Visnyk of the National Bank of Ukraine, 2022, No. 254, pp. 29–42

CRYPTO CURRENCY PRICE FORECAST: NEURAL NETWORK PERSPECTIVES

YURIY KLEBAN[a], TETIANA STASIUK[a]

[a]National University of Ostroh Academy
E-mail: yuriy.kleban@oa.edu.ua
E-mail: tetiana.v.stasiuk@oa.edu.ua

Abstract This study examines the problem of modeling and forecasting the price dynamics of crypto currencies. We use machine-learning techniques to forecast the price of crypto currencies. The FB Prophet time-series model and the LSTM recurrent neural network were used to conduct the study. Using the example of data from Binance (the most popular exchange in Ukraine) for the period from 06.07.2020 to 01.04.2023, prices of Bitcoin, Ethereum, Ripple, and Dogecoin were modeled and forecasted. The recurrent neural network of long-term memory showed significantly better results in forecasting according to the RMSE, MAE, and MAPE criteria, compared to the results from the Naïve model, the traditional ARIMA model, and the FB Prophet.

JEL Codes C45, C53, G11

Keywords crypto currency, forecasting, time series, neural network

1. INTRODUCTION

The rapid development of digital currencies over the past decade is one of the most controversial and unpredictable innovations in the global economy to date. Significant fluctuations in the exchange rates of crypto currencies, the possibility of market perturbations due to false information and a lack of transparency, doubts over the legality of their use related to the anonymity of owners, as well as incomplete and contradictory legislation mean there are significant risks related to investing in crypto assets.

As for Ukraine, the discussion around the crypto currency market intensified with the adoption of the Law «On Virtual Assets» (2022) and the registration in November 2023 of draft Law No. 10225-1 on amendments to the Tax Code of Ukraine and other legislative acts of Ukraine regarding the regulation of the turnover of virtual assets in Ukraine (Verkhovna Rada of Ukraine, 2023). The most complex and most discussed issues of this draft law are related to tax conditions for individuals and businesses operating in the field of virtual assets (Malinovska, 2023). Crypto assets are becoming increasingly popular among economic agents as a form of investment asset (Corbet et al., 2019). Significant price volatility for crypto assets makes it necessary to develop and use models for forecasting prices effectively. The current economic literature actively applies various traditional statistical approaches and machine-learning methods (link) to assess the ability to forecast the prices of various kinds of digital currencies over a range of horizons.

In this study, we contribute to this large body of the literature, and analyze the effectiveness of machine-learning methods such as FB Prophet and LSTM in forecasting the price of crypto assets. The selected crypto currencies are Bitcoin and Ethereum, which have the largest capitalization, Ripple, a popular low-cost currency that is actively used by businesses around the world and has an affordable set of high-quality tools for managing financial resources, and Dogecoin, the so-called Meme-coin that applies to crypto assets, the value of those is determined mainly by community interests and online trends. Crypto currencies of a range of capitalizations and behaviors were chosen, as investors adjust their portfolio preferences depending on the market situation. It is also important to establish whether different forecasting methods should be used for crypto currencies of differing characteristics. The data is daily and was obtained from the service binance.com for the period 06.07.2020 to 01.04.2023.

The Naïve and ARIMA models demonstrated fairly high forecasting accuracy for crypto currencies with low prices that are affected most by market volatility (Ripple, Dogecoin). Prophet is best used for Bitcoin and Ethereum, the crypto currencies that effectively set the trends of the crypto currency market.

According to the results of the study, the recurrent neural network LSTM demonstrated the best forecast accuracy for all crypto currencies. Thus, the paper demonstrates that LSTM, despite the complexity of its use, is a powerful tool for

© National Bank of Ukraine, Y. Kleban, T. Stasiuk, 2022. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. Available at https://doi.org/10.26531/vnbu2022.254.03.

modeling volatile and complex phenomena such as crypto currency prices.

This rest of this paper has the following structure: The second section provides a brief overview of the main characteristics of the crypto currency market. The third section reviews the literature that studies the problems of forecasting the prices of crypto assets. The fourth section describes the study methodology. The fifth section is devoted to the data used. The sixth section presents the key results of the study and evaluates the quality of the constructed models and forecasts, their description and quality analysis. The final section, conclusions, summarizes the results of this study.

2. THE NATURE OF THE CRYPTO CURRENCY MARKET

As in the case of traditional financial assets, the operation of the crypto currency market is based on the principle of a balance between supply and demand: when demand (supply) for crypto currencies increases (decreases), prices usually increase, and vice versa. Supply and demand, in turn, are influenced by various price factors (price stability of the market, the exchange rate price for Bitcoin), and non-price factors (crypto currency issues, news, legal restrictions).

The crypto currency market, where crypto currencies are bought and sold, is decentralized. While similar to a traditional market, unlike traditional ones the crypto currency market is available for trading 24/7. There are several specific features of the way the market functions:

- Decentralization. The crypto currency market is decentralized, meaning it does not have a central body that controls and regulates its operation. First, this technology does not have a central issuing authority, such as a central bank in the case of traditional currencies. Thus, the lack of centralized control contributes to greater autonomy and distribution of ownership. Second, all contracts and transactions made on the crypto currency network are reflected in a blockchain, which is a distributed database. This means that each node in the network has a copy of this database, so no centralized authority can manipulate or control these transactions. The third important aspect is the development of decentralized exchanges, where crypto assets are traded without the mediation of centralized structures. This is effected through smart contracts that ensure the execution of transactions directly on the blockchain, without the need for the involvement of a third party. Thus, decentralization in the context of crypto currencies and the blockchain technology allows for greater autonomy, security and transparency in financial and trading operations, avoiding centralized control and risks related to it. Crypto asset trading often takes place on centralized exchanges, where participants exchange crypto currencies under established rules, using infrastructure that is usually owned by a centralized company or organization. Despite the advantages of centralized exchanges, such as high liquidity and fast transaction execution, there are risks related to centralization, including the possibility of system hacking, fraud, and access restrictions. The disadvantages of centralized exchanges have led to the emergence of decentralized alternatives, where transactions are made directly between users using smart contracts on the blockchain.

- High volatility. Crypto currencies are known for their high volatility. This means that their prices can change very quickly depending on various factors, such as news, industry events, etc.

- Algorithmic trading. Algorithmic trading is widely used in the crypto currency market. This means that many users use special applications that analyze the market and automatically buy or sell crypto currencies depending on various factors.

- Risks. The crypto currency market carries certain risks, including those related to price volatility, possible cyber-attacks and security hacks, as well as legislative changes.

The level of legalization of virtual assets in different countries differs, because the lack of knowledge of the problem, the high risks of these assets and other factors related to the internal development of states do not allow the full implementation and use of crypto currencies as financial instruments. Legislation on virtual assets in various countries regulates the activities of crypto currency market participants. That means market regulators and other public authorities may exercise certain controls over the operations and activities of market participants that perform operations with virtual assets, in order to reduce asset volatility and reduce risk.

Most countries around the world have recognized crypto currencies as virtual assets and legalized them at the legislative level (Amase, 2023). In Ukraine, the legislative framework for regulating the turnover of virtual assets is still at the stage of formation and discussion by society.

The formation of legislation to regulate the crypto currency market is a reaction by countries to the spread and impact of blockchain technologies in the modern world. The main reasons include combating illegal activities, protecting investors and consumers, ensuring financial stability, implementing an effective tax policy, and so on.

3. LITERATURE REVIEW

The main limitation when forecasting prices for crypto assets is their high volatility and the difficulty of determining the main factors influencing the exchange rates of crypto currencies. Because of their high risk, investments and other operations in crypto currency require reasonable risk management and balanced management strategies, since quite often the future financial stability of the investor or shareholder depends on them.

Usually, forecasting the price of crypto currencies is considered as a time series problem (Persson, 2022). Using time series, the forecast of future values is based on previous observations for consecutive time intervals.

An important concept in time series analysis is stationarity, when dynamic series have the same behavior and the same statistical properties over a time period (Whittle, 1953). However, it is worth noting that crypto currency prices are non-stationary (Couts et al., 1966).

Pronchakov and Bugaienko (2019) compared several types of moving averages (simple, weighted, and exponential) for forecasting the prices of digital currencies during their study. They concluded that all moving averages had approximately the same trend, but the exponential model was closest to actual values and adapted faster to price changes.

