Research Paper
Academic Platform Journal of Engineering and Smart Systems (APJESS) 11(2), 47-61, 2023
ISSN: 2822-2385
Received: 19-Apr-2022 Accepted: 01-Feb-2023
https://doi.org/10.21541/apjess.1106001

Improving the Prediction Accuracy in Deep Learning-based Cryptocurrency Price Prediction

*1Furkan BALCI

[1] Department of Electrical and Electronics Engineering, Faculty of Technology, Gazi University, Ankara, Türkiye, furkanbalci@gazi.edu.tr

Abstract

Cryptocurrencies are popular today even though they do not have a physical form with their high-profit rates and increasing daily usage. However, the volatility of cryptocurrencies is higher than physical currencies. Furthermore, these volatilities change with the effect of social media rather than changes in exchange rates of physical currencies. For this reason, in this study, using Twitter data, one of the most widely used social media tools, real-time analysis of the values of four cryptocurrencies with the highest market value and the change in the estimated success compared to classical approaches were examined. This study's basic steps are obtaining Twitter data and financial data, performing sentiment analysis using Twitter data, and making predictions on MM-LSTM architecture. The approach is aimed to be a predictive method open to online learning. Furthermore, various filter steps were applied to remove the effect of bot users on Twitter that could prevent the prediction performance on the created data set, and the impact of the method on accuracy rate was tried to be reduced by eliminating the activity of bot accounts.

Keywords: Forecasting; Twitter Sentiment Score; MM-LSTM; Cryptocurrency; Deep Learning

1. INTRODUCTION

Cryptocurrencies have become a popular topic nowadays. Cryptocurrencies, which have varieties such as Bitcoin and Ethereum, do not have any physical form. After Bitcoin, Satoshi Nakamoto created in 2009, different cryptocurrencies were developed [1]. Cryptocurrencies other than Bitcoin are called altcoins. According to CoinMarketCap data, there are over 5500 cryptocurrencies today. The use of cryptocurrencies as a medium of exchange in various sectors, such as automotive and food, especially physical money, has made it widespread. As of October 2021, the total market capitalization of cryptocurrencies has exceeded $2 trillion. The market value of Bitcoin, the first cryptocurrency, is at the level of $1 trillion. The top 5 cryptocurrencies by market cap are Bitcoin (BTC), Ethereum (ETH), Cardano (ADA), XRP (XRP) and Solana (SOL).

The value of Bitcoin, which was at the level of 800 dollars at the beginning of 2017, increased to 18000 dollars at the end of the same year. This significant change has been heard worldwide thanks to social media and has aroused everyone's curiosity about cryptocurrencies. Regulations by states for the cryptocurrency market are pretty limited. Although there is no institution or state behind it, it is frequently used as an investment tool today, with the growth of the crypto money market and the high profits of many people. Although it allows investors to earn high profits in some periods, cryptocurrencies' values are uncertain today. It is complicated to know the main factor causing this uncertainty. However, today, where interactions on social media affect daily life, the values of cryptocurrencies are also affected by these interactions. Twitter is the platform where text-based interaction is most common in social media. With the results of the studies, it has been revealed that social media has a guiding effect even in presidential elections [2]. Apart from the agendas created by many people, the tweets of some critical people play an essential role in the changes in the values of cryptocurrencies [3]. Television channels about the exchange rates of physical currencies broadcast 24 hours a day, 7 days a week. However, no publications can inform the public about the changes and developments in the cryptocurrency market. That's why people tweet about their feelings and thoughts about cryptocurrencies using social media platforms like Twitter. In behavioral economics, it has been proven that different people's emotions impact the individual's decision-making mechanism. For this reason, it is argued that tweets shared on Twitter create an agenda and affect people's interest in cryptocurrencies [4, 5]. The values of physical currencies are generally constant over time. These currencies are produced by states and protected by the same condition by various laws. Cryptocurrencies lack these safeguards [6]. As a result of the research on Binance, half of the Bitcoin (on average 6 million) held by users is used for short-term (< 1 year), and the remaining half is used for long-term (>1 year) investment purposes. Most users do not use cryptocurrencies as a medium of exchange. The U.S. Securities and Exchange Commission has classified cryptocurrencies as commodities, not securities [5]. Speculation-free currencies are always more robust and more stable. There are studies on how cryptocurrencies are affected by speculation [7]. However, in some studies, it is revealed with GARCH models that cryptocurrencies have similar characteristics with gold and USD [8]. While investigating the effects of manipulations, an abnormal finding was encountered. Between 2010 and 2013, Bitcoin increased from 150 dollars to 1000 dollars with the purchase and sale of one person. Users who own the majority of cryptocurrencies are called whales. These users are open to manipulation as a large amount of buying and selling transactions are not controlled by specific rules. It has been determined that the increase in cryptocurrencies between 2017 and 2018 occurred due to extensive manipulation [9]. There are various studies to make the crypto money exchange a regular use tool. The most critical work among financial institutions is establishing the Exchange-Traded Fund (ETF). With this money exchange method, more profit can be made by taking advantage of the differences between different prices in the stock markets [10].

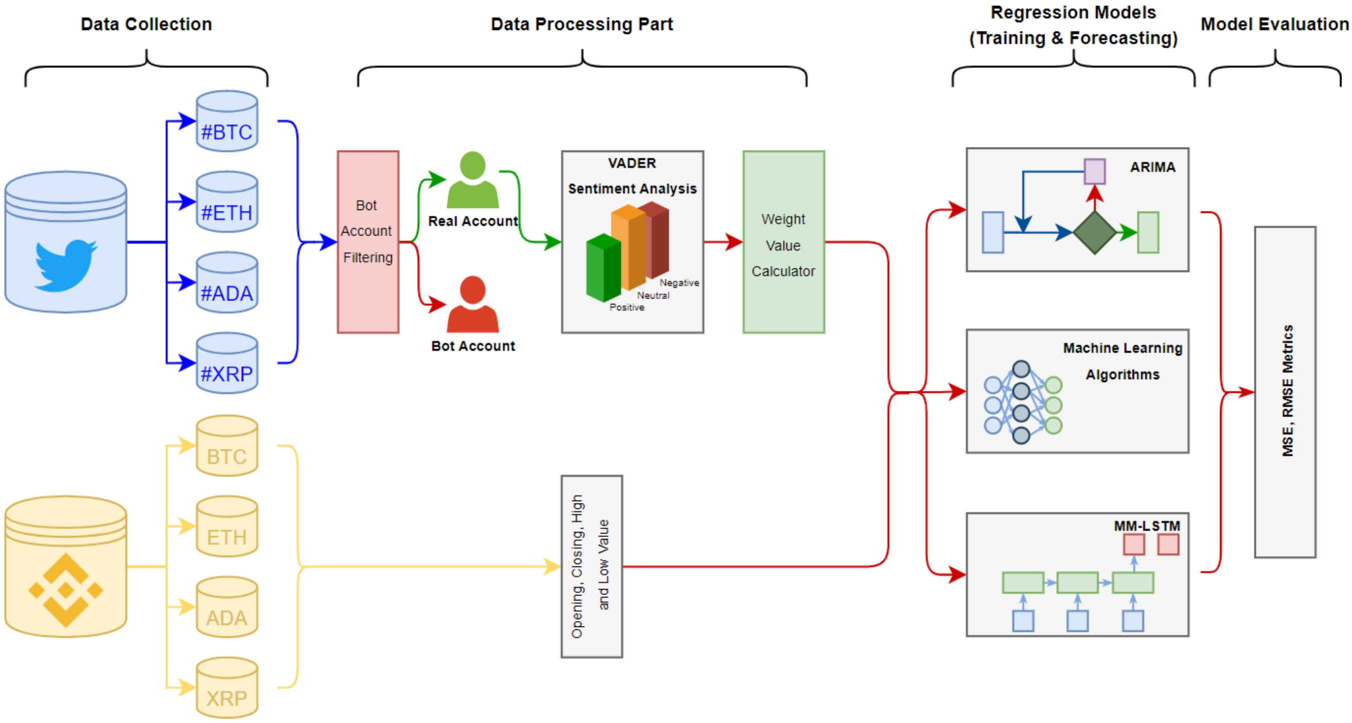

Figure 1. A general overview of the proposed methodology

It is complicated to predict the values of cryptocurrencies. For this reason, artificial intelligence techniques have been used frequently in addition to some classical methods in the financial field. These applications have had sufficient success in the short-term and low-volatility timeframes. Studies in the literature can be divided into three main categories: analysis with classical finance methods [7-10, 23], artificial intelligence-based forecasting, and artificial intelligence studies using different data (social media data, forum site data) in addition to financial data [16, 22, 26-28]. Applications that can interact with people over the internet, such as Twitter, allow the transfer of information and thoughts between people. Tweets can be shared about daily topics, politics, the economy and many similar topics. While sharing about issues, hashtags containing that topic are generally used. Various studies have been done on the push factors of the crypto money market [11–14]. Multiple factors were considered in these studies. Some of these factors are gold prices, the USD/EUR exchange rate, and the S&P 500. Twitter may also be among these factors. In the studies in the literature, it has been determined that investors are affected by the news and apply their investments in this direction [4]. The cryptocurrency market is instantly affected. Therefore, data consisting of various surveys and news in which investors' tendencies are analyzed is somewhat cumbersome in the analysis of instantaneous changes. Thus, tweets shared on Twitter allow investors to access their ideas instantly. Since each tweet is 280 characters long, there are limits to conveying emotions.

