Engineering Applications
https://publish.mersin.edu.tr/index.php/enap
e-ISSN 2979-9201

Forecasting through neural networks: Bitcoin price prediction

Katerina Zela *[1], **Lorena Saliaj** [2]

[1]*Mediterranean University of Albania, Department of Information Technology, Albania, katerina.male@umsh.edu.al*
[2]*Mediterranean University of Albania, Department of Business Informatics, Albania, lorenasaliaj@umsh.edu.al*

Cite this study: Zela, K., & Saliaj, L. (2023). Forecasting through neural networks: Bitcoin price prediction. Engineering Applications, 2 (3), 218-224

Keywords
Bitcoin
Time series forecasting
ANN
ARIMA

Research Article
Received: 07.03.2023
Revised: 15.08.2023
Accepted: 06.09.2023
Published:14.09.2023

check for updates

Abstract

This paper concerns the problem of daily Bitcoin price prediction, aiming to find the best predictive model among the linear and nonlinear forecasting models. Finding the most accurate forecasting model would help investors take important decisions about taking the next step when investing. We compare the forecasting performance of linear and nonlinear forecasting models using daily Bitcoin price data for the period between 31 December 2017 until 24 November 2021. We discuss various forecasting approaches, including an Autoregressive Integrated Moving Average (ARIMA) model, a Nonlinear Autoregressive Neural Network (NARNN) model, a TBATS model and Exponential Smoothing on the data collected from 31 December 2017 to 24 November 2021 and compared their accuracy using the data collected from 01 June 2021 to 09 June 2021, choosing the model with the lowest Mean Absolute Percentage Error (MAPE) value. The chosen model has been used for daily Bitcoin price forecasting for the next 60 days without any additional intervention. The forecasting model can be applied to other cryptocurrencies available on the global cryptocurrency market cap.

1. Introduction

Since its creation, Bitcoin turns out to be the most traded currency in the world and occupies a significant part of the cryptocurrency market. Its birth marked the launch of a new asset class and a major step forward in forms of centralized control. Unlike other currencies, Bitcoin does not have a central bank that regulates the distribution of the currency, but is based on two principles: a network of nodes, composed of computers and cryptography to make transactions valid and secure. The value of Bitcoin has gone from 0, in 2009, to 57,873 in November 2021. Price prediction can be very useful, as it can help the decision-making process regarding possible investments in purchasing the currency.

In this article, we have tested the accuracy of predictions obtained from the most proposed models in the literature, including linear and nonlinear prediction models. The aim of this paper is to identify the most suitable model for predicting future values, starting from its value reported every day since January 2018, through the use of four different prediction models in the Bitcoin price time series. and comparing the validity of these models to analyze its progress for the next 60 days.

2. Material and Method

In this article, we have considered the data published online on Coin Market Cap, the website that reflects the daily price performance of cryptocurrencies, for the period from December 31, 2017 to November 24, 2021, considering the time series of the closing price for the period December 31, 2017- November 24, 2021, the price of the last 8 days for testing the validity of the model and the last 60 days for predicting the future price.

Forecasting was performed using the R software forecasting package, which provides methods and tools for univariate time series forecasting. We implemented an ARIMA model, a Nonlinear Autoregressive (NNAR) model, a TBATS model as well as a Linear Exponential Square and selected the best model among them, considering the mean percentage error (MAPE).

2.1. The forecasting models

2.1.1. NARNN model

Artificial Neural Networks are predictive models inspired by biological neural networks. They identify and model non-linear relationships between the dependent variable and its predictors. A set of neurons, grouped into input, hidden and output layers to form the artificial network, can perform a large number of complex tasks, quite efficiently. This makes ANNs a powerful tool, capable of learning from previous examples and improving its performance, giving them the ability to analyze new data based on previous results. Artificial neural networks are nonlinear models that transform a set of inputs into a set of output variables, through hidden layers of neurons. An ANN is composed of several layers:

