Bitcoin Price Analysis with Deep learning models

Doddamani John Vestly
Department of Computer Science and Engineering
Vardhaman College of Engineering
Hyderabad, India
doddamanijohn123@gmail.com

Shanmugasundaram Hariharan
Department of Computer Science and Engineering
Vardhaman College of Engineering
Hyderabad, India
mailtos.hariharan@gmail.com

Neelima Thangallapelly
Department of Computer Science and Engineering
Vardhaman College of Engineering
Hyderabad, India
neelimathangallapelly@gmail.com

Vinay Kukreja
Department of Computer Science and Engineering
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
vinay.kukreja@chitkara.in

Gandra Sai Manish
Department of Computer Science and Engineering
Vardhaman College of Engineering
Hyderabad, India
saimanishgandra7485@gmail.com

Jayakshatha PR
Department of Information Technology
SRM Easwari Engineering College
Chennai, Tanilnadu, India
Jayakshata.prcseb@gmail.com

***Abstract*—Cryptography is used by cryptocurrencies to safeguard and authenticate transactions, as well as to regulate the creation of new units. Cryptocurrencies, as opposed to traditional money, which is often issued by a central authority like a government are decentralized meaning they are not managed by any one entity or group. One of the most well cryptocurrency is "Bitcoin". Cryptography is used by Bitcoin, a decentralized digital currency, to safeguard transactions and regulate the generation of new units. It is exchanged on different cryptocurrency exchanges and runs without the help of a central bank. However the future currency is the ultimate digital currency. In this paper we are going to inculcate with the neural network models like sequence-to-sequence, ARIMA(Auto Regressive Integrated Moving Average), LSTM(Long Short Term Model). This process is certainly made to analyze the bitcoin price over the past couple of years. We trained our model accordingly and found out 96.21% accuracy for the sequence-to-sequence and 94.2% for the ARIMA as well. Eventually this paper and its work insinuate regarding the highly accurate performing algorithm and its analysis as well**

Keywords— Cryptocurrency, Bitcoin, Neural Network, Digital Currency, accuracy, Sequence-to-Sequence, ARIMA

I. INTRODUCTION

Digital or virtual currency known as cryptocurrency uses cryptography for security and is not controlled by a central bank or government. To keep track of transactions and manage the generation of new units, it makes use of blockchain technology [1]. Due to their decentralized structure and possibility for transaction secrecy, cryptocurrencies have grown in favuor as a replacement for traditional money. There are millions of cryptocurrencies, with Bitcoin being the most well-known. Other coins include Ethereum, Litecoin, and Ripple [2].

As a new type of money that runs decentralized, Bitcoin has drawn the interest of investors, traders, and the general public in recent years [3]. Cryptography is used by Bitcoin, a decentralized digital currency, to safeguard transactions and regulate the generation of new units. It was developed in 2009 as an alternative to conventional money that runs without the help of a centralized organization like a government or bank. The use of encryption makes the Bitcoin network secure. As a means of payment for products and services as well as a speculative investment, bitcoin has grown in popularity. Many people find it to be an appealing investment option because of its distinctive qualities, but the price volatility of Bitcoin continues to be of concern [4]. Over the past several years, there have been huge variations in the price of Bitcoin, which has fluctuated from less than $1,000 in early 2017 to about $65,000 in April 2023. Making educated judgments regarding whether to purchase, sell, or keep Bitcoin can be difficult for investors and traders due to the unpredictability of its price movement [5]. Therefore, it is essential to create trustworthy techniques for assessing Bitcoin's price movement in order to forecast its future value [6]. Investors and traders are nonetheless concerned about its price volatility. Therefore, it is necessary to create trustworthy techniques for assessing Bitcoin's price changes and forecasting its future value [7].

Decentralized currencies have the potential to upend established financial systems, making them a desirable option for people and companies seeking an alternative to conventional banking [8]. Bitcoin can be used as a form of payment for products and services, and it has benefits over conventional payment methods such quicker and less expensive transactions [9]. Blockchain, the technology that

underpins it, has also demonstrated potential for usage in a number of sectors outside of finance, including voting systems and supply chain management[10]. Additionally, as Bitcoin has grown, new industries have emerged, including mining and cryptocurrency exchanges [11].