Nevertheless, Twitter data contains a rich source of information about the changes in instant market values. Extensive studies use Twitter data to forecast financial markets [4–15]. High accuracy rates were achieved in these studies. These obtained accuracy rates have high accuracy for predictions 1 – 2 days later [16]. While most of these studies in the literature test the usability of Twitter data to increase prediction accuracy, various regression models and causality tests are applied. Long-term (30 – 90 days) and short-term (1 – 7 days) data are used to forecast financial data. Prediction studies using artificial intelligence techniques reach high accuracy. Methods such as Support Vector Machine (SVM) and Vector Autoregressive (VAR) are examples of these artificial intelligence techniques. However, when the studies in the literature are examined, some limitations are encountered. There are not enough estimation studies about altcoins in the literature [29, 31-37, 41-43, 45]. Generally, there are prediction studies about the

Bitcoin market. Although the findings obtained in the studies are valid during the research, the hypotheses produced due to the sudden changes in the crypto money market lose their validity. Another limitation is that there are limits on the data obtained through the Twitter API. There is a monthly limit of 5 million tweets for academic studies. Therefore, the sentiment analysis accuracy in the tweets obtained significantly affects the studies.

In this study, cryptocurrencies, including altcoins, are estimated by producing various solutions to the deficiencies of the studies in the literature. There are many suggested approaches with data from social networks such as Twitter. However, the most important of the shortcomings in these approaches is not paying attention to the fact that the effect of every tweet or every shared post is not the same. Unlike the studies in the literature, only tweets shared at the same minute were not taken as supporting data. Instead, a scoring approach was developed to reveal the effect of each tweet and its predictive performance was tested. Classical economics approaches such as ARIMA and regression algorithms such as Decision Tree, Random Forest, Linear Regression, and Logistic Regression were used for comparison. Against these algorithms, the prediction success of the MM-LSTM architecture fed by the Twitter score obtained by the weight value from Eq.1. In summary, forecast accuracy has been increased with a VADER sentiment score-based weight value equation developed for the problem of low forecast accuracy due to the extreme volatility of cryptocurrencies. To test the effectiveness of this developed sentiment analysis approach, both classical methods, machine learning methods and deep learning methods were used. The method with the lowest error rate among the methods used is MM-LSTM.

The remainder of this paper is structured as follows: Section 2 will explain the related works. Section 3 will explain the methodology. In Section 4, the experimental results of the proposed method and the outputs of the analyses. The results obtained will be discussed in Section 5. Finally, the findings of this study are summarized in Section 6.

2. RELATED WORKS

Studies in the literature follow two different methods: classification and regression. It provides information for future investments in the outputs of algorithms in studies that make the classification. In these studies, he usually makes predictions about whether the value of the cryptocurrency will increase or decrease for certain intervals such as 1 minute, 1 day, 15 minutes. Studies that predict the next 1 minute are very few due to the running times of the algorithms, and generally, simple machine learning methods are used in these studies.

The first studies on cryptocurrencies are about whether these assets are usable [17-20]. Wen et al. predicted that there would be some problems with cryptocurrencies in the coming years [21]. The biggest of these problems is that Chinese companies invest enough to determine the cryptocurrency market.

In the following years, with the increase in the value of cryptocurrencies, research topics began to be about predicting future values. Therefore, correlation studies have started between different data and the values of cryptocurrencies. The relationship of cryptocurrencies between Google Trends and Wikipedia search numbers was examined. A high correlation was found in this relationship [22]. In different studies, relations with varying assets with a market have been discussed. It does not correlate with assets such as Financial Stress Index and gold price [23]. The most comprehensive study analyzes 21 different parameters and shows that Google Trends searches have the highest correlation [24]. The relationship with the prices of assets such as stocks, oil barrel prices, and exchange rates is relatively low both in the expected period and when cryptocurrencies change abruptly [25]. The general result obtained in these studies is as follows, a high correlation is seen in data such as social media shares, Google search trends, and opinions on forum sites [7]. Studies about cryptocurrencies with high social media shares have been concentrated in the literature. Some studies forecast daily bitcoin, ripple and Ethereum prices with data from forum sites [26]. Different estimation methods, such as the Markov model, were also used in these studies [27]. It has been determined that the Coronavirus epidemic, which is one of the current issues, affects prices and expands transaction volumes [28].

Recent studies have revealed approaches based on predicting the prices of cryptocurrencies. In Table 1, studies in the literature are listed chronologically. The studies are divided into two subclasses, classification or regression, if the table is examined in detail. It contains brief information about the methods used. In each study, methods that will increase the predictive power, in general, are revealed. Different machine learning and deep learning methods are used in the literature. But there is no consensus on which way gives the highest accuracy. It is understood that only classical methods (such as ARIMA) have lower performance than machine learning and deep learning methods. Most studies only research on Bitcoin. Studies similar to the one performed in this article were performed in [29] and [30]. In these studies, different machine learning methods were used in the study.

Our main difference from the studies in the literature is not only on Bitcoin but also on Ethereum, XRP and Cardano with the Twitter sentiment. In addition, unlike the literature, estimates are not made for time intervals such as 15 minutes, 1 hour, and 1 day, but operations are carried out 1 minute later. Due to the extreme volatility of cryptocurrencies, classical economic forecasting methods (AR, MA, ARIMA, etc.) predict cryptocurrencies with high errors. In general, since artificial intelligence-based studies have a higher ability to learn the volatility of cryptocurrencies, there are often artificial intelligence-based studies in the literature.

3. MATERIALS AND METHODS

In this study, the usability of public Tweets shared on Twitter as a feature to increase the prediction of cryptocurrencies was investigated using deep learning architecture. The proposed method of the study is shown in Figure 1.

Table 1. Literature review

#	Cryptocurrency	Prediction	Data Length	Type	Method
[29]	Bitcoin	Daily	2011-2018	Regression Classification	It shows that the Long Short-Term Memory (LSTM) algorithm for regression gives better results than the Deep Neural Network (DNN) algorithm for classification.
[31]	Bitcoin	10 minutes	Since the discovery of Bitcoin	Classification	It uses Bitcoin closing price and 16 different features of Bitcoin as data. Both 10-minute and 10-second data were used as an additional amount of data. Classification results with 10-second data are higher. Random Forest and Binomial Logistic Regression (BLR) algorithms were used.
[26]	Bitcoin, Ethereum, Ripple	Daily	2013-2016	Classification	The classification was made with various commercial data and comments collected from online forums. Data collected from forum sites provided satisfactory improvement.
[32]	Bitcoin	15 minutes	1 month	Classification	Ten different technical analysis indicators about Bitcoin are used as features. In addition, a study on the average return was carried out with the algorithm developed based on Volume Weighted-SVM.
[33]	12 Cryptocurrencies	30 minutes	2015-2016	Regression	A regression-based on Convolutional Neural Networks (CNN) was made. In addition, forecasts of 12 different cryptocurrencies were carried out.
[34]	Bitcoin	Daily	2011-2017	Regression	Regression was performed using Bayesian neural networks (BNN), linear regression and SVM. In the study, it is argued that the BNN architecture is the best predictive method.
[35]	Bitcoin	Daily	2013-2016	Classification Regression	Regression was performed using Bayes repetitive nerve (RNN) and LSTM algorithms. It has been determined that the most suitable data amount for the LSTM algorithm is 100-120 days, and the most appropriate for the RNN algorithm is 15-20 days.
[36]	Bitcoin	15 minutes	2016-2018	Classification	The classification was made using artificial neural networks (ANN). 4 different technical analysis indicators were used.
[37]	Bitcoin	1 minute	2012-2017	Classification	The classification was made using Random Forest (RF). He used five different technical analysis indicators. He argues that 15-minute forecasts are more accurate.
[38]	1681 Cryptocurrencies	Daily	2015-2018	Regression	A collection of regression trees created by XGboost and the long-term memory network is used. As a result, a profitable forecasting system has been developed against up to 0.2% transaction fees.
[39]	Bitcoin, Ethereum, Litecoin, Ripple	Daily	2011-2017	Classification Regression	A hybrid neuro-fuzzy model (PATSOS) has been developed, and higher results have been obtained than classical machine learning methods.
[40]	Bitcoin, Ethereum, Litecoin, Ripple	Daily	2015-2017	Regression	Linear univariate and multivariate regression models and their selections and combinations were tested individually. It was observed that the results of combinations of univariate models were lower.
[41]	Bitcoin	Daily	2013-2018	Regression	SVM and ANN algorithms are used. It has been observed that the SVM method makes more protective decisions.
[42]	Bitcoin	Daily	2011-2018	Classification Regression	DNN, LSTM, CNN, Deep Residual Network (ResNet), CNNs and RNNs (CRNN) and their combinations were used. LSTM for regression achieved higher accuracy than DNN for classification.
[30]	Bitcoin, Ripple	Daily	2011-2018	Regression	A comparison was made between LSTM and Generalized Regression Neural Networks (GRNN). The prediction accuracy of LSTM is higher than GRNN.
[43]	Bitcoin	Daily	2011-2017	Classification Regression	Artificial neural networks (ANN) and SVM are used. It is argued that SVM is the best regression model. OHLC prices and various external financial variables are used as data.
[44]	100 Cryptocurrencies	1 minute	Until 2018	Classification	Logistic regression (LR), RF, SVM and Gradient Tree Boosting algorithms were used.
[45]	Bitcoin	5 minutes Daily	2017-2018	Classification	LR, Linear Discriminant Analysis (LDA), RF, XGBoost (XGB), SVM and LSTM algorithms are used. The LSTM achieved the highest accuracy in the 5-minute data.
[31]	42 Cryptocurrencies	Daily	6 months	Classification	LightGBM, SVM and RF methods were used for 2-week forecasts. The LightGBM method achieved the best results.