A separate place in the studies is occupied by complex methods of forecast extrapolation, among the most common varieties of which are moving average and exponential smoothing methods (Pronchakov and Bugaienko, 2019; Pilipchenko et al., 2021; Gagnidze and Iavich, 2020). They are commonly used for noise smoothing, identifying fracture points, and short-term forecasting. For example, the intersection of moving averages is an important technical indicator according to the authors: when the moving average for a short period intersects with the moving average for a long period, this is a signal to buy or sell an asset.

Derbentsev et al. (2019) used time-series models, Binary Auto Regressive Tree (BART) and Autoregressive Integrated Moving Average (ARIMA), to build short-term forecast models for the crypto currencies with the highest market capitalization. The time periods contained different types of dynamics (stable, decline, growth, trend change). The results showed that the errors in the BART model were half those compared to the ARIMA model. However, the authors note that all models showed worse results during periods of sharp changes in trends.

The artificial intelligence sub-sector is also often used to forecast the price of crypto currencies: i.e. machine learning. It is based on the use of statistical methods by which a computer acquires the ability to "learn" from a data set.

Popular methods include Long Short-Term Memory and models derived from it (Livieris et al,, 2020; Luo et al., 2022; Ammer and Aldhyani, 2022) and Gated Recurrent Unit (Al-Nefaie and Aldhyani, 2022; Aljadani, 2022)

Aljadani (2022) used a Bidirectional LSTM and a Gated Recurrent Unit in his research. For the Bitcoin price time series, the highest value of the average absolute percentage error was 0.26% for the BiLSTM model and 0.22% for the GRU, given that for the Mean Absolute Percentage Error or MAPE indicator, a value of less than 10% means that the forecast model is considered to have a high level of accuracy.

In the work of Al-Nefaie and Aldhyani (2022) the use of deep learning to forecast the value of Bitcoin was studied, with the aim of helping investors make informed decisions and aiding authorities in evaluating crypto currencies. The authors used a GRU (Gated Recurrent Unit, a type of recurrent neural network (RNN) that was designed to work with sequential data such as text or time series) and an MLP (Multilayer Perceptron, a type of artificial neural network that consists of multiple layers of neurons, where each neuron in the previous layer is connected to each neuron in the next layer, creating a deep multi-layer architecture) models to analyze Bitcoin price time series between January 2021 and June 2022. Based on the results of the study, it was found that the MLP model achieved high regression efficiency with R2 = 99.15% at the training stage and R2 = 98.90% at the testing stage. The authors believe that these models may have a significant impact on portfolio management and optimization in the face of the unpredictability of the crypto currency market. However, it is necessary to keep in mind the need to use such models cautiously in conditions of high volatility in the crypto currency market, and take into account possible limitations and uncertainties when considering their results.

Research by Garlapati et al. (2022) was conducted into forecasting the value of Bitcoin using Facebook's Prophet and ARIMA. The authors compare the effectiveness of these two methods on the same data set covering the period from May 2016 to March 2018. To improve the accuracy of forecasting, control variables selected on the basis of correlation studies between crypto currencies and real currencies are added to the model. According to the test results, Prophet is more effective than ARIMA, demonstrating an R2 value of 0.94, compared to 0.68 for ARIMA.

A study by Cheng et al. (2024) uses empirical financial time series analysis and machine learning to forecast the price of Bitcoin using LSTM, SARIMA, and Facebook's Prophet models. The results show that LTSM has a marked improvement over SARIMA and Prophet in terms of MSE and MAE. Furthermore, the result confirmed that Bitcoin values are seasonally volatile and random, and are often influenced by external variables such as news, crypto currency laws, investments, or social media rumors.

The approaches used by researchers to building crypto currency price forecasts have shown that they may be quite effective in forecasting for the short term. However, there is a need to combine and compare these methods to ensure the stability and reliability of forecasts.

4. DATA

Four crypto currencies were selected for building models and making forecasts: Bitcoin (BTC) and Ethereum (ETH) having the highest capitalization, Ripple (XRP) is cheaper, but quite developed and popular among businesses, and the Dogecoin (DOGE) meme-coin. This was done in order to help investors choose the right asset to invest in, as most initial investors prefer both already well-capitalized and new, cheap crypto assets.

Historical data for the selected assets against the Tether (USDT) stable coin (which, in fact, is a virtual dollar) was taken from the most popular exchange in Ukraine – Binance, using the Application Programming Interface (API) of the exchange and the Python programming language. The paper uses 1,000 observations of the prices of the following trading pairs: BTC-USDT, ETH-USDT, XRP-USDT, DOGE-USDT. The data was obtained on a daily basis for the period 06.07.2020 to 01.04.2023. The data set consists of five characteristics: open – opening price, close – closing price, high – maximum price, low – minimum price, volume – amount of money in circulation. Descriptive statistics of the collected trading pairs are shown in Table 1 (Appendix A), and Figure 1 (Appendix B) illustrates the closing price of each of the crypto currencies for the selected period.

Before starting to model a time series, it is important to understand some of its basic properties. Understanding the properties of a time series may help determine which methods and models will be effective for forecasting. For further modeling, the closing price is selected as the main dependent variable.

The first step is to normalize the data. One of the most common methods of data normalization is min-max normalization: for each attribute, the smallest value is replaced with 0, the largest value is replaced with 1, and all other values are in the range from 0 to 1. The values are calculated using the following formula:

$$X_0 = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)}, \tag{1}$$

where *min*(X_i) and *max*(X_i) are the minimum and maximum values, 0 and 1, respectively.

An additional step in data preprocessing was checking for anomalies and outliers. In order to identify values that have statistical differences, we use an uncontrolled learning technique, namely neural networks (NNs). We used LSTM RNN and autoencoders to build a model of unsupervised learning.

Figure 2 (Appendix B) shows the distribution of the Mean Absolute Error (MAE) in the training and test datasets. In the training set, values greater than 0.2 are seen as unusual. This value was thus set as a threshold for outliers, i.e. values exceeding it will be outliers.

Figure 3 (Appendix B), we can see that there are abnormal values in the Bitcoin training sample. A total of 10 anomalies were detected in the training sample, while there were zero in the test sample. Similar steps have been applied to other crypto currencies. For clarity, the outliers were shown on graphs (Figure 4 in Appendix B).

Thus, Ethereum has 22 outliers, Ripple – 25, and Dogecoin – 20. No outliers were recorded in any of the test samples. For more accurate forecast models, anomalous values were smoothed out using a simple moving average (SMA).

Then, we check the time series for stationarity using the Augmented Dickey-Fuller Test (ADF), since for certain modeling methods, the stationarity of the data is a necessary condition. Table 2 shows the test results.

From the Table above, we can see that the selected price series for crypto currencies are characterized by non-stationarity. The data is reduced to a stationary form using a difference operation with different integration coefficients. The results of the data stationarity test after integration are shown in Table 3.

5. METHODOLOGY

To study, model, and forecast time series (for prices), we start with a naive method and use its results as a baseline for subsequent models, as this is primitive and the goal is to improve the accuracy of forecasts for other models. We make a naive forecast based on the last value in the data set, that is, we assume that the price of crypto currency tomorrow will be the same as today. The obtained forecasts based on the naive method shall be further compared with the results of more complex models in order to evaluate their effectiveness.

At the next stage, ARIMA models were built (Box et al., 2015) for the stationary time series of the prices for each of the crypto currencies. ARIMA is used to analyze and forecast series that may have a trend and/or seasonality. The main stages of build-up include data standardization, ensuring the stationarity of the time series. Then the parameters of the ARIMA model are determined, such as the autoregression order (p), the integrability order (d), and the moving average order (q). This is done on the basis of an analysis of autocorrelation and partial autocorrelation graphs. After that, the best model is selected and parameters are selected using criteria such as AIC or BIC. After the selection of the parameters, an ARIMA model is built and evaluated on the training dataset. Then, diagnostics and the verification of residual errors are performed to determine their randomness and the possibility of improving the model. Finally, the constructed model is used to forecast future values of the time series. This process may require iterations and parameter adjustments to achieve an optimal model for certain time series.

After building and evaluating ARIMA models, a forecast based on stationary and non-stationary data was made for the Facebook Prophet model, which is a both simple and powerful tool from Meta (Taylor and Letham, 2017). It can be used to forecast a time series without the user having to be an expert in data analysis. It was created because working with neural networks often used in forecasting, and this is quite difficult without having proper knowledge about the architecture of these networks.

Prophet's main algorithm is a generalized additive model, which can be decomposed into three main components: trend, seasonality, and public holidays. As mentioned above, seasonality and trend are two important but difficult-to-quantify components of a time series analysis, but Prophet takes both into account perfectly.