The analysis and forecasting of Bitcoin price movements have benefited greatly from the use of neural networks. These algorithms can recognise patterns and trends that are difficult to spot using conventional statistical techniques by learning from historical pricing data [12]. Researchers have been able to develop models for forecasting future price movements that are more precise thanks to the usage of neural networks in Bitcoin price analysis. Long Short-Term Memory (LSTM) and Sequence-to-Sequence (Seq2Seq) models are two typical neural network models used in Bitcoin price analysis. Although developing and improving these models can be difficult, the possibility for precise price forecasts has made them an important tool in the cryptocurrency sector [13].

The paper is organized in a scholarly way by presenting an introduction in section 1. Section 2 presents literature survey and Section 3 deals with basic flow of road navigation system, Section 4 outlining the results and discussion. Section 5 gives the conclusion and future scope.

II. Related Work

Now-a-days everything can be resolved through perfect analysis and differentiation. At some point the comparison [1] is made between sliding window and whole set normalization techniques for forecasting daily Bitcoin OHLC prices using LSTM and GRU neural networks.In order to determine the Stackelberg equilibrium, a resource pricing and scheduling problem was modeled as a multi-leader, multi-follower Stackelberg game with three stages.To show the huge performance gain achieved by the suggested method,several simulations were run [2]. The consequences of adding new cryptocurrency mining loads to the Texas power grid are examined in this research,[3] as well as the possible financial gains from utilizing demand flexibility provided by cryptocurrency mining operations in the electrical market.

The annual revenues of bitcoin mining machines engaging in various demand response programmes for data centers are analyzed. The objective of this project [4] is to create a system that can gauge the sentiment of input and anticipate the price of a cryptocurrency(Bitcoin). The Cryptonic API,which will retrieve the most recent Bitcoin-related news, will be used to provide this input to the model.The performance of alternative algorithms, such as CNN, GRU and RNN was compared before we made the decision to develop this model using LSTM.

An analytical approach is made by [5] presenting a neural network-based autoregressive model that uses 100 historical observations of bitcoin prices. The suggested method uses a nonlinear autoregressive neural network with external input (NARX) specifically by reserving the maximum regression coefficient corresponding to the greatest prediction accuracy. Another study consisted of using [6] sentiment analysis on tweet data acquired from Twitter along with a deep neural network model, LSTM, and historical bitcoin values.

Sentiment analysis, on the other hand, looks into sentiment on Twitter to ascertain how sentiment and changes in bitcoin prices are related.In order to attempt to make short-term forecasts about the price of the currency, the aim of this study [7] is to go and identify the influences of the Bitcoin market through the analysis of the opinions and feelings of the vast communities that belong to the social network Reddit. The findings demonstrate the viability of a Recurrent Radial Basis Function Network (RRBFN)-based method for performing digital asset prediction starting with sentiment analysis of online chats.

The price of bitcoin has [8] undergone significant variations throughout time, choosing eleven alternative regression models, examining these models, and then determining which regression-based model is the most accurate at predicting bitcoin prices. The study's findings show that the Bayesian ridge regressor performs better than every other regression-based model except for the linear regressor. The effectiveness of four different machine [10] learning models for forecasting the return rate and price trend of Bitcoin is compared. Four input feature sets are used to form the data:historical bitcoin exchange data. Random forest, Decision Tree, AdaBoost, and Support Vector Machine are the four machine learning models that were applied. [9].

The efficacy of four different machine learning models on forecasting the Bitcoin return rate and price trend using data from the Bitcoin exchange, COVID-19, and Twitter from January 2020 to July 2020. Four input feature sets are used to form the data: historical bitcoin exchange data, historical bitcoin exchange data plus COVID-19 data (recovery, confirmed, death), historical bitcoin exchange data plus Twitter, and historical bitcoin exchange data plus COVID-19 data plus Twitter. Random forest, Decision Tree, AdaBoost, and Support Vector Machine are the four machine learning models [10] that were applied.

III. Methodology

A. Data Collection

The data for experimental illustrations is obtained from Kaggle (2018). Historical Bitcoin data taken from kaggle [19]. For the proposed study on the analysis and forecasting of the price of Bitcoin. The dataset includes past Bitcoin prices and the columns(Attributes) as Timestamp, Open, High, Low, Close, Volume as well as different information like trade activity and market capitalization. We sought to determine the method for forecasting Bitcoin prices that was most accurate by examining the performance of a couple of models. For the validity of our findings, the use of a trustworthy dataset is essential and Kaggle offers a practical and dependable source of data.