3.1. Financial Data

This study focuses on the improvement that can be made to the forecast of the four largest cryptocurrencies by market capitalization. Particularly focused on cryptocurrencies with high interaction. These cryptocurrencies are Bitcoin (BTC), Ethereum (ETH), Cardano (ADA), XRP (XRP) and Solana (SOL). Binance was used to collect financial data. Financial data were collected between September 6, 2021, and October 3, 2021. This period is a period of 4 weeks. Hourly and minute data was received from Binance. In order to express the value of cryptocurrencies, the physical currency USD, which has general usage in the world, was used. The labels of the financial data used as time series are used with the date and time of the relevant day.

3.2. Twitter Data

In order for the interaction with financial data to be examined, the time interval of the Twitter data and the time interval of the financial data must be the same. A data set was created by collecting tweets separately for each cryptocurrency. Twitter API developed by Twitter for researchers was used to collect Twitter data. Tweets shared publicly on Twitter can be accessed using the Twitter API. On Twitter, users use the hashtag (#) prefix before some keywords related to the topics they are talking about. Thanks to this use, both the agenda is created, and the Tweets shared about that subject can be easily accessed. The dataset was created using hashtags for each cryptocurrency. The prefix before some keywords are related to the topics they are talking about. The dataset was created using hashtags for each cryptocurrency. However, today, artificial agenda creation studies are also carried out for the crypto money market with the use of bot accounts. For this reason, it is necessary to clean the shares made by bot accounts before using the dataset. Bot accounts were identified and filtered by analyzing Tweets and the number of people who shared these Tweets, and the number of people following these users. There are many studies on the detection of bot accounts. In the detection of bot accounts, first of all, it should be understood how the sharing movements of bot accounts are. In this regard, some basic generalizations have been considered. These generalizations are: Tweets with content about free cryptocurrency giveaways, bot account link sharing, almost zero follower/follower ratio. By using these generalizations, the dataset is free of bots. Today, the intense use of social media has made it easier for these environments to be neglected by malicious people. Bot accounts can spread fake news using Twitter. Known for its corporate identity, Twitter social media application is used by every segment of society. The data shared by Twitter has more than 400 million users. According to Alexa, developed by Google, although Twitter ranks 13th among the most visited sites in the world, abuses on this social platform have also increased. The use of Twitter as a usable identity at the entrance of some applications makes it possible to access the Twitter application from different platforms easily. This situation attracts malicious users more. Generally, negative Twitter accounts have goals such as gaining more followers, influencing a specific community to make people join their organization, manipulating people for the stock market, spreading fake news, and blackmailing people using private information. Bots in social networks are computer software that shows automatic reactions that mimic human behavior [48]. Today, it is thought that 15% of active Twitter users are bots [49]. Zi et al. used an entropy-based layered architecture to detect Twitter bot-human-Cyborg [50]. Software frameworks that account for Twitter bots' characteristics and detect accounts by automatic feature extraction have also been developed [49]. In the dataset used in another study, Twitter profiles were discussed according to the number of different followers [51]. The study used all samples in the data set for training and testing, thanks to the cross-validation method. In addition, the diversity of features in the dataset was also crucial in bot detection. The most comprehensive competition for detecting Twitter bots is “The DARPA Twitter Bot Challenge”, held during the 2016 American Presidential Elections. The six teams participating in the competition worked on 7,038 accounts and approximately 4 million tweets [52]. Bot accounts on Twitter were detected with the "warped correlation" technique developed in another study. With this method, 544,868 bot accounts were detected annually, and the success rate was 94% [53]. In the Twitter environment, bot detection is essential for the security of society. Recently, new deep-learning networks have started to be used in Twitter bot detection [54]. Both machine learning and deep learning methods use metrics such as accuracy, precision and F-score in detecting bot accounts. Fred et al. contributed from a different point of view in evaluating these metrics [55]. In some approaches in the literature, many parameters are used to determine bot accounts. Some of these parameters are:

- Age of the account: The age of the account in days is higher for real users.
- The ratio of followers to friends: It is observed that the ratio of real users is close to 1.
- User favorites: The number of favorite tagging tweets by users is higher in real users.
- The number of retweets per tweet: Real users get more retweets due to original content and variety.
- Number of URLs: Real users share fewer URLs.
- Frequency of tweeting: Bot users share more tweets during the day.

To be able to do sentiment analysis later, the text content must be made available for Sentiment analysis. Some steps must be followed for this process. These steps are shown in Table 3. on a sample tweet, respectively. These steps were applied to all tweets, and as a result of these pre-processes, the dataset was ready for sentiment analysis. Statistical data and general information about the data obtained with the access permission obtained from Twitter are shown in Table 1. Some of the statistical information is important for artificial intelligence techniques. Especially the number of data is essential for training and testing. Therefore, the number of Tweets is shared in the table. Twitter can reshare another person's tweet. Therefore, information about unique shared tweets is also provided. There is standard deviation (STD) information about the daily data, which includes the average number of tweets and the difference from the average daily tweets. The average number of tweets and standard deviation information provide us with information about how volatile the agenda is. If the table is examined in detail, Bitcoin is the most talked about cryptocurrency because it is both popular and dominates more than half of the market. There is not enough sharing to create an agenda on Twitter about other sub-cryptocurrencies. That's why only these four cryptocurrencies have been selected.

Table 2. Statistics of the filtered Twitter datasets

	Number of Tweets	Number of Unique Tweets	Mean Daily Tweets	STD of Daily Tweets
BTC	3.101.251	1.352.077	110.759	33.985
ETH	738.192	379.124	26.364	3.846
ADA	486.224	203.218	17.365	5.199
XRP	375.117	79.826	13.397	2.491

Not every Tweet has the same effect. Some people can affect the agenda more due to the high number of followers. This effect needs to be measurable. Using a scoring system to understand the impact of users' Tweets makes Twitter's impact more meaningful. There are various scoring systems in the literature. VADER (Valence Aware Dictionary and Sentiment Reasoner) approach will be used in this study to measure the effect of Tweets shared by users publicly. In the VADER method, Tweets shared by users are analyzed and classified according to their emotions. The words in the dictionary within the VADER structure play an essential role in this scoring process. The compound score is calculated with this dictionary. The punctuation marks used are necessary for the compound score. The Tweet's language was translated into English using the Python module developed by Google on Tweets shared in different languages. The compound score takes a normalized value between -1 and +1. A value of -1 means the most extreme negative, and a value of +1 means the most positive. Positive sentiment analysis means buying, neutral sentiment analysis means waiting, and positive sentiment analysis means selling cryptocurrencies. In this scoring system, an equation developed outside of the VADER structure will be used to turn the effect of the Tweet shared by the user into usable numerical data in the deep learning structure. In Eq. (1), the compound score (CS), the number of followers (FC), the number of likes (TL) and the number of retweets of the shared Tweet (RT) are used as a multiplier. This designed score number is updated instantly and creates a supporting feature for the cryptocurrency closing data, which is planned to be predicted for the next minute. To investigate the contribution of this data, prediction performances on different algorithms will be compared. A comparison of supported and unsupported algorithms will also be made with this data.

$$WeightValue = CS * FC * (TL + 1) * (RT + 1) \tag{1}$$

Table 3. Example of pre-processing steps to be applied to tweet for Twitter score

Process Number	Process Defination	Example
0	Initial Tweet	RT @ethereum https://t.co/FT/status/98342573428 Bitcoin will make us richeeeerrr than yesterday!!! Go and buy lol, end of the year it will be $100000 #BUY #ETHEREUM $BTC $ETC
1	Removing "RT" text	@ethereum https://t.co/FT/status/98342573428 Bitcoin will make us richeeeerrr than yesterday!!! Go and buy lol, end of the year it will be $100000 #BUY #ETHEREUM $BTC $ETC
2	Removing links	Bitcoin will make us richeeeerrr than yesterday!!! Go and buy lol, end of the year it will be $100000 #BUY #ETHEREUM $BTC $ETC
3	Reducing repeated characters	Bitcoin will make us richeerr than yesterday!!! Go and buy lol, end of the year it will be $100000 #BUY #ETHEREUM $BTC $ETC
4	Uppercase character conversion	bitcoin will make us richeerr than yesterday!!! go and buy lol, end of the year it will be $100000 #buy #ethereum $btc $etc
5	Removing hashtags	bitcoin will make us richeerr than yesterday!!! go and buy lol, end of the year it will be $100000 $btc $etc
6	Expansion of abbreviations	bitcoin will make us richeerr than yesterday!!! go and buy laughing out loud, end of the year it will be $100000 $btc $etc
7	Removing punctuation	bitcoin will make us richeerr than yesterday go and buy laughing out loud, end of the year it will be 100000 btc etc
8	Removing numeric expressions	bitcoin will make us richeerr than yesterday go and buy laughing out loud, end of the year it will be btc etc

3.3. Statistical Analysis

The value of cryptocurrencies and the number of tweets shared are time-dependent data. These data types are called time series. The stationarity of the time series is one of the most important criteria. Stationarity is when the mean and variance of the time series are not dependent on time. The covariance value depends on the mean and variance values. Various statistical models should be used first to see that the approach can benefit the deep learning model in examining the relationship between the cryptocurrency values data used in this study and the number of tweets shared in the relevant period. Correlation analysis gives us information about the relationship between two data and the direction of this relationship. Simple correlation analyzes may be insufficient because there will be a time difference between the two data.