- The first layer, known as the input layer, is the one that receives the incoming data. The last layer, called the output layer, provides the results of the analysis or the solution to the problem. Data flows from the input layer to the output layer through one or more intermediate layers called hidden layers. In them the data is analyzed and the required results are obtained. Hidden layer nodes reveal features in the data model and nonlinear relationships between them. Next, the required result is sent from the hidden layer to the output layer. To build a neural network, we need to define the following variables:
- Number of input nodes: corresponds to the number of input layer variables that are used to predict future values. In a time series forecasting problem, the number of input nodes corresponds to the number of lags considered for forecasting. It is preferable to use a small number of input nodes to detect data features, as too few or too many input nodes can affect the learning or prediction ability of the network [1].
- Number of hidden layers and hidden nodes: usually, one hidden layer is sufficient for most prediction problems. Two or more hidden layers are preferred over one hidden layer, especially when a hidden layer network has many nodes, which can lead to unsatisfactory results.
- Number of output nodes: corresponds to the prediction horizon, which can be one-step-ahead (using one output node) or multi-step-ahead prediction.

In our work, the NAR network was developed using the nnetar function of the R software package "caret" to fit a neural network model to a time series [2] developed by Hyndman, O'Hara, and Wang. An NNAR (p,k), where p denotes the number of lags used as input and k the number of nodes in the hidden layer, can be described as an autoregressive process with nonlinear functions. We chose a (1-5-1) network, with 1 delay as input node and 5 hidden layer nodes. It has the form of a three-layer ANN, where neurons have a single connection with neurons of other layers [3]. The data set was divided into the training set (70%), the test set (15%), while the data of the last 8 days was used for testing the validity of the model.

2.1.2. ARIMA

ARIMA (Auto-Regressive Integrated Moving Average) is the most widely used linear model for time series forecasting. This model represents the time series as a function of its past values, called lags, and the lags of the error term. An ARIMA model consists of three terms p,d,q:

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_p \varepsilon_{t-q} \tag{1}$$

where, p is the order of the autoregressive (AR) term, which refers to the number of lags y to be considered as predictors, q is the order of the moving average (MA) term, which refers to the number of error lags used as predictors, while d is the number of differentiations needed to make the time series stationary. The most common approach to making a time series stationary is to subtract the previous value of the time series from the current one. So, d is the minimum number of differentiations that must be made to have a stationary time series. If the time series is stationary, then d=0.

The main objective of the ARIMA model is to predict future values through the stochastic mechanism of time series. Although this model is widely used for time series, it is not easy to choose the right order for its components, so we proceeded to determine the order automatically, by using the auto.arima function in the forecasting package of R, which gave us the best fitting ARIMA model as a result. This involves identifying the most appropriate lags for the AR and MA components and deciding whether the variable needs differentiation. The model that best fitted our time series was ARIMA (0,2,1). This model was used in this study to predict the price of Bitcoin currency for the next few days.

2.1.3. Exponential smoothing

Exponential smoothing, also known as Holt's model, uses double exponential smoothing parameters to predict future values: the first parameter is used for the general smoothing, while the second parameter is used for the trend squaring equation. So, this approach involves a prediction equation and two smoothing equations. The current value of the time series is obtained by considering the last squared value for the trend of the last period and updating the trend for the last period and the following periods, expressing it as the difference between the last two squared values:

$$\hat{y}_{t+h|t} = l_t + h b_t \tag{2}$$

where $l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$ [3] refers to the first equation, while $b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$ is the trend equation, where α refers to the squared parameter, lt refers to the value of the time series at time t, while bt is the trend of the time series at time t.

2.1.4. BATS

The third model is BATS (Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal component), which uses a combination of the Fourier term with an exponential square model and Box-Cox transformation, in an automatic way. The unit of time used in the study was "day". The predictive ability of these models was evaluated through the mean absolute percentage of the MAPE error, as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{3}$$

where, n is the total number of observations, at is the actual value and Ft is the value predicted by the model.

3. Results

The values of the validity indicators of the models, RMSE, MAE, MPE; MAPE, ME, and MASE are shown in Table 1. They were used to measure the performance of the models built for the Bitcoin price time series, taking into account the data used to train the models. Apart from the graph, where it can be clearly seen, the values in the following tables show that the NARNN model has given more accurate prediction results than the other linear models. Based on the MAPE values, NARNN improved the forecasting accuracy by 39% compared to the ARIMA model.

The NARNN model gives better results in almost all the indicators considered, with a significant difference from the indicators of other models. This model has improved the performance of predictions based on the ME indicator by 21% compared to the BATS indicator.

The selection of the best predictive model was based on the value of MAPE, since it is recommended to use this indicator as a comparative unit between predictive models for time series, considering as the most accurate model the one with the lowest value of MAPE. Based on the above, the NNAR model has the lowest value of MAPE (14.2%).

