Time-Series Real-Time Bitcoin Price Prediction using LSTM

Omkar Bhattarai	Raj Chaudhary	Ganesh Gupta
Department of the Computer Science and Engineering, Sharda University Greater Noida, India	Department of the Computer Science and Engineering, Sharda University Greater Noida, India	Department of the Computer Science and Engineering, Sharda University Greater Noida, India

ABSTRACT

Bitcoin, the biggest decentralized cryptocurrency till date, is an extremely volatile digital currency in which new forms/units of currency are generated by the mechanical/computational solution of mathematical problems. In recent years its popularity can't be ignored as it is very significant and evolutionary in its own way and its price has shown high volatility. Due to its volatility and instability of prices, there is a high risk of investing in it if ones don't have any knowledge about it. Bigger players can manipulate the market or they barely go into losses. But for smaller players, it's a high chance to lose their hard-earned money. Therefore, it has become necessary to develop an efficient and more accurate predictive model for forecasting bitcoin prices. Accurate prediction of its price is of great interest to investors and traders as well. In this research paper, Long Short-Term Memory (LSTM), a type of Recurrent Neural Network (RNN), has been used for predicting Bitcoin time series data. We have used real-time data on an hourly basis. Bitcoin, a decentralized digital currency, We collected and pre-processed historical Bitcoin price data from the coin desk, an API for cryptocurrency live data dashboard, and split it into training and testing sets. We have designed an LSTM model and trained it on the training data. We evaluate the performance of the model on the testing data and use it to make predictions on new Bitcoin price data. The experimental results of our experiment demonstrate that the LSTM model outperforms other traditional time series models and provides a promising approach for Bitcoin price prediction with 96.88% accuracy for new data.

Keywords- Bitcoin, LSTM, RNN, Cryptocurrency;

1. INTRODUCTION

The cryptocurrency market is very popular in today's world. Although it is not very tranquil. There is always something happening in the crypto world. The growth of this market is something that nobody had imagined. It has grown so big in an unpredictable way. January 2009, is the date when bitcoin was first publically released by an anonymous developer using the name Satoshi Nakamoto. Since its release, it has become the most well-known cryptocurrency in the world. As of now(2023), after the bitcoin was released, over 22000 other cryptocurrencies have been developed but unfortunately, very few cryptocurrencies have become successful. As for bitcoin it had seen its peak in 2021($68770). After that its price started declining. Cryptocurrency is the game changer of this new world as not any central bank entity available to control and manage the cryptocurrency asset, unlike the flat currency[18]. People are making huge money investing in cryptos. As of 2023, there are about 420 million global crypto users, still, most people are taking chances with it but are very afraid as cryptocurrencies are not very stable and forgiving. Many people made millions but many nearly went bankrupt due to its unpredictable nature[16].

Due to its rapid growth in the last 5 years, it has taken the financial world by storm. As a result, the enthusiasm for in investing cryptocurrency is growing day by day. Covid has played a big role in attracting new generational risk-tolerant investors prepared to play in the crypto market [17]. To help understand the nature of cryptocurrencies and give them a push forward much open research is required to be done which has been absent in this field. Bitcoin is an intriguing subject since it is a time series prediction issue in a market that is still in its early stages.

Conventional time series prediction techniques, like the HoltWinters exponential smoothing models, rely on linear hypotheses and call for data that may be categorized as trend, seasonal, or noise (Chatfield and Yar, 1988). This method works best for projects that involve seasonal influences, such as estimating sales Because of the Bitcoin market's extreme volatility and lack of seasonality, these strategies are unsuccessful. Deep learning is an interesting technical alternative given the complexity of the challenge, based on its success in related domains (McNally et al., 2018).