If the values in two different time series have a relationship over time concerning a particular lag time, and it is desired to determine the direction of this relationship, one of the tests that can be used is the Granger causality test. The Granger causality test can be defined as follows: If an X variable causes a change in a Y variable, changes in X will lead to changes that will occur in Y after a while. That is, if the estimation is significantly improved when variable X or the delayed values of variable X are added to the regression of variable Y with other variables, then variable X is the Granger cause of variable Y [46].

Although there are many views on the definition of causality, there is a common view on establishing a cause-effect relationship between the variables. The existence of a strong relationship or correlation between the variables does not necessitate the existence of causality. For example, regression analysis establishes a statistical relationship between variables and does not deal with cause-effect relationships.

In this context, Granger developed a relatively simple test in 1969 to reveal the cause-effect relationship between the variables. According to Granger, if the success of the estimation increases with the inclusion of the past values of the X independent variable in the estimation of any Y variable, the X variable is a cause of the Y variable. The Granger causality test is expressed in the context of linear regression models. For example, it is described as a linear autoregressive model with two variables such as X_1 and X_2. It is expressed mathematically as follows [46]:

$$x_1(t) = \sum_{j=1}^{p} A_{11} X_1(t-j) + \sum_{j=1}^{p} A_{12} X_2(t-j) + E_1(t) \tag{2}$$

$$x_2(t) = \sum_{j=1}^{p} A_{21} X_1(t-j) + \sum_{j=1}^{p} A_{22} X_2(t-j) + E_1(t) \tag{3}$$

In the equation above, p represents the maximum number of delayed observations included in the model. If the variance of E_1 decreases when the term X_2 is included in the equation above, it is understood that X_1 is the Granger cause of X_2. Likewise, if the variance of E_2 decreases when the term X_1 is included in the equation above, it is understood that X_2 is the Granger cause of X_1. In other words, X_2 becomes the Granger cause of X_1 if the coefficients in A_{12} together are unique in relation to zero. This situation can be understood by performing the F-test of the H_0 hypothesis with when A_{12} is 0, when the assumption of covariance stationarity for X_1 and X_2 is accepted. The magnitude of a Granger causality relationship is measured by taking the logarithm of the corresponding F-statistic. Choice of criteria methods like Hannan-Quinn Information Criteria (HQ), Schwartz Information Criteria (SIC) or Akaike Information Criteria (AIC) can be utilized to decide the fitting number of delays.

3.4. Multi-step Multivariate Long Short-Term Memory

This paper created a deep learning structure to predict the value of cryptocurrencies. This deep learning structure is based on the Recurrent Neural Network (RNN) structure. The value to be estimated in RNN structures does not only analyze based on the current value but also based on historical data. Therefore, RNN structures are frequently used in time series data [47, 56]. RNN structures do not delete old data, such as the work of the human brain. Classical neural network structures delete after using old data in the weighting setting [56]. RNN structures have been developed to cover this gap. The data used in RNN-based structures are stored in memory units until the cycle is completed. RNN structure is basically the same as classical neural networks [57]. This structure is formed by listing the same networks. The input of each network depends on the output of the previous network. There are varieties of RNN structures. Long Short-Term Memory (LSTM) structure is used in this paper. LSTM structure has begun to be used widely in estimation processes based on historical data. RNN structure has a single-layer network structure.

The LSTM structure has a four-layer network structure. There are structures called gates in the LSTM structure. These gate structures perform tasks such as adding and removing information to the neural cell. The sigmoid function used in the neural network layer gives values between 0 and 1. The sigmoid function determines how much of the signal is allowed to pass. This value varying between 0 and 1 is used as a ratio. The first of these gates is called the forget gate. This gate investigates the prior output and the current input and produces a value between 0 and 1. On the off chance that the produced value is 0, it signifies “forget this state”; assuming the produced value is 1, it represents “keep this state” [47]. The forget gate is indicated by ft. Eq. (4) shows the equation of the forget gate result.

$$f_t = sigmoid(W_f[h_{t-1}, x_t] + b_f) \tag{4}$$

Another gate layer is the input gate. This gate structure decides which new values to keep. In this step, both sigmoid and tanh functions are used. The sigmoid structure produces the value to be updated, and the tanh structure produces the intermediate value C_{tx}. Eq. (5) shows the equation of the sigmoid function. Eq. (6) shows the equation of the tanh function. Then these values are combined.

$$i_t = sigmoid(W_i[h_{t-1}, x_t] + b_i) \tag{5}$$

$$C_{tx} = tanh(W_c[h_{t-1}, x_t] + b_c) \tag{6}$$

Using it and C_{tx} values, C_t is generated, allowing the old data to be transferred to the next cell. For example, Eq. (7) shows the current data equation C_t obtained with old data and new entries.

$$C_t = f_i * C_{t-1} + i_t * C_{tx}) \tag{7}$$

In the next step, the output of that cell must be calculated. This calculated output is branched to be used in the next cell. Finally, deciding which data will be used as output from the cell is necessary. The sigmoid function is used to make this decision. Eq. (8) shows the equation of the sigmoid function. The tanh function converts the sigmoid function's result to -1 and 1. The final cell output is obtained after this transformation. Eq. (9) shows the equation of the last cell output.

$$o_t = sigmoid(W_o[h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t * \tanh(C_t) \tag{9}$$

Table 4. Pseudocode of the proposed MM-LSTM architecture

Algorithm 1. *MM-LSTM Architecture*

Input: *Multivariate Cryptocurrency Data, Learning Rate(λ), Twitter Score*

Output: *Predicted Value* ($x_{N+1}^{(1)}$)*, Test MSE, Test RMSE*

Initialize the parameters of MM-LSTM

test size ← *length(data)* × *(100 – λ)*

for i ∈ (1, …, epoch) **do**

$\hat{x}_t^{(1)}$ ← *LSTM*($x_t, \ldots, x_{t-slidingwindowsize+1}$)

Loss ← $\left\| \hat{x}_t^{(1)} - x_t^{(1)} \right\|$

Optimize the parameters based on Loss

Back propagation method with λ

end for

Predicted Value List ← $x_{N+1}^{(1)}$

Test RMSE ← $\sqrt{\frac{\sum_{i=1}^{testsize}(x_{N+1}^{(1)} - x_{N+1})^2}{testsize}}$

return *Predicted Value List, Test RMSE*

The Multivariate Multi-step LSTM structure used in this study is similar to the classical LSTM structure described above. However, it has 2 differences from the classical LSTM structure. These differences:

- The dependent variable estimation result is estimated with more than one independent variable.
- The estimated value includes more steps than a single step.

LSTM architecture is used in the study carried out in various academic fields, such as biomedical data in time series, economic data, and time series data related to the process [58 – 61].

4. RESULTS AND ANALYZES

Two separate data sets are needed to carry out this study. First of all, the dataset that needs to be pre-processed and analyzed is the Twitter dataset. Table 2 contains statistical data on Twitter data. In light of these data, it has been determined that the tweets of bot users at various rates (1% - 17%) set the agenda. These accounts, created to share false information and manipulate, need to be filtered out. The posts of bot accounts are mostly about low-value cryptocurrencies. If Table 2 is examined, the effect of bot accounts on each cryptocurrency is also shown as a percentage distribution. Some essential criteria detect these bot accounts. It can be said that there may be some bot accounts that exceed these criteria. During the detection of bot accounts, attention was paid to the fact that the follower/following ratio was almost zero and that the tweet content included giving away crypto money. This study's sentiment analysis on publicly shared Tweets investigates the predictability of cryptocurrencies' values. The analysis method used to examine this relationship is Granger causality analysis. To search for causality between series, stationarity information is needed. If the series are stationary of the same order, a cointegration relationship can be sought between them. If a cointegration relationship is not observed, causality can be investigated in the order that the series are stationary. Thanks to the Granger causality test, it is revealed whether tweets shared by users cause a change in the value of cryptocurrencies. It plays a vital role in seeing the relationship between the scoring process using the compound score calculated with the VADER method and the value of the relevant cryptocurrency at that moment. With these scores, the Granger causality test was applied. Here, the Granger causality test was applied for both cases to understand whether Twitter data contains causality on cryptocurrencies or on Twitter data. The Granger causality test results, the maximum order of integration and p values for each cryptocurrency are shown in Tables 5 to 8. The section shown in red shows the Twitter score calculated per minute, obtained within the framework of our approach. The part shown in blue is the closing data of the cryptocurrency received on Binance at that minute. It is important here that there are statistically significant relationships ($p < 0.05$). Cryptocurrencies with lower market caps are more susceptible to manipulation. For this reason, if we look at the results of the Granger analysis, the causality relationship is higher for cryptocurrencies with lower market capitalization.