Table 1. Accuracy of predictions for training set.

Model	ME	RMSE	MAE	MPE	MAPE	MASE
ARIMA	11.3	840	436.21	0.07	23.4	1
BATS	32.26	836.48	431.04	0.03	32.1	0.98
HOLT	11.96	842.12	435.83	0.73	33.48	1.09
NNAR	9.37	759.4	432.12	0.19	14.2	0.65

Table 2 shows the MAPE values for the last 8 days used as data for testing the validity of the models. Once again, we can conclude that the NARN model is the best among those used. This confirms our presumption about the model that should be used for such predictions.

Table 2. MAPE (%) for the last 8 days Bitcoin price.

Model	NNAR	HOLT	BATS	ARIMA
MAPE (%)	7.47	11.89	11.06	12.05

The above models were used to predict the price performance of Bitcoin for the next 60 days and compared the data obtained from the forecast to the actual data for an 8-day period (June 25 - July 02). The values of MAPE for each model are presented in Table 3. From the results of the analysis, we can conclude that the results predicted by the model are similar to the real values of the time series. In particular, the NNAR model gave more accurate predictions, since the MAPE values for it were lower compared to the values of the other models.

Table 3. MAPE (%) for model predictions accuracy.

Model	ARIMA	BATS	HOLT	NNAR
2022-06-25	4.61	4.45	4.57	3.45
2022-06-26	9.76	9.25	9.69	7.46
2022-06-27	9.86	9.00	9.75	6.57
2022-06-28	15.05	13.80	14.91	10.67
2022-06-29	12.71	11.13	12.53	7.56
2022-06-30	12.53	10.60	12.32	6.60
2022-07-01	14.31	12.00	14.06	7.54
2022-07-02	17.55	14.83	17.27	9.88

A graphical representation of the predictions is shown in the Figure 1 and Figure 2.

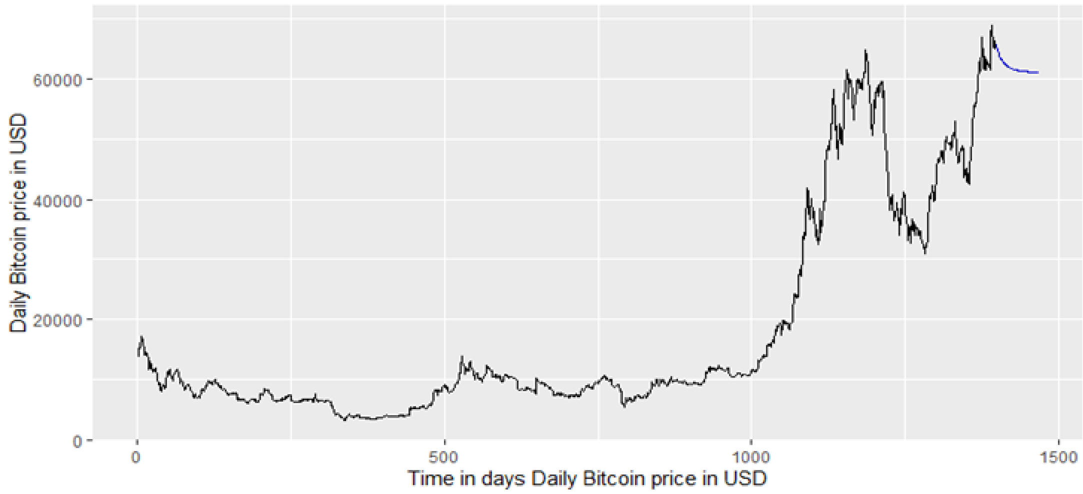


Figure 1. Forecasting according to NARNN model.