Because the model can be decomposed, it is relatively easy to obtain model coefficients to understand the impact of seasonality, trends, holidays, and other variables. For example, by forecasting the price of a crypto currency, we can obtain a demand ratio to find out how much demand affects price changes.

Prophet also contains a built-in cross-validation function for measuring forecast error using historical data. This is done by selecting boundary points in the history, and for each of them, a model is built using data only up to this boundary point. Then we can compare the forecasted values with the actual ones.

Table 2. Checking Time Series for Stationarity

	Bitcoin	Ethereum	Ripple	Dogecoin
ADF	-1.57	-1.77	-2.31	-2.46
p-value	0.49	0.39	0.16	0.12
Stationary	–	–	–	–

Note: calculated by the authors based on the collected data.

Table 3. Reducing the Time Series to Stationary Ones

	Bitcoin	Ethereum	Ripple	Dogecoin
ADF	-31.05	-11.31	-6.69	-5.76
p-value	0.00	0.00	0.00	0.00
Stationary	+	+	+	+

Note: calculated by the authors based on the collected data.

It is worth noting that Prophet is great for stationary data, i.e. time series that have the same behavior and the same statistical properties over a time period (Sivaramakrishnan et al., 2021). Prophet may also be used for non-stationary time series, but it does not always allow accurate forecasts to be made (Vasselin and Bertrand, 2021). In our study, the accuracy of the Prophet model was quite low for both stationary and non-stationary data.

Deep learning methods are also used to study time series. They are used to create multi-layered neural networks that ensure the high accuracy of results.

Recurrent Neural Networks (RNNs) are a popular class of neural networks that allows cyclic connections between nodes to be created. That is, the output of a node can affect subsequent inputs to the same node (Figure 5 in Appendix B). This allows them to use their internal state to process sequences of input data. By the same principle, a person reading a book understands each word by relying on knowledge gained earlier.

Nodes at different levels of the neural network are compressed and form a single layer of the recurrent neural network. At any given time t, the input data is a combination of the input data x<t> and x<t-1>, and the output data is returned to the network to improve the results.

Standard RNNs have some problems when exploring long-term dependencies and remembering them for a long time. These problems do not occur when using a Long Short-Term Memory (LSTM) – this is their typical behavior.

All RNNs have the same shape, i.e. a chain of repeating neural network modules. Regular models will have a simple structure with a single tanh (hyperbolic tangent) level (Feng et al., 2019), the LSTM, in turn, will have four interacting levels with unusual coupling (Figure 6 in Appendix B). Technical independent variables (lag values of the opening price, maximum and minimum price, and transaction volumes) are also used to build the LSTM model, which improves the forecasting ability of the model. At the same time, macroeconomic variables were not used as variables due to their low frequency and, consequently, their irrelevance in forecasting daily data.

To avoid overtraining with the LSTM network, the training sample was pre-tested through cross-validation.

The training sample was divided into five parts (as shown in Figure 7 in Appendix B), where green indicates the elements used for training the model, and blue indicates the elements used for testing. The metric for evaluating the model was set to Mean Absolute Percentage Error (MAPE).

For example, for the Bitcoin price, the mean error for the five evaluation stages was 2.80%, with a standard deviation of 0.82%. All models gave approximately the same results, which indicates the high stability of the model.

It is worth noting that models based on non-stationary data (Facebook's Prophet, LSTM) forecast the closing price of a candlestick (from a Candlestick Chart visual representation of trading data) for the day, while models based on stationary data (Naïve, ARIMA) forecast price movement between two candlesticks. Because of this, there are problems when converting results to their previous form as closing prices. In other words, models based on stationary series showed more accurate results for the average daily price, and forecasted changes in the movement of candlesticks, but these models could not forecast the closing price, so non-stationary data was used to compare the forecast strength of Prophet and LSTM, which does not violate the requirements for these models.

There are quite a few metrics for estimating the accuracy of mathematical models (Plevris et al., 2022). It is on the basis of such metrics that the best models are selected and their further use or implementation in production processes. Most of them may be divided into two categories based on the types of forecasts in ML models: classification and regression. Since the problem we are considering is a regression, we use popular indicators to evaluate forecast models, namely: RMSE, MAE, and MAPE.

MAE and RMSE are used together to diagnose changes in errors in the forecast set. RMSE will always be greater than or equal to MAE: the greater the difference between them, the greater the variance of the errors in the sample. RMSE and MAE measure the error in units of measurement of forecasted variables, while MAPE displays the result as a percentage, determines the accuracy of the forecast model, and allows it to be determined how accurate the forecasted values were compared to the actual ones.

In general, the machine-learning methods chosen have certain advantages over other models. They have the ability to analyze large amounts of data and make forecasts based on them. The algorithms built into them can significantly improve the accuracy of forecasting of values, and identify complex dependencies and their degree of impact on the resulting variable – in our case the closing price.

6. STUDY RESULTS

During modeling, the collected time series were divided into training and test samples in a 90/10 ratio (900 observations for training models and 100 for test ones). The large amount of training data gives the selected models a wider horizon for studying time-series patterns. Furthermore, given that cross-validation is applied to the selected methods and cross-validation of models occurs during their training, the error is averaged after each iteration of the training, and we obtain a more reliable estimate of the error and the accuracy of the model, respectively. It is worth noting that only the data of the training sample is involved in training the model using cross-validation, while the test sample is not accepted in this process and is used at the stage of checking the model for quality.

A basic forecast for the future was made with the assumption that the price of crypto currencies tomorrow will not differ from yesterday (Figure 8 in Appendix B). Figure 9 (Appendix B) shows the results of forecasting the price of crypto assets based on ARIMA.

The results of Prophet forecasting are shown below. The blue curve shows the real price, and the red curve shows the simulated price. The light-blue field is the 95% confidence interval. Figure 10 shows that Prophet performed well when forecasting with only historical crypto currency data.

During modeling, the settings of the model parameters were changed for each of the different crypto currencies. For example, for Ethereum, the influence of the trend and the number of points of change were increased. As a result, MAPE was significantly reduced, and a rather accurate forecast was obtained – although confidence boundaries widened. For Dogecoin, in contrast, the influence of the trend was weakened. As a result, the forecast gives good results, taking into account the abnormal values.

Increasing the change points makes the model more adaptive to dynamics – especially if there are significant price changes in different periods. The parameter that determines the impact of a trend characterizes how much the model will pay attention to possible changes in the trend and seasonality. In other words, this parameter is responsible for regulating the flexibility in finding change points in trend and seasonality.

Let's look at the results of long-term memory network modeling in the test sample (Figure 11 in Appendix B). The information is displayed only for the test sample, because due to the low error in the training sample, the lines are superimposed, which makes it difficult to visually analyze the series.

The errors of the LSTM model during cross-validation on the training and test data are almost identical, which indicates the stability of the model on the sets that were used in its building, and on the new data for the forecast. Although it is quite easy to visually trace the convergence of actual and modeled values, the graph still shows some lag between the real values of the crypto currency price and the simulated ones.

Let's compare the built models in more detail using different metrics. Tables 4 to 7 below show the relative percentage deviation of the RMSE, MAE, and MAPE values of the ARIMA, Facebook's Prophet, and LSTM models from the Naïve model values. The least successful model compared to the Naïve model is highlighted in red, and the most successful model is highlighted in green.

For the Bitcoin crypto currency, the forecasts based on the Naïve model and ARIMA had relatively similar results and had satisfactory accuracy indicators. LSTM proved to be the best model for forecasting, showing the lowest errors and the highest accuracy among the methods studied.

For Ethereum, Naïve and ARIMA showed better results than for Bitcoin, but as in the previous example, LSTM has significantly lower errors and higher accuracy in its forecast values.

The Ripple crypto currency has slightly different results. The worst model was Prophet, and the best one was LSTM. An interesting observation is how well the Naïve and ARIMA models performed. Both models have high accuracy rates, and their forecast errors are 2–3 times lower than for Prophet.

Table 4. Choosing the Forecast Method for Bitcoin

	RMSE	MAE	MAPE
ARIMA(1,2,1)	-13.2	-13.4	-13.4
FB Prophet	-48.8	-54.6	-48.8
LSTM	-94.1	-95.0	-94.7

Note: calculated by the authors based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 5. Choosing the Forecast Method for Ethereum

	RMSE	MAE	MAPE
ARIMA(2,2,3)	-15.0	-14.9	-15.0
FB Prophet	-36.3	-44.3	-43.0
LSTM	-85.0	-84.8	-83.5

Note: calculated by the author based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 6. Choosing the Forecast Method for Ripple

	RMSE	MAE	MAPE
ARIMA(3,2,3)	-16.7	-25.0	-20.7
FB Prophet	166.7	250.0	236.0
LSTM	-81.7	-75.5	-77.0

Note: calculated by the author based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 7. Choosing the Forecast Method for Dogecoin

	RMSE	MAE	MAPE
ARIMA (2,2,3)	-11.1	-14.3	-2.8
FB Prophet	33.3	57.1	57.4
LSTM	-64.4	-58.6	-57.5

Note: calculated by the authors based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Dogecoin has similar results to the Ripple currency. As in the case of Dogecoin, Prophet did the worst job with forecasting, and LSTM did the best.