Therefore, this paper seeks to assist financial experts, students, and cryptocurrency investors in selecting the optimal price prediction model for these coins Because Bitcoin (BTC) has the biggest market capitalization as of 2023. Hari Krishnan Andi, [15] The paper has used Linear Regression, Lasso Regression and compared it with the logistic regression with LSTM which has the increased accuracy to 97.2% from 89.1% in Linear Regression and 91.2% in Lasso algorithm. Precision has also shown the same rate of increment. However they have used the smaller dataset, therefore their future work will be of large dataset.

Maham Chaudhary et. al[19] predicted the trend and direction of cryptocurrency market through ML. Granularity, accuracy of Different Machine models were used and also surveyed different researched papers. Because there was no benchmark model, comparing the various article models was challenging[14]. Andi et. al[9] used large dataset with LSTM model, but has over-fitting issues. Suhwan Ji et al[5] LSTM, CNN, DNN, ResNet, CRNN and different Ensemble models were used, but few had outperformed the current model. Do-Hyun Kwon et al. [6] employed the LSTM model once again to forecast the price of bitcoin, but they had trouble identifying the performance distribution and tendency in relation to time units.

Paper Structure: Section 2 focuses on the related studies and work made in the same field and the contribution made by them. Section 3 explores the methodology proposed in this research paper which will explain our working model, section 4 will explain the experimental findings and results, section 5

will explain the conclusions and section 6 will be of references of research papers.

2. RELATED WORK

E. Sin and L. Wang use a Genetic Algorithm-based Selective Neural Network Ensemble technique using a Multi-Layered Perceptron as the basic model for each neural network in the ensemble to investigate the relationship between Bitcoin features and the difference in its price the following day. The data set was than retrieved directly from Block-chain.info. The Levenberg-Marquardt (LM) algorithm was used to train all of the MLPs due to the difficulty of the ensemble and the computational cost of back-testing[1].

S. Velankar et al. Used Bayesian Regression and GLM(Generalized Linear Model) or Random Forest model for the model construction and testing where they used simple pre-processing techniques to refine the data thereby focusing mainly on normalization. They have used Quandl and CoinMarketCap data sets for their projects[2]. Amin Azari et al. Applied Auto-regressive Integrative Moving Average, called ARIMA approach to forcast the price of crypto-currency bitcoin in their research paper. Their approach is based on the dataset from 3 years-long periods, during which the price of bitcoin had experienced fluctuating behavior, but large prediction errors were introduced. Their study highlights the interaction of the three things: prediction accuracy, choice of (p, q, d), and window size 'w'[3].

K. Rathan et al. Their paper applied Regression and Regression techniques to predict the Crypto-Currency price. Bitcoin dataset was considered from 2011 to 2019 price and above machine learning methods was applied. Linear regression outperformed the other, but accuracy didn't meet the expected level as it is supposed to be[4]. K.Raju, S. M., et al.[20] investigated several machine learning methods employing supervised learning to create a prediction model and offer insightful analyses of potential market price changes. It is frequently quite difficult to make suitable forecasts because it is difficult to assess the precise nature of a Time Series(ARIMA) model. The implementation of recurrent neural networks (RNN) with long short-term memory cells (LSTM) was then continued. They compared the predictability of bitcoin price and sentiment analysis of bitcoin tweets to the conventional method (ARIMA) and used long short-term memory (LSTM) techniques to assess the time series model prediction of bitcoin prices more effectively. But given the vast range of bitcoin prices and their volatility, it wasn't enough.

R. Albariqi et al. Presented the early neural network models for predicting both short- and long-term changes in Bitcoin prices in their study. The Multi-layer Perceptron (MLP) and Recurrent Neural Network models are employed (RNN). A total of 1300 data points covering the period from August 2010 to October 2017 were gathered from the blockchain of Bitcoin for the study. Each data point was separated by two days. The models created can forecast price changes from two days to sixty days in the future. MLP had more accuracy(81.3%) than the RNN model. But, accuracy could be improved[5].