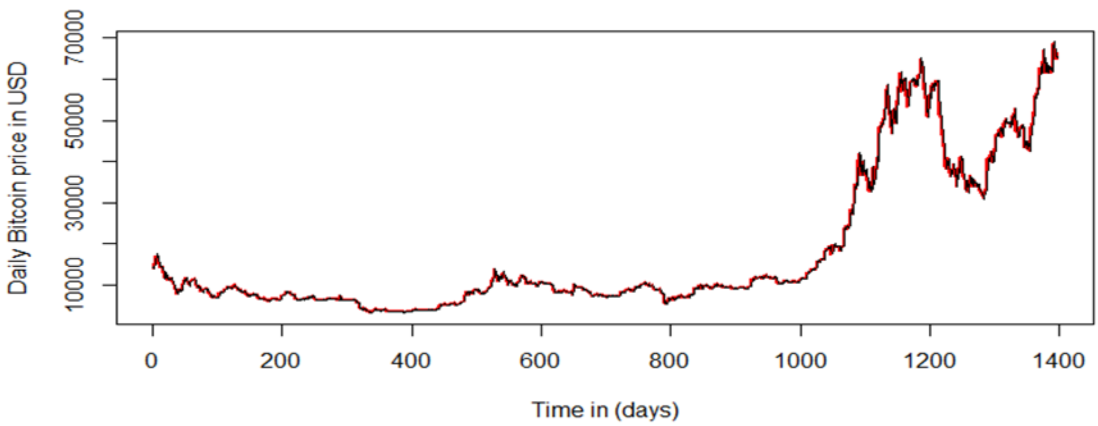


Figure 2. Time series values against predicted values.

For the construction of the NARNN model, the time series data were divided into two groups; model training set and test set. The training set was used to create the model, while the test set was used to evaluate the created model [4]. The mesh structure was chosen based on the results in Zhang et al. [1], who showed that the optimal structure of neural networks contains 1 hidden layer. Since the layer with 5 hidden neurons performed better than those with 1, 2, 3, and 4 hidden neurons, we chose 5 hidden nodes for our model because it had a lower RMSE compared to other models. The neurons of the input layer were selected through the nnetar and accuracy functions. These functions resulted in a neural network with 28 input nodes, resulting in a model (1-5-1). Figure 1 presents the graphical output of Bitcoin price predictions made by the NARNN model for the next 60 days. The model manages to follow the trend of the time series well, thanks to the training and learning process, which enable the model to better understand the characteristics of the time series. Figure 2 shows the fit of the NARNN model to real time series data. All components are well represented and the difference between predicted and observed values tends to zero, thanks to the model's ability to identify non-linear relationships between observations. The ability to learn, to work with parallel and multiple inputs are some of the characteristics that make neural networks efficient in generating models suitable for predicting time series [1].

Figure 3 presents the forecasting according to ARIMA (0,2,1) model, from which we find that there is an upward trend for the next 60 days, with future Bitcoin price values ranging from $65,000 to $85,000. According to the ARIMA model, there will be an increase in the price of Bitcoin during the month of July and August. The confidence interval (colored in blue on the graph) shows that the accuracy of the predictions can vary within that interval. If we compare the values of the 8 days that were used as a test group, we find significant differences between the values predicted by the ARIMA model and the values observed by the time series. This is highlighted by the MAPE value for the 8 days of the test, which for the ARIMA model reaches 17.55%.