When building models, various parameters and data sets were used (without and with smoothed anomalies), it turned out that for all crypto currencies forecasts based on smoothed data had a higher error.

From Tables 4 to 7, we may see that LSTM demonstrated the best results among the studied models. In half of the cases, Prophet showed the worst forecast results, while MAPE for Naïve and ARIMA did not exceed 11%.

7. CONCLUSIONS

This study examines the problems of forecasting the price dynamics of crypto currencies. The spread and impact of crypto currency technologies in the modern world is causing heated discussions about the place and role of crypto currencies in the modern economy. Research into methods of forecasting crypto currency prices is of great importance for the scientific community, financial analysts, investors, and traders.

In the course of the study, data was collected, cleaned, normalized, and selected as the resulting basis for forecasting the closing price of crypto currencies. When studying the time series data, it turned out that it is non-stationary, which limits the range of possible approaches for modeling. In order to use ARIMA, the data was transformed into a stationary time series, but experiments have shown that this, firstly, complicates the process of calculating the resulting variable closing price, and secondly, does not improve the accuracy of models. The final mathematical models selected as the best ones are built using machine-learning techniques and using non-stationary time-series prices.

The recurrent neural network of long-term memory showed significantly better results in forecasting, according to the calculated errors, compared to the naive forecast, and for all ARIMA models, as well as the results of Facebook's Prophet. It is worth noting that in half of the cases, even Naïve and ARIMA showed more accurate results than Prophet.

Modeling and forecasting the price of crypto currencies is a quite promising and still under examined area of scientific study. To improve the process of forecasting crypto assets in the future, it is necessary to take into account fundamental factors (news, events in the field of technology, regulation, etc.), and to study relationships with other financial markets and economic trends.

Mathematical models for forecasting crypto currency prices have already become the foundation for developing trading algorithms and bots based on them, portfolio management tools, and for budget planning. With the development of modeling methods themselves and the growth of computing power, the accuracy of forecasting in the short term, even in very volatile markets, will increase.

REFERENCES

Aljadani, A. (2022). DLCP^{2}F: a DL-based cryptocurrency price prediction framework. Discover Artificial Intelligence, 2(20). https://doi.org/10.1007/s44163-022-00036-2

Al-Nefaie, A. H., Aldhyani, T. H. H. (2022). Bitcoin price forecasting and trading: data analytics approaches. Electronics, 11(24), 4088, 2–18. https://doi.org/10.3390/electronics11244088

Amase, W. (2023). Countries where cryptocurrency is legal vs illegal. CoinGecko, 11 December 2023. Retrieved from https://www.coingecko.com/research/publications/crypto-legal-countries

Ammer, M. A., Aldhyani, T. H. H. (2022). Deep learning algorithms to predict cryptocurrency fluctuation prices: increasing investment awareness. Electronics, 11(15), 2349. https://doi.org/10.3390/electronics11152349

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control, 5th ed. New Jersey: Wiley.

Cheng, J., Tiwari, S., Khaled, D., Mahendru, M., Shahzad, U. (2024) Forecasting Bitcoin prices using artificial intelligence: combination of ML, SARIMA, and Facebook Prophet models. Technological Forecasting and Social Change, 198, 122938. https://doi.org/10.1016/j.techfore.2023.122938

Corbet, S., Lucey, B., Urquhart, A., Yarovays, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. International Review of Financial Analysis, 62, 182–199. https://doi.org/10.1016/j.irfa.2018.09.003

Couts, D., Grether, D., Nerlove, M. (1966). Forecasting non-stationary economic time series. Management Science, 13(1), 1–21. http://www.jstor.org/stable/2627909

Derbentsev, V., Datsenko, N., Stepanenko, O., Bezkorovainyi, V. (2019). Forecasting cryptocurrency prices time series using machine learning approach. SHS Web of Conferences, 65, 02001. https://doi.org/10.1051/shsconf/20196502001

Feng, J., He, X., Teng, Q., Ren, C., Chen, H., Li, Y. (2019). Reconstruction of porous media from extremely limited information using conditional generative adversarial networks. Physical Review E, 100, 033308. https://doi.org/10.1103/PhysRevE.100.033308

Persson, E. (2022). Forecasting Efficiency in Cryptocurrency. Markets A Machine Learning Case Study. Retrieved from https://www.diva-portal.org/smash/get/diva2:1711510/FULLTEXT01.pdf

Gagnidze, A., Iavich, M. (2021). Time-series forecasting of Bitcoin prices. East European University Collection of Scientific Papers, 2(2020), 46–53. https://orcid.org/0000-0003-4728-3709

Garlapati, A., Krishna, D. R., Garlapati, K., m. Srikara Yaswanth, N., Rahul, U., Narayanan, G. (2021). Stock price prediction using Facebook Prophet and Arima models. 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021. https://doi.org/10.1109/I2CT51068.2021.9418057

Lindemann, B., Maschler, B., Sahlab, N., Weyrich, M (2021). A survey on anomaly detection for technical systems using LSTM networks. Computers in Industry, 131, 103498. https://doi.org/10.1016/j.compind.2021.103498

Livieris, I. E., Pintelas, E., Stavroyiannis, S., Pintelas, P. (2020). Ensemble deep learning models for forecasting cryptocurrency time-series. Algorithms, 13, 121. https://doi.org/10.3390/a13050121

Luo, C., Pan, L., Chen, B., Xu, H. (2022) Bitcoin price forecasting: an integrated approach using hybrid LSTM-ELS models. Mathematical Problems in Engineering, 2022, 2126518. https://doi.org/10.1155/2022/2126518

Malinovska, A. (2023). Crypto currency tax: percentage, who should pay and features (In Ukrainian). Fakty, 1 December 2023. Retrieved from https://fakty.com.ua/ua/ukraine/ekonomika/20231201-podatok-na-kryptovalyutu-vidsotok-hto-maye-splachuvaty-ta-osoblyvosti/

Pilipchenko, A., Kuzminsky, V., Chumachenko, O. (2021). Using methods of technical analysis to forecast the cryptocurrency market. Včenì zapiski unìversitetu “KROK”, 4(64), 28–35. https://doi.org/10.31732/2663-2209-2021-64-28-35

Plevris, V., Solorzano, G., Bakas, N., Ben Seghier, M. (2022). Investigation of performance metrics in regression analysis and machine learning-based prediction models. ECCOMAS Congress 2022, 5, 9 June 2022. Oslo. https://doi.org/10.23967/eccomas.2022.155

Pronchakov, Y., Bugaienko, O. (2019). Methods of forecasting the prices of cryptocurrency on the financial markets. Technology Transfer: Innovative Solutions In Social Sciences and Humanities, 13–16. https://doi.org/10.21303/2613-5647.2019.00927

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D: Nonlinear Phenomena, 404, 132306. https://doi.org/10.1016/j.physd.2019.132306

Sivaramakrishnan, S., Fernandez, T. R., Babukarthikm R. G., Premalatha, S. (2021). 4 forecasting time series data using ARIMA and Facebook Prophet models. *In* Big Data Management in Sensing: Applications in AI and IoT, 47–60. River Publishers. Retrieved from http://ieeexplore.ieee.org/document/9606779

Taylor, S. J., Letham, B. (2017). Forecasting at scale. PeerJ Preprints, 5, e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2

Vasselin, H., Bertrand, J. (2021). Is Facebook Prophet suited for doing good predictions in a real-world project? Artefact. Retrieved from https://www.artefact.com/blog/is-facebook-prophet-suited-for-doing-good-predictions-in-a-real-world-project/

Verkhovna Rada of Ukraine (2023). The Draft Law of Ukraine on Amendments to the Tax Code of Ukraine and Other Legislative Acts of Ukraine Concerning the Regulation of Virtual Asset Turnover in Ukraine No. 10225-1 dated 17.11.2023 (In Ukrainian). Retrieved from https://itd.rada.gov.ua/billInfo/Bills/Card/43232