B. Data Pre-processing

The MinMaxScaler technique was applied to the research's data preprocessing stage to normalize the data. This method rescales the data to a predetermined range between 0 and 1. It

was picked because of how easy it is to use and how well it preserves the distribution of the original data. The neural network models were then trained and tested using the scaled data. Overall, it was found that the MinMaxScaler method was a useful preprocessing step for the analysis and forecasting of Bitcoin prices.

C. Basic algorithm & background:

The model in this paper are chosen to inculcate with output and proposed couple of models are as follow as:

i. LSTM:

Recurrent neural networks (RNNs) of the sort known as Long Short-Term Memory (LSTM) have received a lot of interest lately because of their ability to efficiently model sequential input. Due to its capability to recognise long-term dependencies in data, LSTM has been extensively used in a variety of domains, including natural language processing, speech recognition, and time-series analysis. To model the time-series data of Bitcoin prices and capture complicated patterns that conventional statistical models could miss in the context of Bitcoin price prediction, LSTM can be a valuable tool. By learning and remembering long-term dependencies, LSTM can aid in the creation of more precise models for forecasting future Bitcoin values. In this study, we examine LSTM's performance in predicting Bitcoin values and compare it with other time-series models.

ii. ARIMA:

Popular time-series forecasting models, such as ARIMA (AutoRegressive Integrated Moving Average), are utilized in a variety of disciplines, including finance, economics, and engineering. It is a statistical method that combines moving average, differencing, and auto regression techniques to analyze and forecast time-series data. ARIMA models are helpful for estimating future values because they can spot patterns and trends in time-series data. The moving average component, the integrated component, and the autoregressive component make up the model. The analysis and forecasting of price changes for financial assets, including virtual currencies like Bitcoin, is frequently done using the ARIMA model.

iii. Sequence-to-sequence (Seq2seq):

The neural network design known as sequence-to-sequence (Seq2Seq) is frequently used in time series analysis, audio recognition, and natural language processing. Seq2Seq models may process sequential input data of different durations and produce output data that is also sequential. Seq2Seq models have been used in the context of Bitcoin price analysis and prediction to estimate future prices based on historical price data.

Seq2Seq models' capacity to capture intricate interactions between variables and their dynamic patterns makes them much superior to conventional time series analysis methods in this regard. Furthermore, Seq2Seq models have the capacity to produce probabilistic forecasts, offering a range of potential outcomes as opposed to a single point estimate. For Seq2Seq models to be useful in analyzing and predicting Bitcoin price, they must be carefully tuned and validated, just like any other model. The pseudo algorithm for Seq2seq is presented in Fig.1.

i.Evaluation metrics:

The evaluation metrics are the most crucial part of any system and architecture. Eventually in this work we used a couple of evaluation metrics in order to evaluate the models we built. The metrics are as follow as:

a. Mean Squared Error (MSE):

A popular statistic for assessing the precision of regression models, particularly those used to forecast Bitcoin prices, is the Mean Squared Error (MSE). Lower values represent a better fit between the model and the data, and it estimates the average squared difference between the predicted and actual values. In the context of predicting the price of Bitcoin, MSE can be used to assess the performance of different machine learning models and compare how well they perform in this regard is presented in Equation 1.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \quad (Y_i - \hat{Y}_i \quad)Y2 \tag{1}$$

Where: = Mean Squared Error, = number of data points, = observed values, = predicted values.

b. MAPE (Mean Absolute Percentage Error)

It is a frequently used evaluation statistic in time series research and forecasting. The average absolute percentage difference between actual and anticipated values is measured by MAPE (presented in equation 2). A lower MAPE indicates better model accuracy; it is represented as a percentage. Although MAPE has certain drawbacks, such as being sensitive to extreme values and undefined for actual values equal to zero, it is nevertheless useful for comparing forecasting performance across various time series datasets and models.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \quad \left|\frac{A_t - F_t}{A_t}\right| \tag{2}$$

Where M = Mean Absolute Percentage Error, n = number of data points, A_t = The actual value of the observation, F_t = The predicted value for the observation.