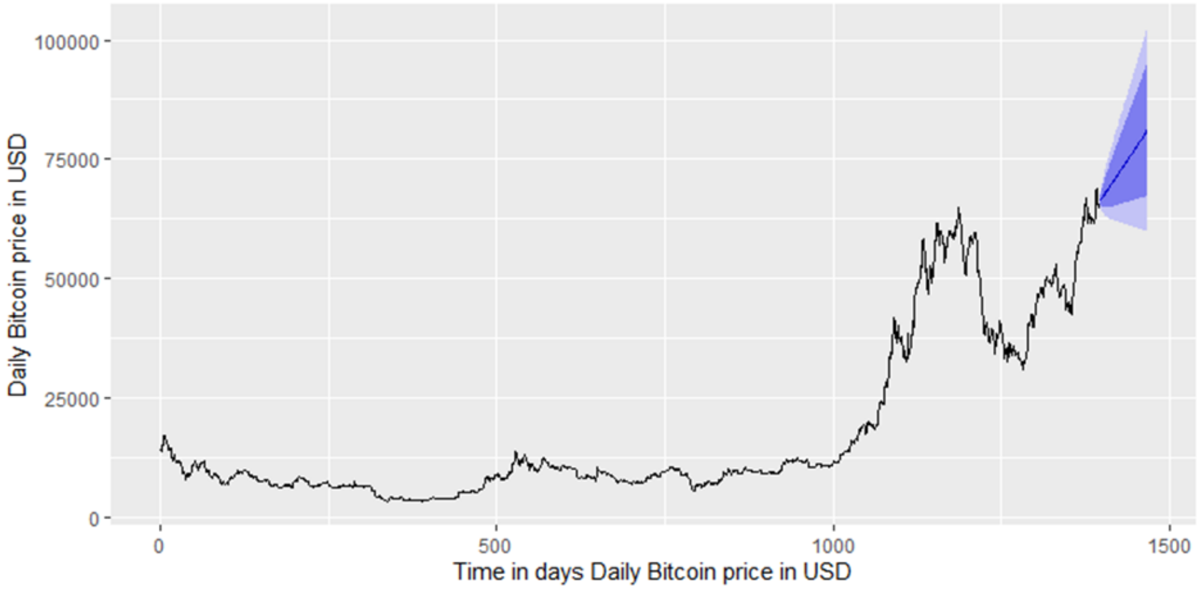
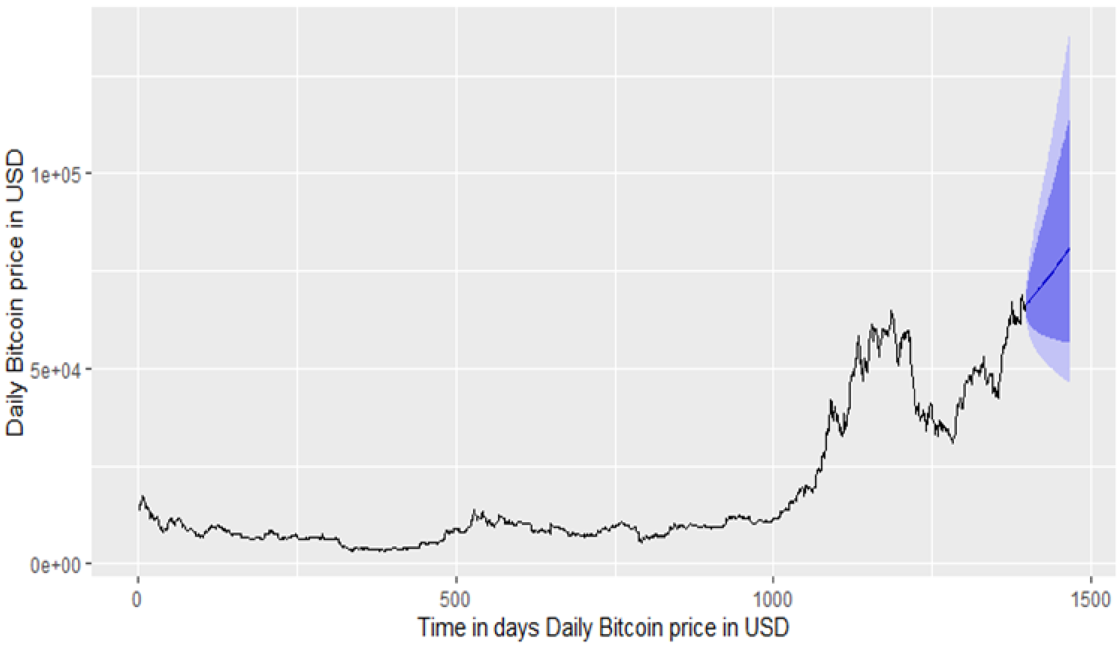
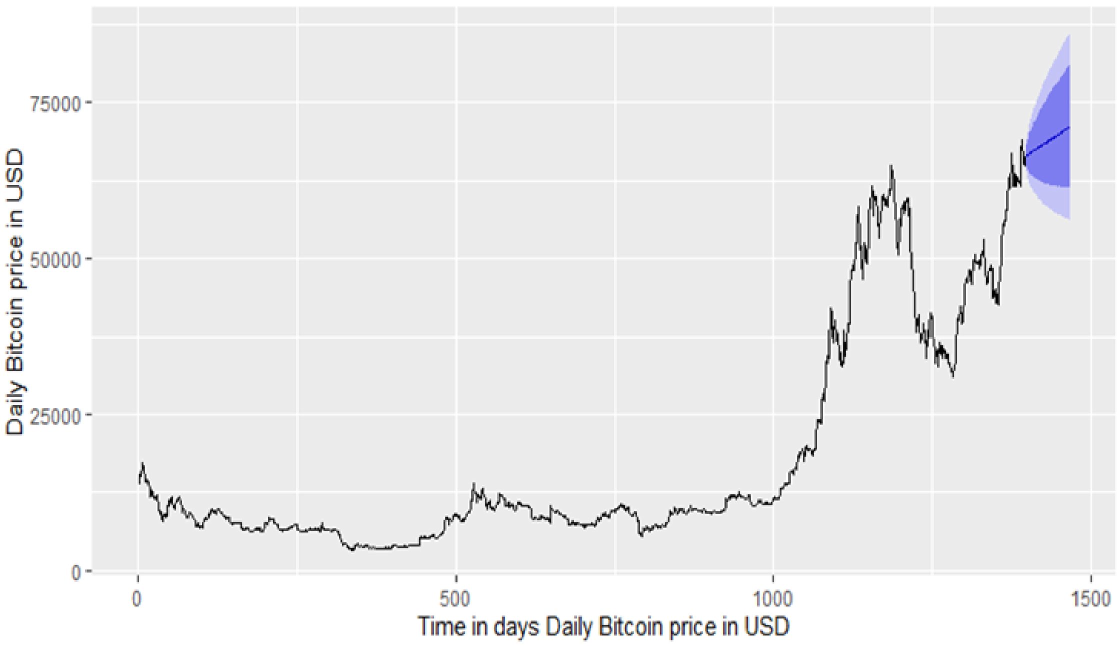