Verkhovna Rada of Ukraine (2024). Law of Ukraine “On Virtual Assets” No 2074-IX Edition of 01.01.2024 (did not enter into force) (In Ukrainian). Retrieved from https://zakon.rada.gov.ua/laws/show/2074-20

Whittle, P. (1953). Estimation and information in stationary time series. Arkiv för Matematik, 2, 423–434. https://doi.org/10.1007/BF02590998

APPENDIX A. TABLES

Table 1. Descriptive Statistics of Collected Trading Pairs

Dataset	Column	count	mean	std	min	25%	50%	75%	max
Bitcoin USD (BTC-USDT)	Open	1000	32,116.56	15,386.39	9,069.41	19,307.39	30,306.58	44,405.35	67,525.82
	High	1000	32,990.02	15,828.64	9,145.24	19,632.81	31,394.45	45,804.72	69,000.00
	Low	1000	31,141.11	14,848.86	8,893.03	18,908.93	29,288.29	43,017.27	66,222.40
	Close	1000	32,136.01	15,369.77	9,069.41	19,314.61	30,306.59	44,404.55	67,525.83
	Volume	1000	116,894.17	109,911.34	15,805.45	47,413.36	73,239.47	146,121.37	760,705.36
Ethereum USD (ETH-USDT)	Open	1000	1,956.26	1,136.44	227.54	1,217.71	1,713.87	2,813.22	4,807.98
	High	1000	2,022.10	1,170.42	229.85	1,259.99	1,777.80	2,946.53	4,868.00
	Low	1000	1,881.64	1,094.06	223.05	1,185.72	1,659.09	2,720.82	4,713.89
	Close	1000	1,957.86	1,135.16	227.56	1,218.31	1,715.80	2,813.22	4,807.98
	Volume	1000	795,863.90	526,527.70	117,762.10	448,910.60	656,849.10	970,714.00	4,309,836.00
Ripple USD (XRP-USDT)	Open	1000	0.59	0.33	0.18	0.35	0.47	0.80	1.83
	High	1000	0.62	0.35	0.18	0.36	0.49	0.83	1.97
	Low	1000	0.57	0.31	0.17	0.34	0.45	0.77	1.65
	Close	1000	0.59	0.33	0.18	0.35	0.47	0.80	1.84
	Volume	1000	55,8190,600.00	627,478,500.00	59,622,710.00	242,894,300.00	365,348,400.00	612,346,800,00	8,608,358,000.00
Dogecoin USD (DOGE-USDT)	Open	1000	0.12	0.11	0.002	0.06	0.08	0.17	0.69
	High	1000	0.13	0.13	0.002	0.06	0.09	0.18	0.74
	Low	1000	0.11	0.10	0.002	0.05	0.08	0.17	0.60
	Close	1000	0.12	0.11	0.002	0.06	0.08	0.17	0.69
	Volume	1000	2,690,987,000.00	6,872,100,000.00	88,706,470.00	653,910,800.00	1,082,091,000.00	2,006,851,000.00	109,073,700,000.00

Note: calculated by the author based on the data.

APPENDIX B. FIGURES

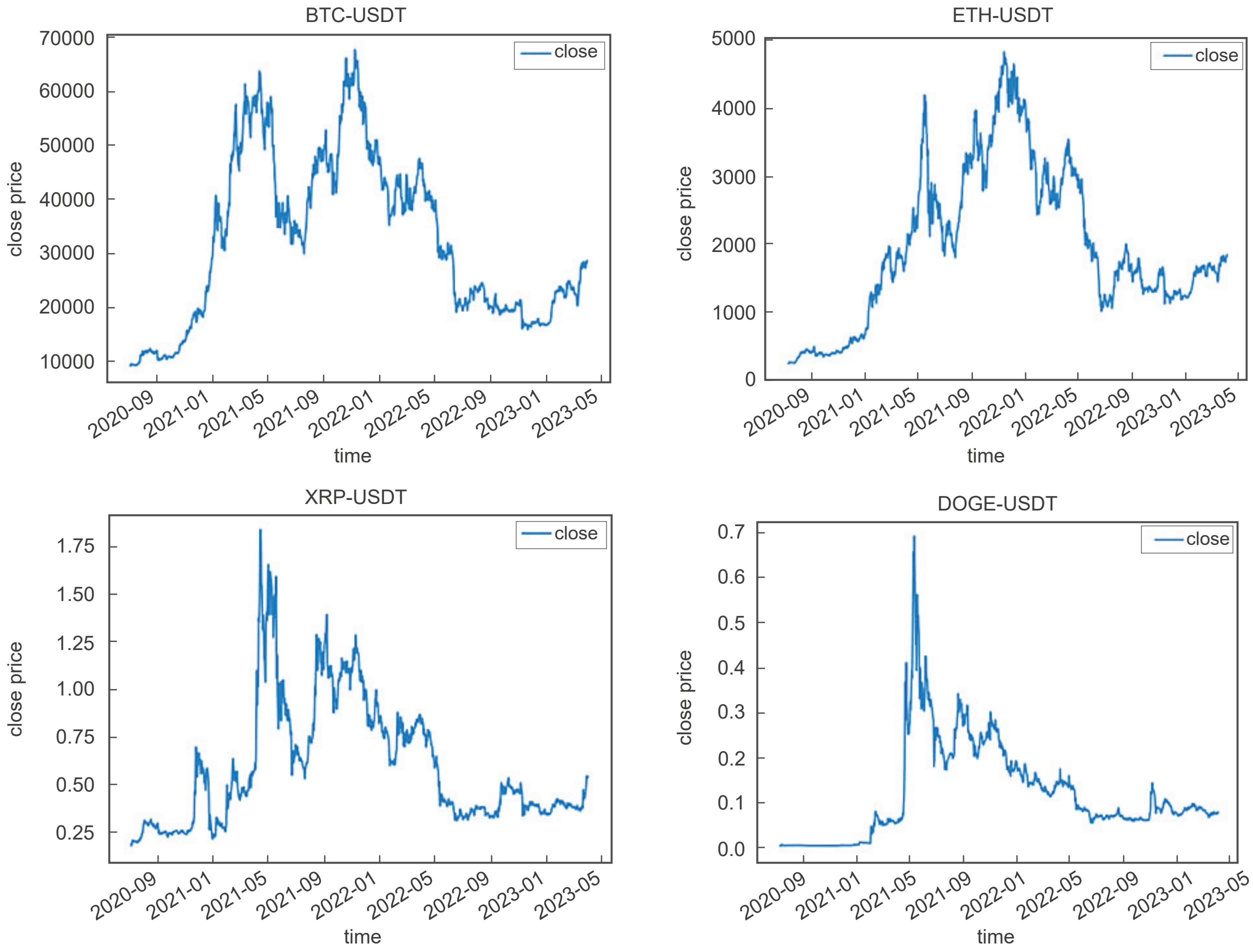

Figure 1. Dynamics of World Prices of Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the author based on the collected data.

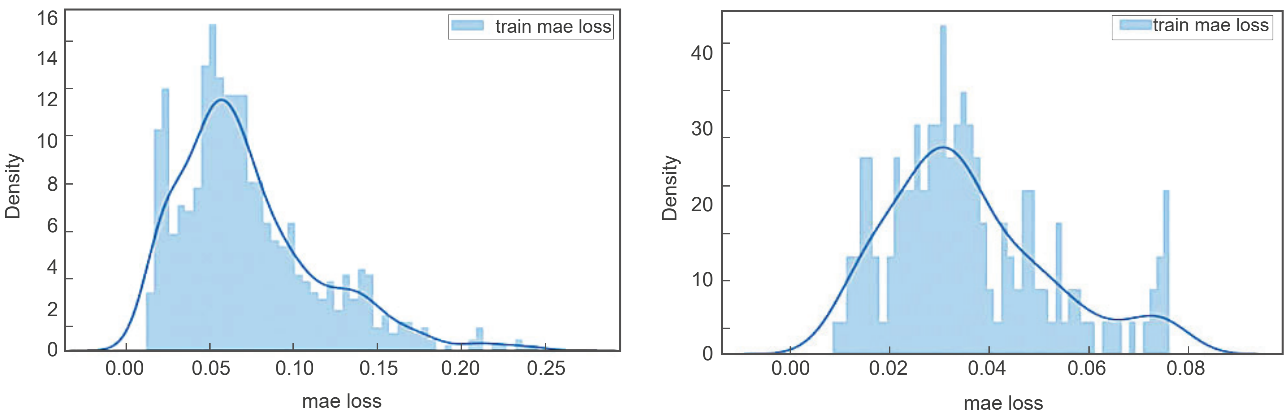

Figure 2. Distribution of the Mean Absolute Error in Bitcoin Samples

Note: built by the author based on the calculated errors.

