi, S. A. (2018). Predicting closed price time series data using ARIMA Model. Modern Applied Science.

Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. Neurocomputing, 79-98.

Yenidoğan, I., Çayir, A., Kozan, O., Dağ, T., & Arslan, Ç. (2018). Bitcoin forecasting using ARIMA and prophet. 3rd International Conference on Computer Science and Engineering (UBMK) (s. 621-624). IEEE.

Zhang, Y. Q., & Wan, X. (2007). Statistical fuzzy interval neural networks for currency exchange rate time series prediction. Applied Soft Computing, 1149-1156.