Nimkova et al. The body of their paper offers convincing proof of the advantages of using COVID-19 data to forecast bitcoin prices. According to their analysis, COVID-19 had a considerable impact on cryptocurrency prices. In their research article, Niamkova et al. used methods like random forest and LSTM model. To respond to their problem statement, they used analytical criteria, such as previous bitcoin values, the source of the COVID-19 data, investor opinion, etc. According to the

paper's conclusion, financial managers might be able to benefit from an increase in the model's predictability.[6]

Taryanto et al. The primary portion of their paper discusses the magnitude of the present and prospective markets for cryptocurrencies. Price forecasting is therefore a hot topic in today's culture. However, given the reality of price volatility, Taranto et al. think that the evidence of long-term support on the cryptocurrency market is dubious. The strategy employed in their study is ARCH-GARCH probability. When calculating the findings of their research, they take into account several variables, including cryptocurrency investments, returns, and price movements. They assert that investors will benefit from their research when learning how to evaluate bitcoin price levels.[7] By examining the price time series over a 3-year period, Azari, A.[21] sought to demonstrate the utility of the conventional autoregressive integrative moving average (ARIMA) model in predicting the future value of bitcoin.

Seighaly et al. Express that the main objective of their research is to study the accuracy and precision of algorithmic trading systems like LSTM and RNN. To predict the low and high prices of cryptocurrencies, the authors of this study used algorithmic trading strategies like LSTM and RNN. The variables for the study are mainly catagorized into two main categories. The target variable is one of the variable that they were seeking to predict. The target variable in this research is the top or bottom price (1 or 0), which is measuring using the Awesome Oscillator (AO) indicator from the closing price of candles. Seighaly et al. conclude that their research in this area will be very insightful and will offer a reliable trading strategy based on the predictions of the RF model[8].

Wu et al. States their study's key argument is that comments posted on social media affect the price of cryptocurrencies. Using a combination of sentiment analysis and machine learning approaches, this study investigates if social media sentiment towards Bitcoin has an impact on the daily price returns of other cryptocurrencies. Wu et al. used machine learning models like neural networks and regression analysis along with sentiment analysis techniques like VADER to estimate cryptocurrency prices. They are trying to fill in knowledge gaps about how changes in cryptocurrency pricing are influenced by social media and public opinion. This study suggests utilizing a Naive Bayes classifier or a FinBERT transformer to improve the quality of sentiment capturing[9].

Petrovic et al.[10] Their essay's main body study makes the case that Cryptocurrencies are electronic money, or more precisely, computer code that acts as a medium of exchange. The factors of security and transparency are ensured by the use of cryptographic techniques. Block-chain, sentiment data, cryptocurrencies, and other characteristics were employed in their paper. In their research, the long short-term memory model and the arithmetic optimization algorithm are the two main techniques used. The very four basic arithmetic operations—addition, subtraction, multiplication, and division—are used. The balance between the exploration and the exploitation in this operation is the operation's primary influencing factor. The authors believe that this paper will help investors by providing with with a sophisticated solution for predicting cryptocurrency prices.

Latif et al. The body of their paper predicts the market value of multiple cryptocurrencies from their peak values and bottom values by using autoregressive integrated moving average (ARIMA) and long-short-term memory (LSTM) models and then compares the results. The metrics MAPE Forecast Accuracy

(Mean Absolute Percentage Error), MAE Forecast Accuracy (Mean Absolute Error), MAD (Mean Absolute Deviation), WAPE (Weighted Absolute Percentage Deviation), and RMSE are used to compare the results of the two models (Root Mean Squared Error). The results and conclusion of this paper shows that the LSTM model outperforms the model called ARIMA with a 99.73% accuracy rate. To find the most accurate and precise findings per Bitcoin value, this study argues that more work may be done to improve the LSTM utilizing real-time data[11].