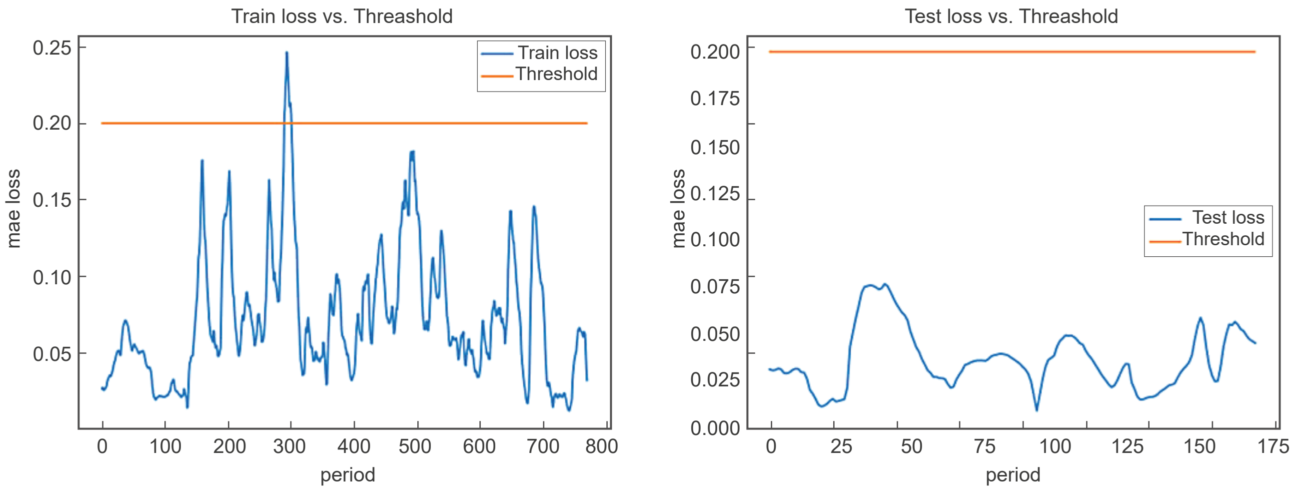

Figure 3. MAE and Bitcoin Sampling Anomaly Threshold

Note: built by the author basis on the calculated errors and the threshold value of the anomaly (Lindemann, 2021).

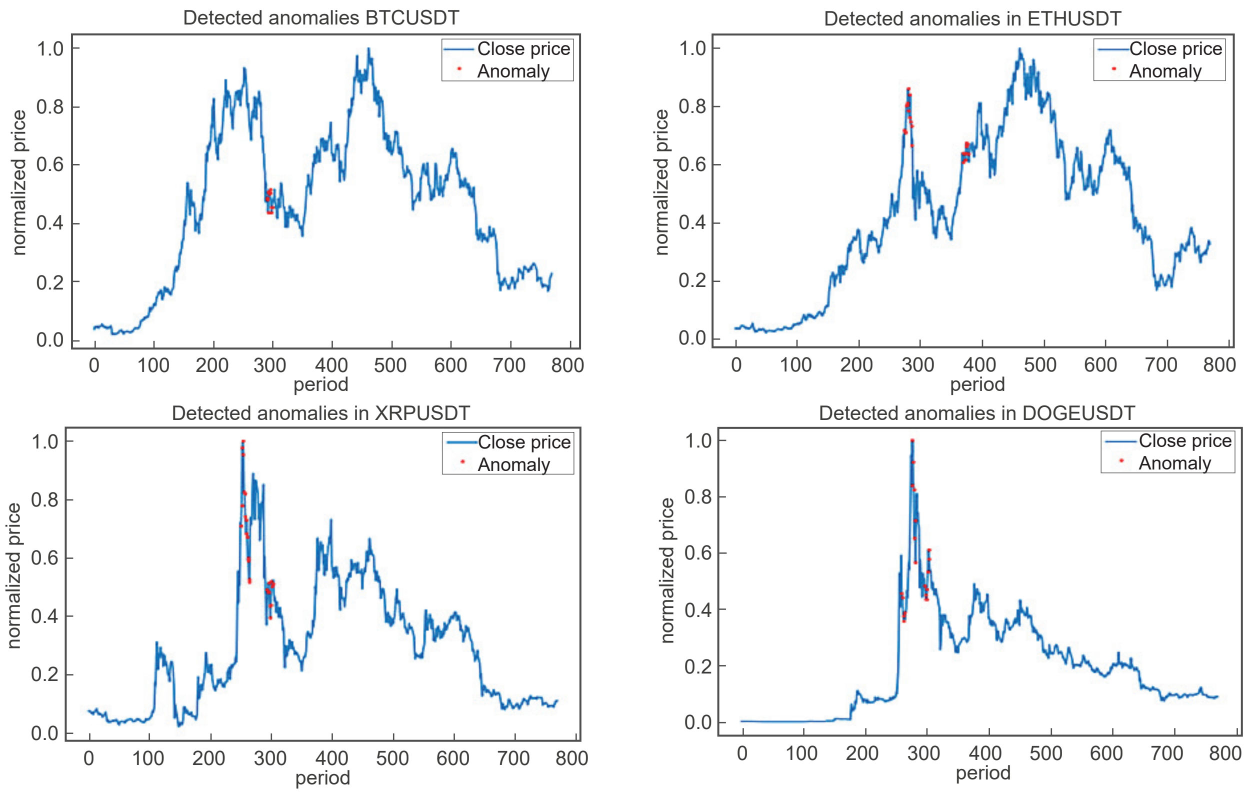

Figure 4. Abnormal Values in the Dynamics of Crypto Currency Prices

Note: built by the authors based on the detected anomalies.

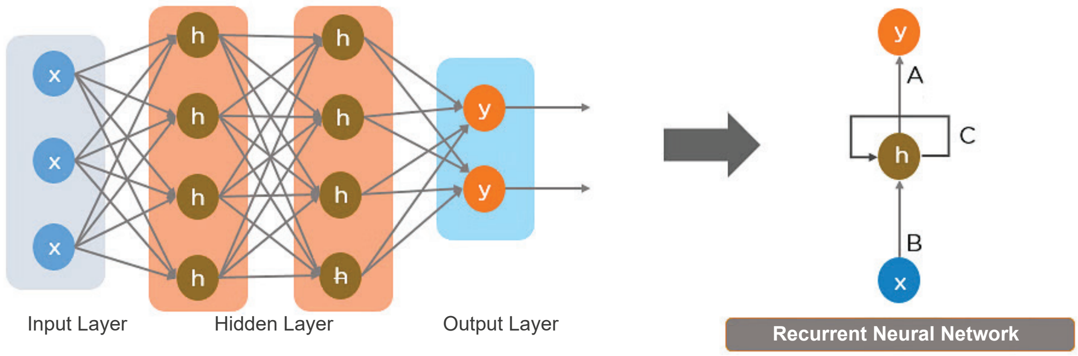

Figure 5. Structure of Simple RNN (Sherstinsky, 2020)

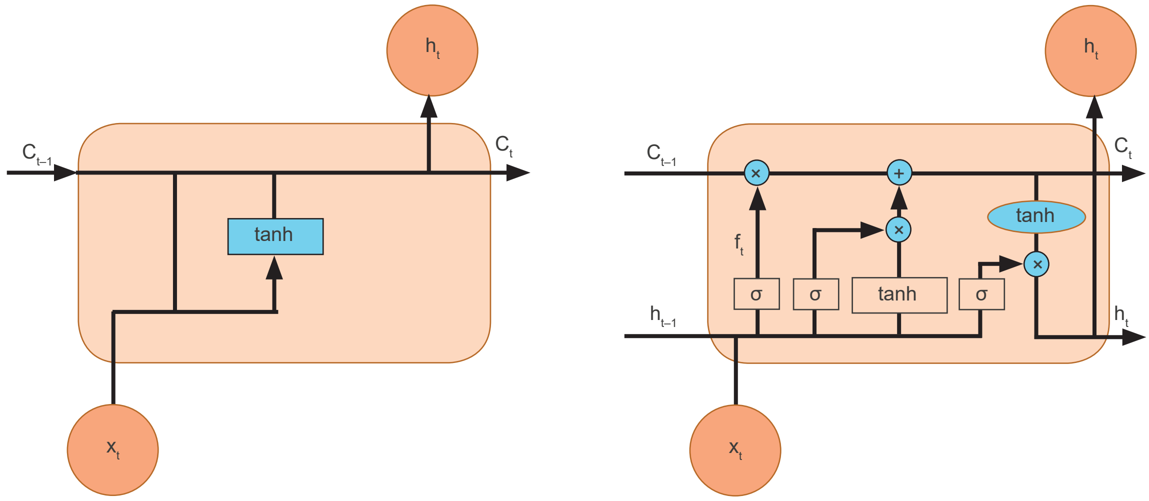

Figure 6. Module Structure for Regular RNN and LSTM Networks (Sherstinsky, 2020)