Figure 3. Time series versus predicted values.

Figure 4 presents the predictions obtained from the model of Holt and BATS [5]. Both of these models presented a poor performance compared to the two models above because they have higher MAPE, RMSE, ME and MAE values than them. They represent an upward trend in the price of Bitcoin in the coming days, accompanied by relatively wide confidence intervals, corresponding to a higher degree of uncertainty about the predictions.



(a)



(b)

Figure 4. Forecasting according to Holt (a) and BATS (b).

4. Conclusion

In this article we have evaluated four different models for predicting the Bitcoin price time series. Our findings highlighted the difference between the accuracy of each model's performance. Using several models allows testing and comparing the accuracy of their predictions and leads to an optimal choice. For our time series, the NARNN model is preferred over other forecasting models. It was selected based on MAPE values, as it had the smallest value among all predictive models. In addition, the NARNN model improved forecast accuracy by 39% compared to the ARIMA model, according to MAPE.

The NARNN model gave better results in almost all the indicators considered, with a significant difference from the indicators of the other models. We selected the NARNN model as the best model based on the MAPE value, considering the model with the lowest MAPE value as the most accurate model. The NARNN model has the minimum MAPE value for the considered period (7.47%). He predicted a downward trend in the price of Bitcoin in the coming days. Forecasts are valid for a short period of time, because for the long term they can be influenced by other external factors, such as inflation, economic crisis, etc. The above models can equally be applied to the data of future periods, for predictions of the future price of Bitcoin, in order to improve the accuracy of the predictions. Predictions with a high degree of reliability about the future price of Bitcoin would be very important for speculators, investors and financial actors in general, to make future projections and prevent possible losses.

Acknowledgement

This study was partly presented at the 6th Advanced Engineering Days [6].

Funding

This research received no external funding.

Author contributions

Katerina Zela: Conceptualization, Methodology, Software. **Lorena Saliaj:** Data curation, Writing-Original draft preparation, Software, Validation.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International journal of forecasting, 14(1), 35-62. https://doi.org/10.1016/S0169-2070(97)00044-7
2. Batista, M. (2020). Estimation of the final size of the COVID-19 epidemic. Preprint. medRxiv
3. Rodríguez Rivero, C., Pucheta, J., Laboret, S., Patiño, D. and Sauchelli, V. (2015) Forecasting Short Time Series with Missing Data by Means of Energy Associated to Series. Applied Mathematics, 6(09), 1611-1619. https://doi.org/10.4236/am.2015.69143
4. Ruiz, L. G. B., Cuéllar, M. P., Calvo-Flores, M. D., & Jiménez, M. D. C. P. (2016). An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. Energies, 9(9), 684. https://doi.org/10.3390/en9090684
5. Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. Future Computing and Informatics Journal, 2(1), 39-47. https://doi.org/10.1016/j.fcij.2017.05.001
6. Zela, K., & Saliaj, L. (2023). Neural networks for bitcoin price forecasting. Advanced Engineering Days (AED), 6, 156-158.