IV. Results And Discussions

The work which we performed demonstrated a set of results. This results section includes the information regarding the evaluation metrics and the visualization aspect as well.

Algorithm 1: Seq2Seq

Input: Source sequence **X = [x1, x2, ..., xn]**, target sequence **Y = [y1, y2, ..., ym]**

Output: Predicted target sequence **Y' = [y'1, y'2, ..., y'k]**

1. Initialize the encoder and decoder neural network models with the parameters θ**enc** and θ**dec**.
2. Encode the source sequence **X** into a sequence of hidden states **H** using the encoder model with parameters θ**enc**.
3. Initialize the decoder hidden state **s_0** to be the last encoder hidden state **h_n**.
4. For each time step t in the output sequence **Y:**
   - **i.** Compute the context vector **c_t** by applying an attention mechanism to the encoder hidden states **H** and the current decoder hidden states **s_{t-1}**.
   - **j.** Compute the decoder hidden states **s_t** and output distribution **P(y_t/y_{<t}, X)** using the decoder model with attention parameters θdec.
   - **k**. Sample the predicted token **y'_t** from the output distribution **P(y_t/y{<t}, X)** using a sampling technique, such as beam search or greedy search.
   - **l.** Update the decoder input **y_t** to be the predicted token **y'_t**.
   - **m.** Append the predicted token **y'_t** to the predicted target sequence **Y'**.
5. Return the predicted target sequence **Y'**.

Fig.1: Seq2seq algorithm psudeo procedure

Table 1 presents the performance of three distinct models—LSTM, ARIMA, and Seq2Seq—for forecasting Bitcoin prices is displayed in the accuracy table. The results show that the Seq2Seq model works well with an accuracy of 96.21%, while the ARIMA model obtains the accuracy with an astounding 94.2%. The LSTM model, on the other hand, has the lowest accuracy at 87%. These results show that the LSTM model may not be as useful for predicting Bitcoin prices as the ARIMA and Seq2Seq models. In order to choose the best model, future research can examine how well these models work in various settings. A popular statistical model for financial forecasting and time-series analysis is the ARIMA model. Deep learning-based Seq2Seq modeling has grown in popularity in recent years due to its capacity to represent intricate temporal connections.

TABLE I. PERFORMANCE METRICS OF DIFFERENT MODELS

Model	Accuracy
LSTM	87.34%
ARIMA	94.26%
Seq2Seq	96.21%

TABLE II. EVALUATION METRICS TABLE

Model	MSE	MAPE
LSTM	0.176	1.19
ARIMA	6.28	0.21
Seq2Seq	0.15	0.18

For a certain assignment, the table includes the assessment metrics of three different models: LSTM, ARIMA, and Seq2Seq. Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) are the metrics used to assess the models. The LSTM model showed the minimum average squared error between the predicted values and the actual values, with an MSE value of 0.176. However, it had the highest MAPE value (1.19), indicating that the absolute error rate of its forecasts was rather significant.

The ARIMA model, in contrast, had a higher MSE score of 6.28, suggesting that its average squared error was higher than that of the LSTM model. However, it had a lower MAPE value of 0.21, indicating that the absolute error rate of its predictions was rather low.

The Seq2Seq model's average squared error was 0.15, which is considered moderate and places it between the LSTM and ARIMA models. It had the lowest MAPE score, 0.18, indicating that the percentage of absolute mistakes in its predictions was relatively low.

In conclusion, the LSTM model did well in minimizing squared errors but had relatively large absolute errors in its predictions. It had the lowest MSE but the greatest MAPE. The ARIMA model fared best in minimizing absolute errors but had a greater average squared error, as indicated by its lower MAPE but higher MSE. The Seq2Seq model has the best performance in terms of both measures, as evidenced by its moderate MSE and minimal MAPE.

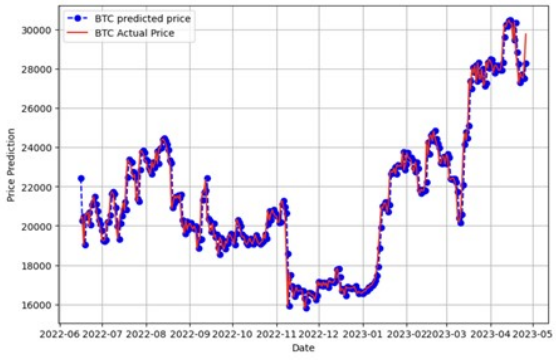

Fig.2: Bitcoin price prediction

Fig.-2 depicts the Bitcoin predicted price and the Bitcoin actual price for the couple of years. The ARIMA model has got certain visualization and it's made an output through the above graph. The predicted price is very keen and close to the actual price. This made a great significance of it. Fig.3 certainly describes the closing Prices and the dates of the training data and testing data as well. This clearly depicts that the closing prices vary upon the certain features and the circumstance actions as well. The ARIMA model is trained in a certain way to get the difference between the training and the testing data.