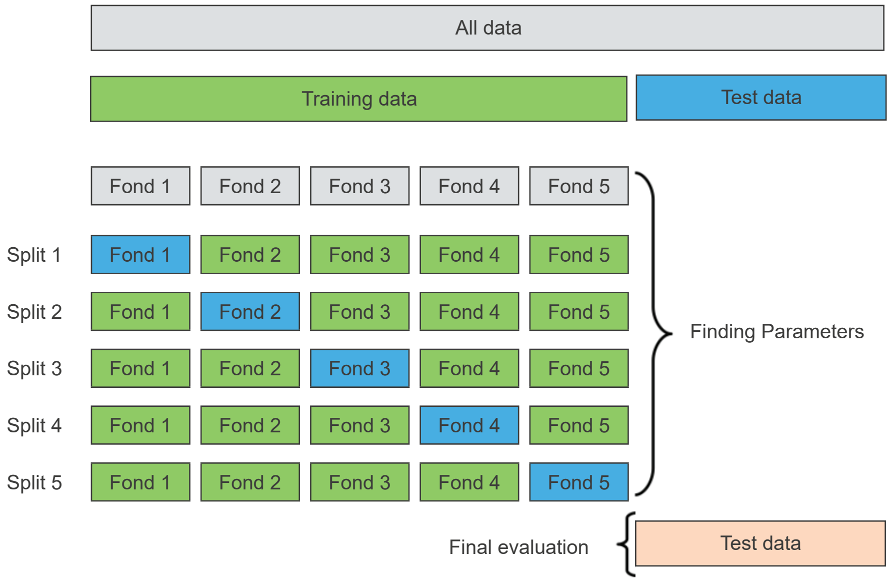

Figure 7. Cross-Validation Scheme *k-fold*[1]

[1] scikit-learn. Cross-validation: evaluating estimator performance https://scikit-learn.org/stable/modules/cross_validation.html#computing-cross-validated-metrics

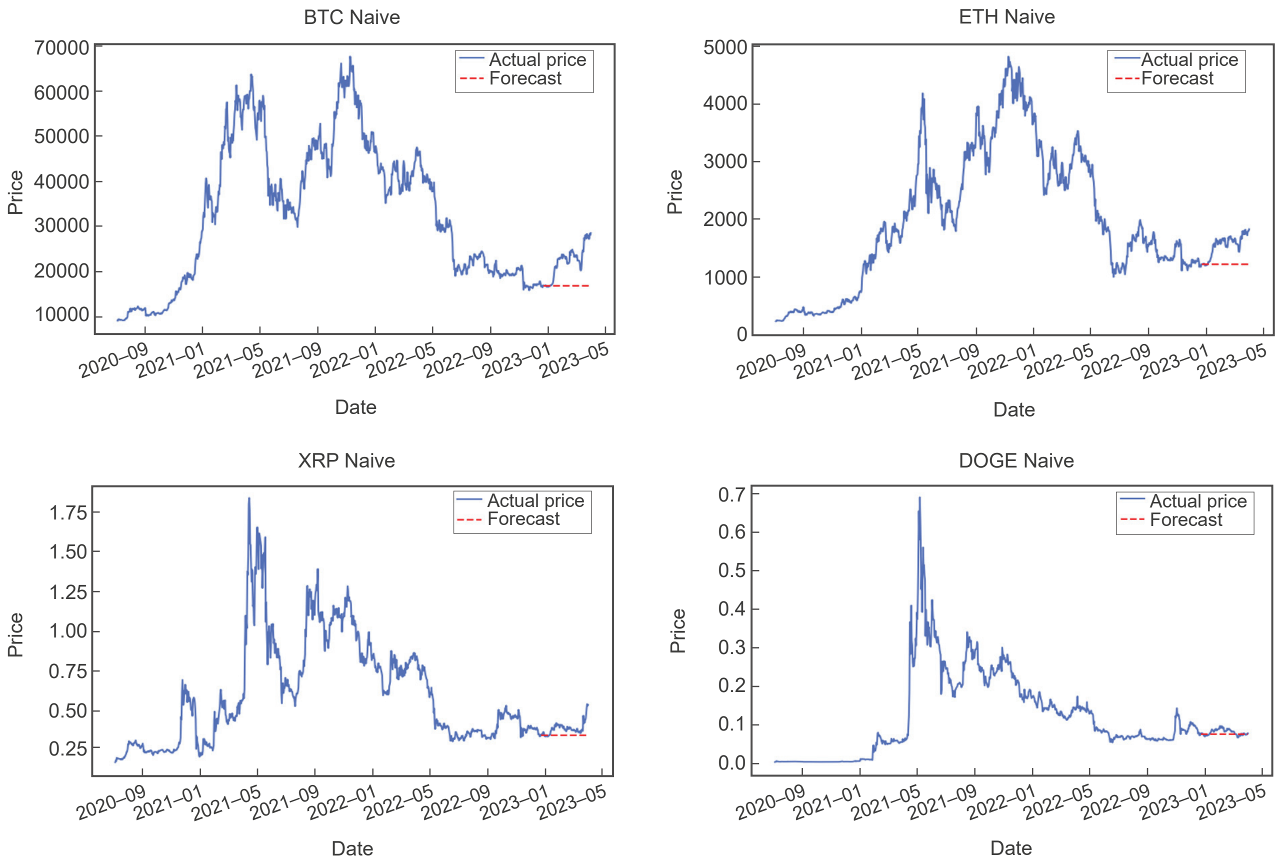

Figure 8. Naïve Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the authors based on the results of modeling.

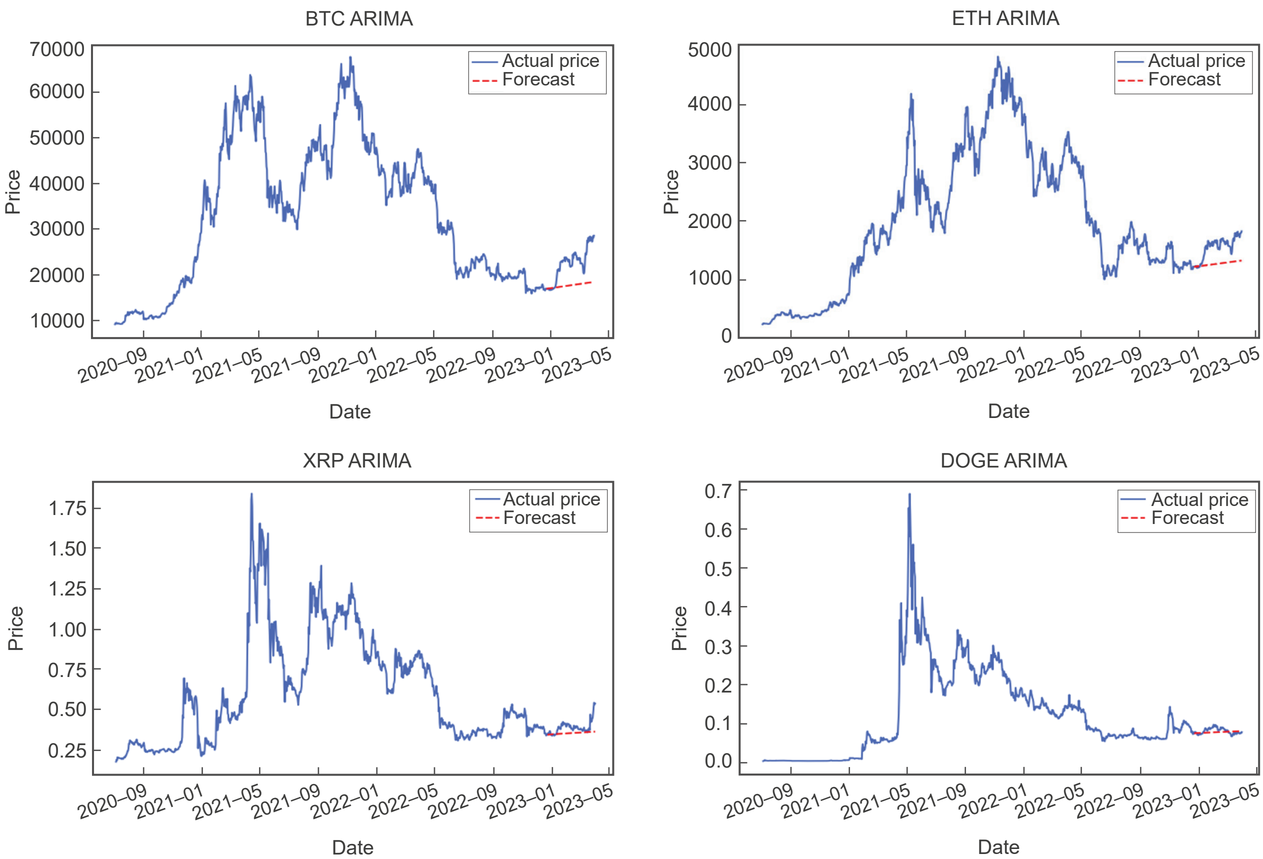

Figure 9. ARIMA Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the authors based on modeling results, ARIMA parameters are indicated in Tables 4 to 7.