Eight different methods were used to analyze the performance of the proposed method. First, ARIMA, LTSM, Support Vector Regressor, Decision Tree Regressor, Random Forest Regressor, Linear Regressor, Logistic Regressor, Gaussian Process Regressor and MM-LSTM architectures were tested only with cryptocurrency closing data. Then these architectures were tested with Twitter sentiment analysis score and cryptocurrencies' opening, closing, high and low data. Next, the Root Mean Square Error (RMSE) calculation was calculated separately for all methods to make the performances comparable. RMSE is a metric used to create a comparison of predicted values. It explains how far each predicted value deviates from the correct value. The RMSE result can start from 0 and take forever. Eq. (10) shows the equation for calculating the RMSE value. Each RMSE result is given in Tables 9 and 10.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicted_i - Actual_i)}{N}} \tag{10}$$

Since different types of data (Cryptocurrency value, Twitter score) are used in the MM-LSTM structure, the data were normalized before using the deep learning model for training and testing. The Min-Max Scaling method was used, one of the various normalization methods in the literature. Due to the use of the minute dataset, there are 40200 samples in each feature for all three methods between the relevant dates. Finally, the closing data of each cryptocurrency is used to make predictions with eight regression algorithms.

ARIMA models are applied to non-stationary series but converted to stationary by differencing. Models used to non-stationary series but converted to stationary by difference-taking are called non-stationary linear stochastic models. These models are AR, applied to series with a d-degree difference, in which the value of the variable in the t-period is expressed as a linear function of a certain number of back-period values and the error term in the same period. The variable's value in the t-period is a linear function of the error term in the same period and a certain number of back-period error terms. They are a mix of MA models expressed as the general representation of the models is ARIMA (p, d, q). Here, p and q are the levels of the autoregressive model (AR) and the moving average model (MA) separately, and d is the level of difference. The general ARIMA (p, d, q) model is formulated as Eq. (11).

$$Z_t = \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} + \delta\alpha_t - \Theta_1\alpha_{t-1} - \Theta_2\alpha_{t-2} - \ldots - \Theta_q\alpha_{t-q} \quad (11)$$

In Eq. (11); Z_t, Z_{t-1}, ... d-degree observation values, Φ_1, Φ_2, ... coefficients for d-degree differentiated observation values, δ constant value, α_t, α_{t-1}, ... error terms and Θ_1, Θ_2, ... error show the coefficients related to the terms.

Table 5. Bitcoin Granger causality test results

BTC Number of Lags: 1		(dfdenom=40196)		BTC Number of Lags: 2		(dfdenom=40193)	
F test on SSR	23.166	p=0.0000	num=1	F test on SSR	8.259	p=0.0003	num=2
chi2 test on SSR	23.168	p=0.0000	df=1	chi2 test on SSR	16.521	p=0.0003	df=2
Likehood ratio test	23.161	p=0.0000	df=1	Likehood ratio test	16.518	p=0.0003	df=2
F test parameter	23.166	p=0.0000	num=1	F test parameter	8.259	p=0.0003	num=2
BTC Number of Lags: 3		(dfdenom=40190)		BTC Number of Lags: 4		(dfdenom=40187)	
F test on SSR	3.347	p=0.0182	num=3	F test on SSR	1.517	p=0.1942	num=4
chi2 test on SSR	10.042	p=0.0182	df=3	chi2 test on SSR	6.069	p=0.1940	df=4
Likehood ratio test	10.040	p=0.0182	df=3	Likehood ratio test	6.069	p=0.1941	df=4
F test parameter	3.347	p=0.0182	num=3	F test parameter	1.517	p=0.1942	num=4

Table 6. Ethereum Granger causality test results

ETH Number of Lags: 1		(dfdenom=40196)		ETH Number of Lags: 2		(dfdenom=40193)	
F test on SSR	103.098	p=0.0000	num=1	F test on SSR	28.923	p=0.0000	num=2
chi2 test on SSR	103.105	p=0.0000	df=1	chi2 test on SSR	57.852	p=0.0000	df=2
Likehood ratio test	102.973	p=0.0000	df=1	Likehood ratio test	57.811	p=0.0000	df=2
F test parameter	103.098	p=0.0000	num=1	F test parameter	28.923	p=0.0000	num=2
ETH Number of Lags: 3		(dfdenom=40190)		ETH Number of Lags: 4		(dfdenom=40187)	
F test on SSR	7.588	p=0.0000	num=3	F test on SSR	2.434	p=0.0451	num=4
chi2 test on SSR	22.767	p=0.0000	df=3	chi2 test on SSR	9.738	p=0.0451	df=4
Likehood ratio test	22.760	p=0.0000	df=3	Likehood ratio test	9.737	p=0.0451	df=4
F test parameter	7.588	p=0.0000	num=3	F test parameter	2.434	p=0.0451	num=4

Table 7. Cardano Granger causality test results

ADA Number of Lags: 1		(dfdenom=40196)		ADA Number of Lags: 2		(dfdenom=40193)	
F test on SSR	24.711	p=0.0000	num=1	F test on SSR	8.605	p=0.0002	num=2
chi2 test on SSR	24.713	p=0.0000	df=1	chi2 test on SSR	17.213	p=0.0002	df=2
Likehood ratio test	24.706	p=0.0000	df=1	Likehood ratio test	17.209	p=0.0002	df=2
F test parameter	24.711	p=0.0000	num=1	F test parameter	8.605	p=0.0002	num=2
ADA Number of Lags: 3		(dfdenom=40190)		ADA Number of Lags: 4		(dfdenom=40187)	
F test on SSR	4.417	p=0.0041	num=3	F test on SSR	2.505	p=0.0401	num=4
chi2 test on SSR	13.254	p=0.0041	df=3	chi2 test on SSR	10.022	p=0.0401	df=4
Likehood ratio test	13.252	p=0.0041	df=3	Likehood ratio test	10.021	p=0.0401	df=4
F test parameter	4.417	p=0.0041	num=3	F test parameter	2.505	p=0.0401	num=4

Table 8. XRP Granger causality test results

XRP Number of Lags: 1		(dfdenom=40196)		XRP Number of Lags: 2		(dfdenom=40193)	
F test on SSR	91.789	p=0.0000	num=1	F test on SSR	48.881	p=0.0002	num=2
chi2 test on SSR	91.796	p=0.0000	df=1	chi2 test on SSR	97.775	p=0.0002	df=2
Likehood ratio test	97.657	p=0.0000	df=1	Likehood ratio test	97.657	p=0.0002	df=2
F test parameter	91.789	p=0.0000	num=1	F test parameter	48.881	p=0.0002	num=2
XRP Number of Lags: 3		(dfdenom=40190)		XRP Number of Lags: 4		(dfdenom=40187)	
F test on SSR	20.967	p=0.0041	num=3	F test on SSR	9.847	p=0.0401	num=4
chi2 test on SSR	62.913	p=0.0041	df=3	chi2 test on SSR	39.398	p=0.0401	df=4
Likehood ratio test	62.863	p=0.0041	df=3	Likehood ratio test	39.379	p=0.0401	df=4
F test parameter	20.967	p=0.0041	num=3	F test parameter	9.847	p=0.0401	num=4

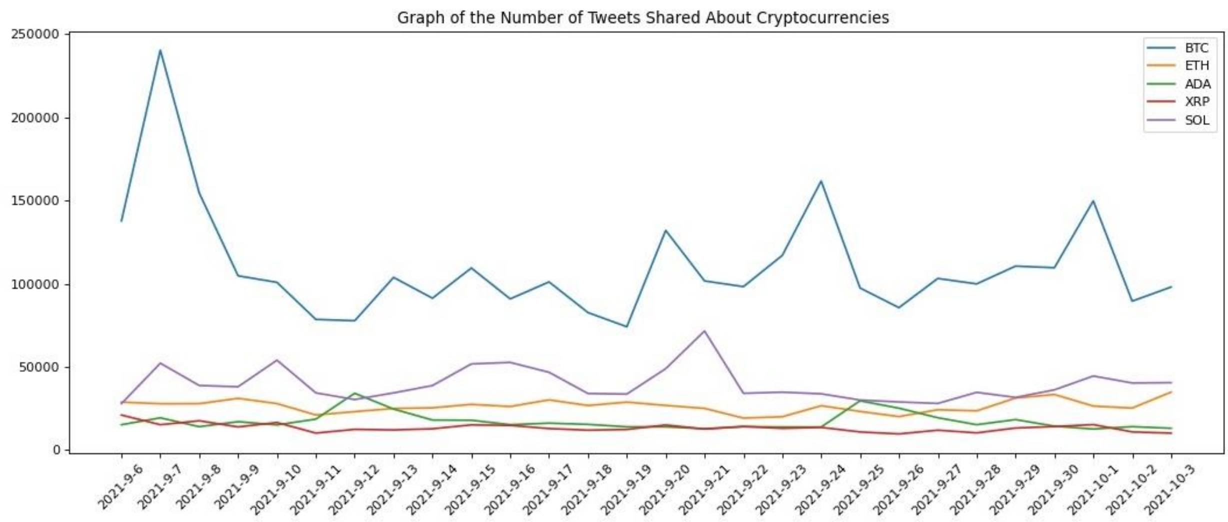

Figure 2. Daily tweet volumes of four cryptocurrencies between September 6 and October 3, 2021