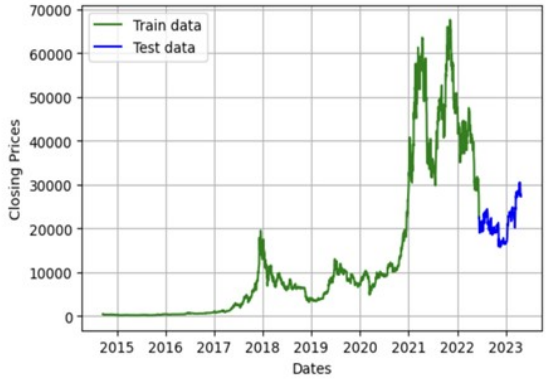

Fig.3: Train data v/s Test data of ARIMA model

Fig.4 clearly describes the comparison between the Closing Prices and dates using the Seq2Seq model. This would certainly assist us with some accurate instances when compared to the ARIMA model. The price took a rise during the year 2021 and 2022.

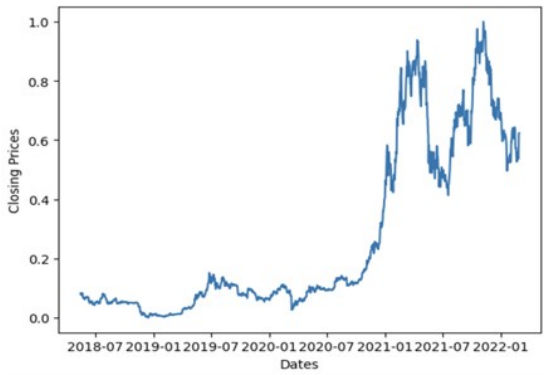

Fig.4: Closing Price v/s Dates of Seq2Seq mode.

The actual and accurate output was certainly made by the Seq2Seq model i.e., from the above visualization aspect. When the training data and the testing data was compared in the aspects of the Closing price and Dates there was a clear spike during the years 2021 and 2022 as well. Fig.5 presents the training and test data of Seq2seq model.

Fig.5: Train data v/s Test data of Seq2Seq model

V. CONCLUSIONS

In this study, combination of methods to examine the price of Bitcoin is observed over the previous few years. In order to forecast Bitcoin's price, we specifically used the sequence-to-sequence model, ARIMA, and LSTM. According to the Evaluation metrics, that is when they are compared the overall credentials were very favorable to the Seq2Seq model. Yes, it is certainly very effective for prediction and to analyze time series forecasting problems. The accuracy of the Seq2Seq model is 96.21%. This depicts and suggests making use of the model in a very certain way to invest and gain in the cryptocurrency market and agencies as well. The Seq2Seq model's greater accuracy rate can be attributed to its capacity to recognise the sequential nature of Bitcoin price data. Additionally, the study has demonstrated that data preprocessing methods like MinMaxScaler can significantly enhance the models' performance. The accuracy of the models has also been assessed using assessment metrics from the study including MSE and MAPE. However, it is crucial to remember that the cryptocurrency market is extremely unpredictable and volatile, so the forecasts provided by the models should only be used with extreme caution

REFERENCES

.

[1] S. Simtharakao and D. Sutivong, "Exploring Normalization Techniques in Neural Networks for Bitcoin Candlestick Price Prediction," 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Bali, Indonesia, 2023, pp. 483-488, doi: 10.1109/ICAIIC57133.2023.10067086.

[2] S. Huang, H. Huang, G. Gao, Y. -E. Sun, Y. Du and J. Wu, "Edge Resource Pricing and Scheduling for Blockchain: A Stackelberg Game Approach," in IEEE Transactions on Services Computing, vol. 16, no. 2, pp. 1093-1106, 1 March-April 2023, doi: 10.1109/TSC.2022.3177438.