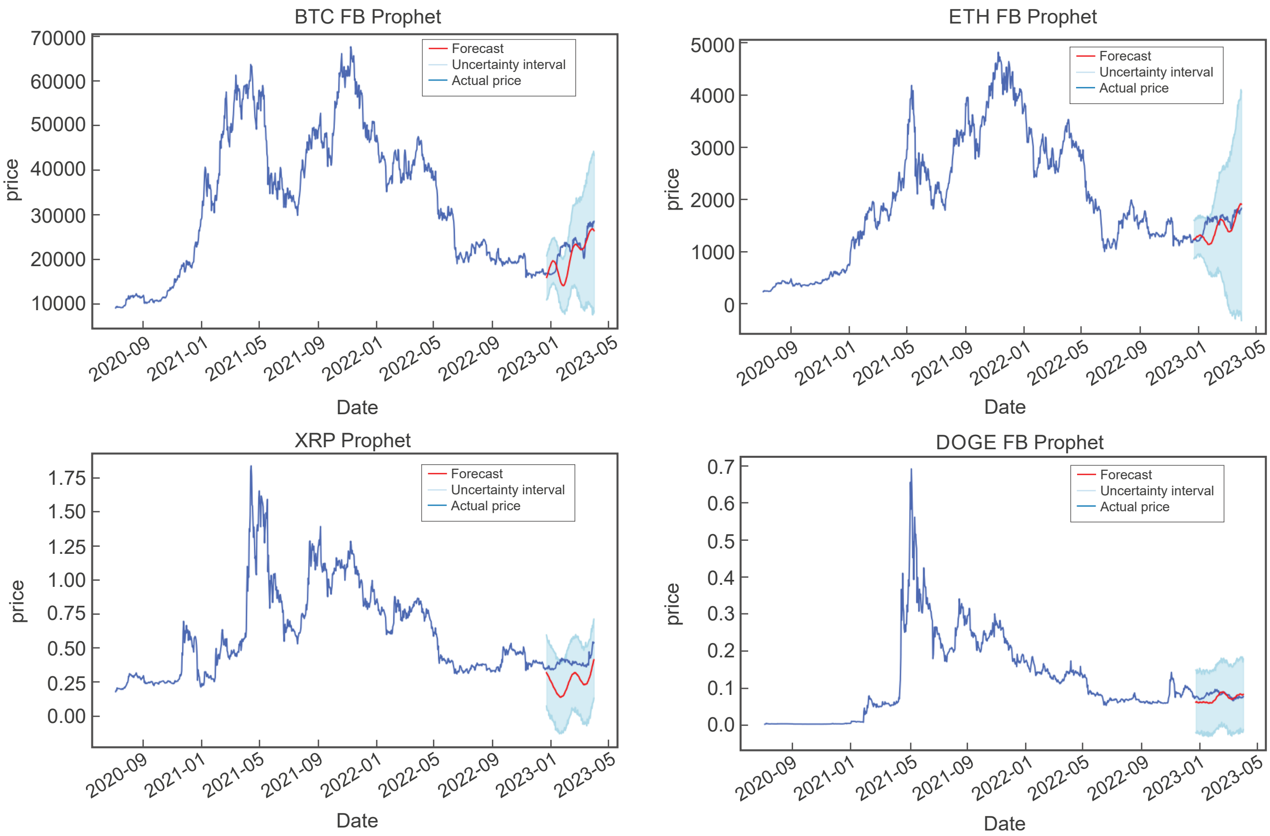

Figure 10. Prophet Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the author based on the results of modeling.

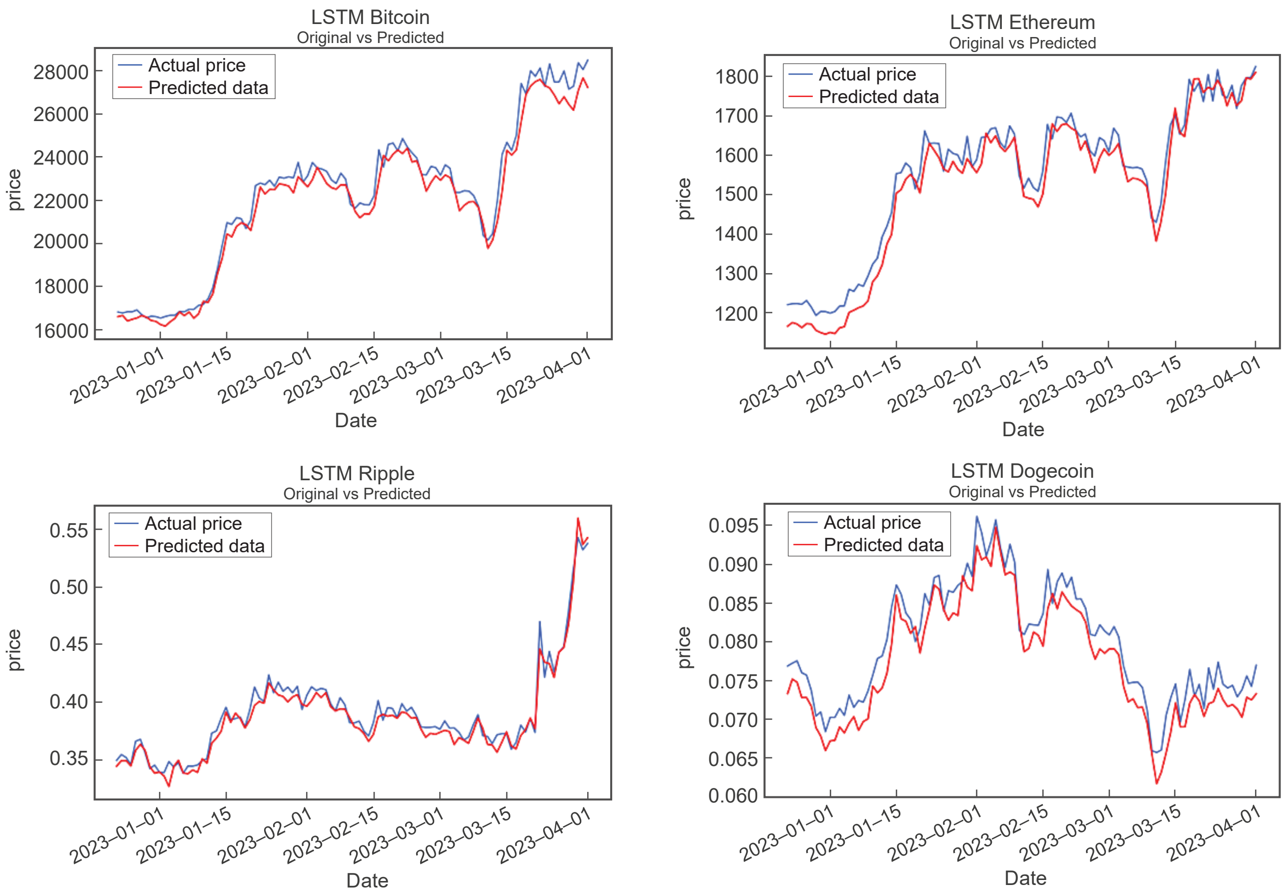

Figure 11. LSTM Forecast Models for Crypto Currencies for 23.12.2022 to 01.04.2023

Note: built by the author based on the results of modeling.