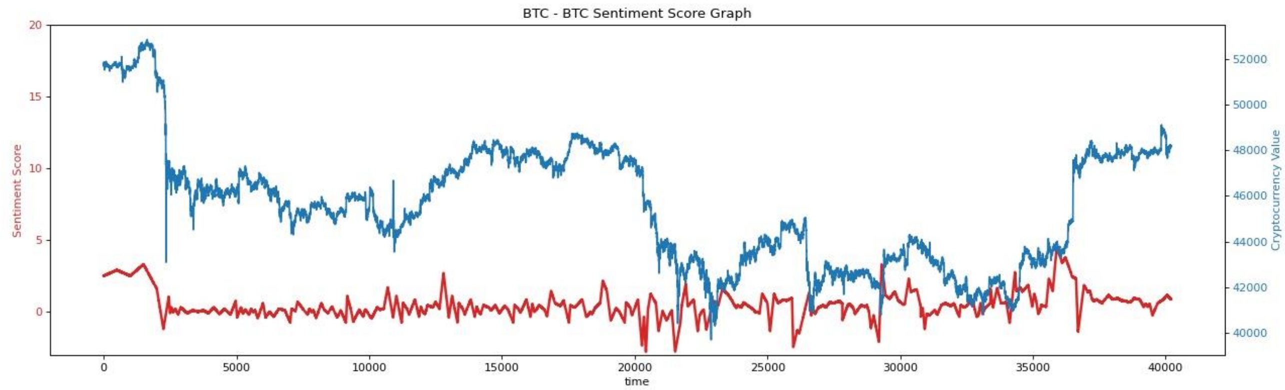

Figure 3. Twitter sentiment analysis score chart with Bitcoin closing data

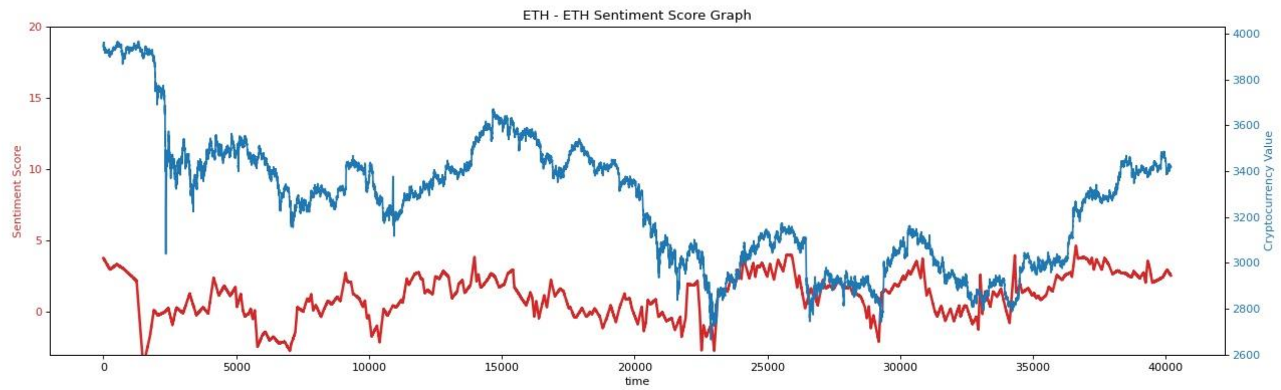

Figure 4. Twitter sentiment analysis score chart with Ethereum closing data

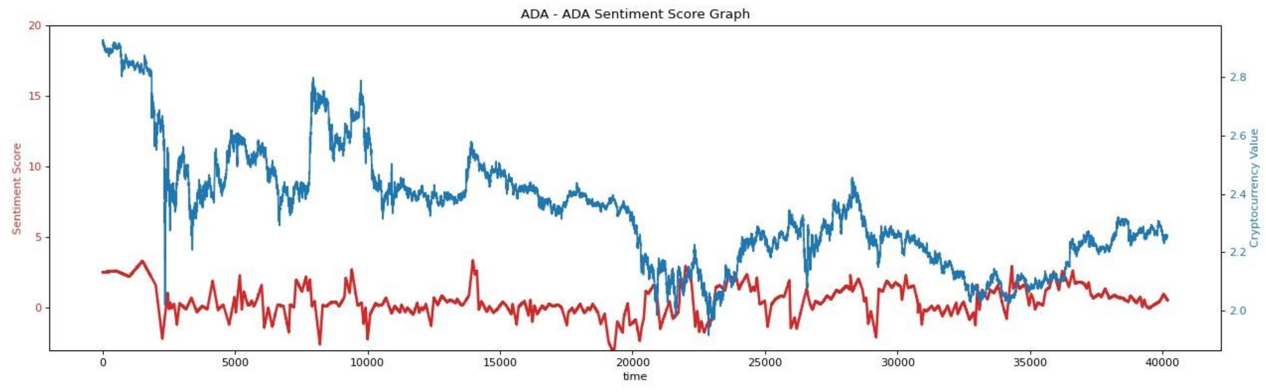

Figure 5. Twitter sentiment analysis score chart with Cardano closing data

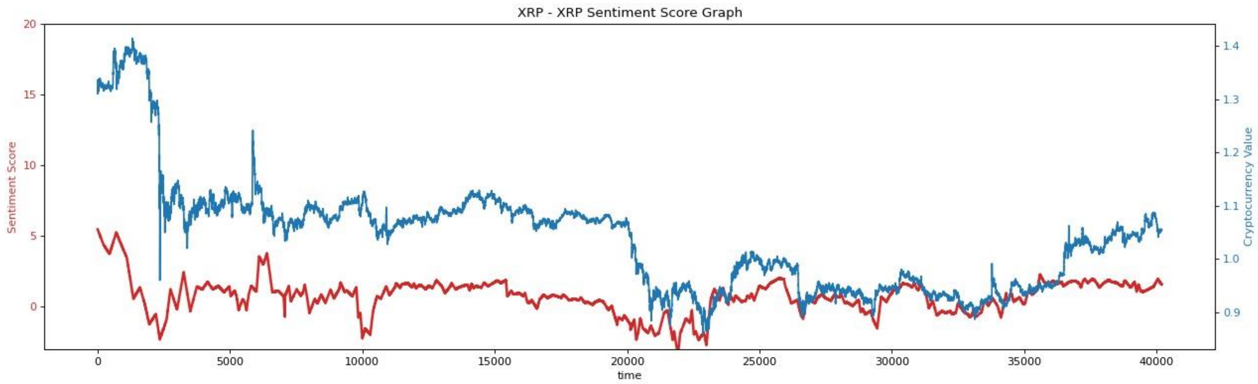

Figure 6. Twitter sentiment analysis score chart with XRP closing data

The total length of the data collected in 4 weeks (28 days x 24 hours x 60 minutes) is 40200. 65% of the whole data set was used as training data, and the remaining 35% as test data. According to these ratios, three different methods were trained and tested. Table 10 shows that the ARIMA method has a higher MSE & RMSE value for each cryptocurrency. Although the ARIMA method is a widely used method for estimating time series, the main reason for having high RMSE values is the volatility of cryptocurrencies. As it is known, the values of cryptocurrencies are not stationary time series. While creating the ARIMA model, p, d, and q parameters were selected as 0, 1 and 0, respectively, from the values showing the highest performance by trial. Examining the prediction performance using the ARIMA method, the primary purpose is to compare LSTM based on deep learning and MM-LSTM, another improved version of LSTM. Parameters of LSTM structures: number of layers is 2, and the number of hidden units is 7. The performances of the algorithms that are frequently used in regression problems (Decision Tree Regressor, Random Forest Regressor, Linear Regressor, Logistic Regressor, and Gaussian Process Regressor) are shown in Table 10. The analysis of the comparative performance of the algorithms is demonstrated with the MSE and RMSE results. According to these results, it is seen that the Twitter score (TS) approach created in equation 1 has an improving effect on the accuracy rate for each algorithm. In addition, the data were divided into 1-7-14 and 28-day parts, and the impact of the amount of data was wanted to be observed. The amount of data used in forecasts, such as cryptocurrency or dollar rate in the literature, is 4 weeks (28 days). Many of the results we obtained were higher in 4-week data.

In order to compare the method applied in this article with similar studies in the literature, a brief summary of the studies in the literature is given in Table 10. At the point when this table is analyzed, it is seen that the applied MM-LSTM structure has achieved results that can rival the LSTM-based approaches.

5. DISCUSSION

Various problems were encountered in the study of this article. First of all, the Twitter API used to collect Twitter data, which forms the main frame of the article, has a data limitation. Due to these limitations, various problems were encountered in accessing historical Twitter data. Despite these limitations, approximately 5 million tweets were analyzed. Another problem is the filtering of tweets shared by bot users. The tweets shared by bot users, which are used to be among the prominent hashtags on Twitter, are effective. Although Twitter is seen as a knowledge bank, it can also turn into a social media network that can cause manipulation with the shares of these bot accounts. Although various approaches are used to detect bot user accounts in this study, some bot accounts got rid of these filters, and some real users were stuck with these filters. For example, the most important of these filters is that the ratio of the number of followers to the number of people followed is close to zero. Since some bot accounts have been active for a long time, this rate has moved away from zero. Apart from this, since some real users share without caring about the number of followers, the ratio of the number of followers to the number of people followed is close to zero.

There are some problems with the filter for accounts in these two groups. Apart from these problems, users sometimes use more than one hashtag while sharing a tweet. Therefore, there are some overlapping tweets in the obtained dataset. The most important of the deficiencies in the studies in the literature is a weighting according to the effect of tweets. Not all tweets shared by users have the same effect. Factors such as the number of people following the user, the number of likes, and the number of retweets determine the interaction. Using these factors, a score was generated for each tweet. One-word tweets shared by users with high followers can be scored with this method. However, even if these tweets have a very high impact factor, they do not have enough words for sentiment analysis. In addition, it will not be possible to conduct sentiment analysis and use the weight value in Eq.1 from tweets (only links, hashtags, etc.) consisting of the parts cleaned in the pre-processing steps in Table 3.