[3] A. Menati, K. Lee and L. Xie, "Modeling and Analysis of Utilizing Cryptocurrency Mining for Demand Flexibility in Electric Energy Systems: A Synthetic Texas Grid Case Study," in IEEE Transactions on Energy Markets, Policy and Regulation, vol. 1, no. 1, pp. 1-10, March 2023, doi: 10.1109/TEMPR.2022.3230953.

[4] K. Nair, A. Pawle, A. Trisal and S. Krishnan, "Bitcoin Price Prediction Using Sentimental Analysis - A Comparative Study of Neural Network Model for Price Prediction," 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 2022, pp. 1-4, doi: 10.1109/ASIANCON55314.2022.9908846.

[5] Q. A. Al-Haija, "Time-Series Analysis of Cryptocurrency Price: Bitcoin as a Case Study," 2022 International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT), Jember, Indonesia, 2022, pp. 49-53, doi: 10.1109/ICEECIT55908.2022.10030536.

[6] M. H. Bin Mohd Sabri, A. Muneer and S. M. Taib, "Cryptocurrency Price Prediction using Long Short-Term Memory and Twitter Sentiment Analysis," 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, Pune, India, 2022, pp. 1-6, doi: 10.1109/ICCUBEA54992.2022.10011090.

[7] M. Casillo, M. Lombardi, A. Lorusso, F. Marongiu, D. Santaniello and C. Valentino, "Sentiment Analysis and Recurrent Radial Basis Function Network for Bitcoin Price Prediction," 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON), Palermo, Italy, 2022, pp. 1189-1193, doi: 10.1109/MELECON53508.2022.9842889.

[8] K. Soni and S. Singh, "Bitcoin Price Prediction- an Analysis of Various Regression Methods," 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Penang, Malaysia, 2022, pp. 271-276, doi: 10.1109/ISCAIE54458.2022.9794532.

[9] T. Li, "Prediction of Bitcoin Price Based on LSTM," 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Guangzhou, China, 2022, pp. 19-23, doi: 10.1109/MLISE57402.2022.00012.

[10] J. Luo, "Bitcoin price prediction in the time of COVID-19," 2020 Management Science Informatization and Economic Innovation Development Conference (MSIEID), Guangzhou, China, 2020, pp. 243-247, doi: 10.1109/MSIEID52046.2020.00050.

[11] 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, pp. 1-5, doi: 10.1109/IC3.2019.8844928.

[12] S. Velankar, S. Valecha and S. Maji, "Bitcoin price prediction using machine learning," 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, Korea (South), 2018, pp. 144-147, doi: 10.23919/ICACT.2018.8323676.

[13] B. Malhotra, C. Chandwani, P. Agarwala and S. Mann, "Bitcoin Price Prediction Using Machine Learning and Deep Learning Algorithms," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-6, doi: 10.1109/ICRITO56286.2022.9964677.

[14] A. Singh, A. Kumar and Z. Akhtar, "Bitcoin Price Prediction: A Deep Learning Approach," 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2021, pp. 1053-1058, doi: 10.1109/SPIN52536.2021.9565988.

[15] I. Ntourmas and D. Sotiropoulos, "Bitcoin Price Prediction and Automated Trading via LSTM Networks and Reinforcement Learning," 2022 13th International Conference on Information, Intelligence, Systems & Applications (IISA), Corfu, Greece, 2022, pp. 1-5, doi: 10.1109/IISA56318.2022.9904355.

[16] A. Demir, B. N. Akılotu, Z. Kadiroğlu and A. Şengür, "Bitcoin Price Prediction Using Machine Learning Methods," 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara, Turkey, 2019, pp. 1-4, doi: 10.1109/UBMYK48245.2019.8965445.

[17] 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 (ICCCS), Kathmandu, Nepal, 2018, pp. 128-132, doi: 10.1109/CCCS.2018.8586824.

[18] H. Kavitha, U. K. Sinha and S. S. Jain, "Performance Evaluation of Machine Learning Algorithms for Bitcoin Price Prediction," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2020, pp. 110-114, doi: 10.1109/ICISC47916.2020.9171147.

[19] https://www.kaggle.com/mczielinski/bitcoin-historical-data (last accessed 05.07.2023)