Table 9. Performance Metrics Results of MM-LSTM with Twitter Score Data (RMSE Results)

Data	1 Day Data	3 Days Data	7 Days Data	14 Days Data	28 Days Data
BTC	7.253	21.803	18.625	9.322	**3.401**
ETH	6.406	18.211	12.376	6.601	**2.729**
ADA	3.782	12.384	9.107	4.598	**1.902**
XRP	2.625	10.206	6.294	2.713	**1.726**

Table 10. Performance metrics results of Twitter score data (RMSE results)

Data	ARIMA	Support Vector Regressor	Decision Tree Regressor	Random Forest Regressor	Linear Regressor	Logistic Regressor	Gaussian Process Regressor	Proposed Method MM-LSTM
	28-Days Data - Prediction Results Without Using Twitter Score Data							
BTC	117.529	19.305	16.263	16.329	27.264	21.407	22.345	13.352
ETH	88.762	16.347	13.229	13.283	14.645	15.346	15.963	12.412
ADA	62.904	10.029	8.156	8.297	9.803	8.206	9.056	5.456
XRP	24.203	4.971	4.294	4.362	4.614	3.742	4.391	4.634
	28-Days Data - Prediction Results With Using Twitter Score Data							
BTC	43.646	13.275	10.857	11.016	18.291	15.386	14.051	**3.401**
ETH	24.016	8.701	7.216	7.615	15.204	10.056	9.108	**2.729**
ADA	12.052	5.104	3.729	4.026	9.381	6.319	5.827	**1.902**
XRP	7.102	1.916	1.895	2.004	3.519	2.542	2.168	**1.726**

Table 11. Relative comparison for cryptocurrency value prediction

Objective	Methodology	Forecast Duration	Results
A study has been made on selecting the inputs in the LSTM structure. [62]	LSTM & AR	71 Days	RMSE: 247.33
A study has been made about the effect of the parameters used in estimating the Bitcoin value in deep learning methods. [63]	CNN, LSTM, GRU	1&3 Months	Gold Price: CNN RMSE: 201.34 LSTM RMSE: 151.67 GRU RMSE: 32.98 Twitter: LSTM RMSE: 32.98
ANN and LSTM method were used to find out how the price dynamics of cryptocurrencies changed in various time intervals. [64]	ANN, LSTM	1, 3, 5, 7, 14 Days	Min-MSE: ~2 Max-MSE: ~66
The performances of ARIMA, LSTM and GRU methods in time series forecasting are compared using only the opening values of cryptocurrencies. [65]	ARIMA, GRU, LSTM	492 Days	ARIMA RMSE: 302.53 LSTM RMSE: 603.68 GRU RMSE: 381.34
The performances of forecasting the price of different cryptocurrencies with the LSTM method were compared. [66]	LSTM	1, 10, 20, 30 Days	1 Day RMSE: 53.30 10 Days RMSE: 67.99 20 Days RMSE: 91.41 30 Days RMSE: 45.71
Proposed Method	ARIMA, LSTM, MM-LSTM	28 Days (4 Weeks)	Twitter Score and MM-LSTM: BTC RMSE: **3.401** ETH RMSE: **2.729** ADA RMSE: **1.902** XRP RMSE: **1.726**

Different amounts of data are used in the literature. Therefore, in this study, first of all, the methods used in the literature were used and it was determined that the most suitable model was MM-LSTM. Table 10 shows the results of different methods. Since these methods are not supported by Twitter data, they estimate with rather high errors. It has been determined that the best model is LSTM. The high volatility of cryptocurrencies negatively affects the success of machine learning methods and methods such as ARIMA. In the literature, it was seen that higher results were obtained in the estimations made using 4-week data. In order to test the accuracy of this, the performance of the MM-LSTM algorithm supported by the Twitter score was tested with 1-3-7-14 and 28-day data. The highest performances obtained are shown in Table 9. Data diversity can be increased by expanding the data collected in four weeks. But here, Twitter's service policy comes into play as a limiting effect. Twitter allows the analysis of 500,000 tweets for hobbyists, and 5 million tweets for academic research users, provided that they prove they are academic staff. In addition, there are not enough tweets about altcoins other than Ethereum, XRP and Cardano. Therefore, no agenda can affect people. Thus, the proposed approach may not be appropriate for cryptocurrencies other than these altcoins. The number of data for 4 weeks (28 days) is approaching 5 million. Therefore, the number of days is limited to 4 weeks. According to these results, the highest performance was obtained using 28 days (4 weeks) of data. Forecast using 1-day data has the second-best accuracy. Since there is less speculation about XRP, the prediction accuracy seems to be relatively high. These results show us that social media data such as Twitter can cause speculation. However, the estimation performance has been considerably improved thanks to the data created with the scoring system developed in this study.

While performing the sentiment analysis, the words are examined one by one. For this reason, typos or unique jargons in tweets affect the scoring. In addition, the model's input is increased by using the MM-LSTM structure instead of the classical LSTM structure. In this way, Twitter scores of cryptocurrencies, opening, closing, highest and lowest values in that minute can be used as the input of the proposed method. The study's primary purpose is to observe the effect of the calculated weight value for each tweet shown in Eq.1 on the RMSE results. Therefore, in the study, it is checked whether the prediction accuracy increases rather than the selection of the regression algorithm. It is observed that our VADER architecture-based weight value generation approach makes improvements in each algorithm. Therefore, future studies can focus on algorithm selection. When the results in Table 10 are examined, it is seen that the ARIMA model, which economists generally use, is unsuitable for cryptocurrency prediction. In addition, the LSTM architecture is a deep learning architecture developed for predicting and classifying time series. The MM-LSTM architecture developed based on LSTM has a lower RMSE value than the machine learning algorithms frequently used in regression studies in Table 11. This is important for the hybrid model development part, which can be applied to us in future studies. In particular, it gives us a clue that the basic algorithm that should be used should be a deep learning algorithm. However, Gaussian Process Regressor and Logistic Regressor methods, which are machine learning methods, have very successful results.

6. CONCLUSION

In this article, the scoring of Twitter sentiment analysis is used as another additional data that increases the performance of methods performed on predicting cryptocurrencies. When the studies in the literature were checked, it was seen that the tweets shared on Twitter had the power to affect other users. With these shared tweets, successful results have been obtained for predicting stock prices, elections and cryptocurrencies. With the Granger causality analysis, it is seen that one of the reasons for the fluctuations in the cryptocurrency market is the effect of tweets shared on Twitter. This article argues that the amount of influence of shared tweets is also a quality. Since it has a data line that can access historical data, the LSTM structure has a high predictive ability in time series data. This forecasting performance can be enhanced with supporting input data. Tweets shared on Twitter were used as supporting data. However, there are tweets (1% - 17%) shared by bot accounts. This article successfully reduced the RMSE value, one of the performance measures frequently used in time series forecasting studies, with the analyzed Twitter score data. This improvement rate is different for each cryptocurrency. Estimation was also made with the ARIMA method, the most widely used method of classical economics. The RMSE performances of these three different methods were compared.

The method applied in this article is open to improvement. For example, improvements can be made to make predictions in real-time in future studies. Furthermore, Twitter is not the only social media network that people interact with. For this reason, the scoring system used in this article can be developed by using data from different social media networks.

Conflict of Interest: No conflict of interest was declared by the authors.
Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

[1] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, Decentralized Business Review, 2008.

[2] H. Wang, D. Can, A. Kazemzadeh, F. Bar and S. Narayanan, A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle, Proceedings of the ACL 2012 System Demonstrations, pp. 115-120, 2012.

[3] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel and B. K. Lama, Recurrent neural network-based bitcoin price prediction by twitter sentiment analysis, 2018 IEEE 3rd International Conference on Computing, Communication and Security, Kathmandu, Nepal, (2018) 128-132.

[4] J. Bollen, H. Mao and X. Zeng, Twitter mood predicts the stock market, Journal of computational science, (2011) 2.1: 1-8.

[5] X. Li and C. A. Wang, The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin, Decision support systems, (2017) 95: 49-60.

[6] P. Ciaian, M. Rajcaniova and D. Kancs, The economics of Bitcoin price formation, Applied Economics, (2016) 48.19: 1799-1815.

[7] D. G. Baur, K. Hong and A. D. Lee, Bitcoin: Medium of Exchange or Speculative Assets?, Journal of International Financial Markets, Institutions and Money, (2018) 54: 177-189.

[8] A. H. Dyhrberg, Bitcoin, gold and the dollar–A GARCH volatility analysis, Finance Research Letters, (2016) 16: 85-92.

[9] N. Gandal, Price manipulation in the Bitcoin ecosystem, Journal of Monetary Economics, (2018) 95: 86-96.

[10] M. Baker and J. Wurgler, Investor sentiment and the cross-section of stock returns, The journal of Finance, (2006) 61.4: 1645-1680.

[11] Y. Sovbetov, Factors influencing cryptocurrency prices: Evidence from bitcoin, Ethereum, dash, litcoin, and monero, Journal of Economics and Financial Analysis, (2018) 2.2: 1-27.

[12] O. Poyser, Exploring the dynamics of Bitcoin's price: a Bayesian structural time series approach, Eurasian Economic Review, (2019) 9.1: 29-60.

[13] Z. H. Munim, M. H. Shakil and I. Alon, Next-day bitcoin price forecast, Journal of Risk and Financial Management, (2019) 12.2: 103.

[14] M. Polasik, A. I. Piotrowska, T. P. Wisniewski, R. Kotkowski and G. Lightfoot, Price fluctuations and the use of bitcoin: An empirical inquiry, International Journal of Electronic Commerce, (2015) 20.1: 9-49.

[15] T. Rao and S. Srivastava, Modeling movements in oil, gold, forex and market indices using search volume index and twitter sentiments, Proceedings of the 5th annual ACM Web science conference, Paris, France, (2013) 336-345.

[16] V. Karalevicius, N. Degrande and J. De Weerdt, Using sentiment analysis to predict interday Bitcoin price movements, The Journal of Risk Finance, (2018) 19.1: 56-75.

[17] D. Yermack, Is bitcoin a real currency? An economic appraisal, In: Handbook of digital currency. Academic Press, London, (2015) 31–43.

[18] G. P. Dwyer, The economics of Bitcoin and similar private digital currencies, J Financ Stab, (2015) 17:81–91.

[19] A. Cheung, E. Roca, J. J. Su, (2015) Crypto-currency bubbles: an application of the Phillips-Shi-Yu (2013) methodology on Mt.Gox Bitcoin prices. Appl Econ 47(23):2348–2358

[20] E. T. Cheah and J. Fry, Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin, Econ Lett, (2015) 130:32–36.

[21] F. Wen, L. Xu, G. Ouyang and G. Kou, Retail investor attention and stock price crash risk: Evidence from China, Int Rev Financ Anal, (2019) 65:101376.

[22] L. Kristoufek, BitCoin meets Google trends and wikipedia: quantifying the relationship between phenomena of the Internet era, Sci Rep, (2013) https://doi.org/10.1038/srep03415

[23] L. Kristoufek, What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis, PLoS ONE, (2015) 10(4):e0123923.

[24] T. Panagiotidis, T. Stengos and O. Vravosinos, On the determinants of bitcoin returns: a LASSO approach. Finance Res Lett, (2018) 27:235–240.

[25] E. Bouri, P. Molnár, G. Azzi, D. Roubaud and L. I. Hagfors, On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?, Finance Res Lett, (2017) 20:192–198.

[26] Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, T. H. Kim, S. J. Kang and C. H. Kim, Predicting fluctuations in cryptocurrency transactions based on user comments and replies, PLoS ONE, (2016) 11(8):e0161197.

[27] R. C. Phillips and D. Gorse, Predicting cryptocurrency price bubbles using social media data and epidemic modelling, In: 2017 IEEE symposium series on computational intelligence (SSCI), (2017) https://doi.org/10.1109/SSCI.2017.8280809

[28] C. Chen, L. Liu and N. Zhao, Fear sentiment, uncertainty, and bitcoin price dynamics: the case of COVID-19, Emerg Mark Finance Trade, (2020) 56(10):2298–2309.

[29] S. Ji, J. Kim and H. Im, A comparative study of bitcoin price prediction using deep learning, Mathematics, (2019) 7(10):898.

[30] T. A. Borges and R. F. Neves, Ensemble of machine learning algorithms for cryptocurrency investment with different data resampling methods, Appl Soft Comput, (2020) 90:106187.

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

[32] K. Żbikowski, Application of machine learning algorithms for bitcoin automated trading, In: Machine intelligence and big data in industry. Springer, Cham, (2016) 161–168.

[33] Z. Jiang and J. Liang, Cryptocurrency portfolio management with deep reinforcement learning, In: 2017 intelligent systems conference (intelliSys). IEEE, New York, (2017) 905–913.

[34] H. Jang and J. Lee, An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information, IEEE Access, (2018) 6:5427–5437.

[35] S. McNally, J. Roche and S Caton, Predicting the price of Bitcoin using machine learning, In: 2018 26th Euromicro international conference on parallel, distributed and network-based processing (PDP), IEEE, New York, (2018) 339–343.

[36] M. Nakano, A. Takahashi and S. Takahashi, Bitcoin technical trading with artificial neural network, Phys A, (2018) 510:587–609.

[37] A. Vo and C. Yost-Bremm, A high-frequency algorithmic trading strategy for cryptocurrency, J Comput Inf Syst, (2018) https://doi.org/10.1080/08874417.2018.1552090

[38] L. Alessandretti, A. ElBahrawy, L. M. Aiello and A. Baronchelli, Anticipating cryptocurrency prices using machine learning, Complexity, (2019) 2018:8983590.

[39] G. S. Atsalakis, I. G. Atsalaki, F. Pasiouras and C. Zopounidis, Bitcoin price forecasting with neuro-fuzzy techniques, Eur J Oper Res, (2019) 276(2):770–780.

[40] L. Catania, S. Grassi and F. Ravazzolo, Forecasting cryptocurrencies under model and parameter instability, Int J Forecast, (2019) 35(2):485–501.

[41] J. B. Han, S. H. Kim, M. H. Jang and K. S. Ri, Using genetic algorithm and NARX neural network to forecast daily bitcoin price, Comput Econ, (2019) https://doi.org/10.1007/s10614-019-09928-5

[42] S. Lahmiri and S. Bekiros, Cryptocurrency forecasting with deep learning chaotic neural networks, Chaos Solitons Fractals, (2019) 118:35–40.

[43] D. C. Mallqui and R. A. Fernandes, Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques, Appl Soft Comput, (2019) 75:596–606.

[44] Z. Chen, C. Li and W. Sun, Bitcoin price prediction using machine learning: an approach to sample dimension engineering, J Comput Appl Math, (2020) 365:112395.

[45] X. Sun, M. Liu and Z. Sima, A novel cryptocurrency price trend forecasting model based on LightGBM, Finance Res Lett, (2020) 32:101084.

[46] C. WJ. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica: journal of the Econometric Society, (1969) 424-438.

[47] L. Deng and D. Yu, Deep learning: methods and applications, Foundations and trends in signal processing, (2014) 7.3–4: 197-387.

[48] M. Shafahi, L. Kempers and H. Afsarmanesh, Phishing through social bots on Twitter, In 2016 IEEE international conference on big data, (2016) 3703-3712). IEEE.

[49] O. Varol, E. Ferrara, C. Davis, F. Menczer and A. Flammini, Online human-bot interactions: Detection, estimation, and characterization, In Proceedings of the international AAAI conference on web and social media, (2017) 11(1):280-289.

[50] Z. Chu, S. Gianvecchio, H. Wang and S. Jajodia, Detecting automation of twitter accounts: Are you a human, bot, or cyborg?, IEEE Transactions on dependable and secure computing, (2012) 9(6):811-824.

[51] Z. Gilani, E. Kochmar and J. Crowcroft, Classification of twitter accounts into automated agents and human users, In Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining, (2017) 489-496.

[52] V. S. Subrahmanian, A. Azaria, S. Durst, V. Kagan, A. Galstyan, K. Lerman and F. Menczer, The DARPA Twitter bot challenge, Computer, (2016) 49(6):38-46.

[53] N. Chavoshi, H. Hamooni and A. Mueen, Debot: Twitter bot detection via warped correlation, In Icdm, (2016) 817-822.

[54] C. Cai, L. Li and D. Zengi, Behavior enhanced deep bot detection in social media, In 2017 IEEE International Conference on Intelligence and Security Informatics (ISI), (2017) 128-130, IEEE.

[55] F. Morstatter, L. Wu, T. H. Nazer, K. M. Carley and H. Liu, A new approach to bot detection: striking the balance between precision and recall, In 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), (2016) 533-540, IEEE.

[56] D. Ciregan, U. Meier and J. Schmidhuber, Multi-column deep neural networks for image classification, 2012 IEEE conference on computer vision and pattern recognition, USA, (2012) 3642-3649.

[57] K. Fukushima and S. Miyake, Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, Competition and cooperation in neural nets, (1982) 267-285.

[58] F. Balcı, Z. Oralhan, LSTM ile EEG Tabanlı Kimliklendirme Sistemi Tasarımı, Avrupa Bilim ve Teknoloji Dergisi, (2020) 135-141.

[59] M. F. Stollenga, B. Wonmin, M. Liwicki and J. Schmidhuber, Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation, Advances in neural information processing systems, (2015) 28: 2998-3006.

[60] S. Song, H. Huang and T. Ruan, Abstractive text summarization using LSTM-CNN based deep learning, Multimedia Tools and Applications, (2019) 78.1: 857-875.

[61] Y. Yin, X. Zheng, B. Hu, Y. Zhang and X. Cui, EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM, Applied Soft Computing, (2021) 100: 106954.

[62] C. Wu, C. Lu, Y. Ma and R. Lu, A new forecasting framework for bitcoin price with LSTM, 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, (2018) 168-175.

[63] A. Aggarwal, I. Gupta, N. Garg and A. Goel, Deep learning approach to determine the impact of socio economic factors on bitcoin price prediction, 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, (2019) 1-5.

[64] W. Yiying and Z. Yeze, Cryptocurrency price analysis with artificial intelligence, 2019 5th International Conference on Information Management (ICIM), Cambridge, UK, (2019) 97-101.

[65] P. Yamak, L. Yujian and P. K. Gadosey, A comparison between arima, lstm, and gru for time series forecasting, 2nd International Conference on Algorithms, Computing and Artificial Intelligence, Sanya, China, (2019) 49-55.

[66] W. Zhengyang, L. Xingzhou, R. Jinjin and K. Jiaqing, Prediction of cryptocurrency price dynamics with multiple machine learning techniques, 4th International Conference on Machine Learning Technologies, Nanchang, China, (2019) 15-19.