Yamin et al. States the main goal of their article is to utilize machine learning to forecast the trend and trajectory of the bitcoin market. The researchers suggested using machine learning models built on statistical, algorithmic, and neural network models as efficient price prediction techniques. Based on a thorough literature assessment of the most recent machine learning-based prediction methodologies, the market for digital assets is expected to expand dramatically between 2014 and 2022, according to the forecast made in this paper. This study will classify, assemble, and discuss the prior research on using machine learning classifiers to forecast the bitcoin market. Future research output will grow as a result of this study. According to the study, profitability cannot be determined just by machine learning performance evaluation criteria. In order to compare performance, the ideal classifier that was previously supplied should be compared to the suggested machine learning models[12].

Kamolov et al. States that the main section of the essay explains how high-frequency trading aids in cryptocurrency price forecasting. They do this by contrasting the performance of the traditional ordinary least squares (OLS) linear regression with that of the elastic net (Enet) regression. The results are unexpected because OLS is more precise than the more complex Enet. They also determined the factors that drive each model. Stock prices, market indexes, exchange rates, and cryptocurrencies are just a few of the many parameters they take into account to arrive at their conclusion. They examined 600K observations and 108 variables from minute-based trade data. Surprising findings were found when researchers compared the performance of the Enet model against OLS. In numerous cases, they discovered that OLS performed better than Enet. They also observed that Bitcoin-related variables such as previous closing, open, and high prices are the most important factors in determining the price in the next time step[13]. Li, X.[22] et al. sought to mine the numerous patterns embedded in bitcoin transactions, and proposed a k-order transaction graph to uncover patterns of varying breadth. To automatically encode the patterns, they developed a transaction graph-based feature. However, it was unable to meet the shifting prices of 2017.

3. PROPOSED METHODOLOGY

In this academic research paper, the model called LSTM , Long Short Term Memory, has been applied, which is one of the recurrent neural networks(RNN) model, which has shown very good results in the forecasting of the time series data, and sequential data. LSTM models are particularly effective in forecasting the price of bitcoin because they can manage non-linear correlations between input variables and output objectives and capture long-term dependencies in the data. The primary function required for the analysis of the time series bitcoin data is data collection. We have collected two lakhs and sixty-four bitcoin data through web scrapping from the open-source API called *https://www.coindesk.com/.* The dataset comprises 3 months of data from 1st January to 20th March of 2023. The dataset has Datetime, Symbol, Open, High, Low, and Close

attributes and we worked on the Close column for the prediction and analysis.

Traditional approaches have utilized several statistical models for the purpose of prediction such as ARMA(autoregressive moving average), ARIMA(autoregressive integrated moving average), GRU(Gate Recurrent Unit), GARCH(Generalized Autoregressive Conditional Heteroscedasticity and so on to find the best estimates, but their accuracy varied due to the lack of either proper dataset or traditional training models. Viewing such a situation of the prediction, we have used the LSTM model for the prediction of the price. It works as:
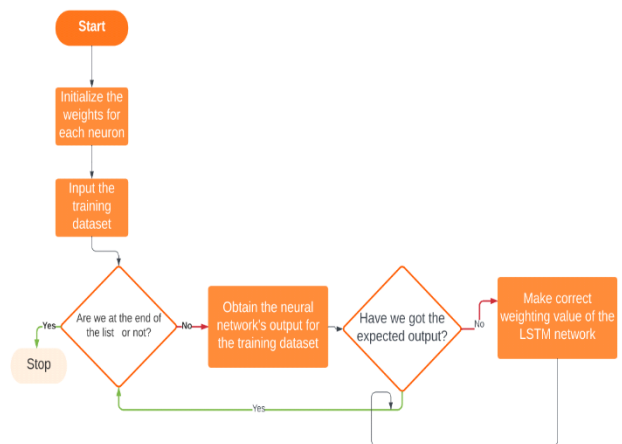

Figure 1. Working of LSTM model in our approach

Before applying the LSTM model, we preprocess the dataset collected and split the data into training and testing data where we train the model on training data using LSTM. In the LSTM model, we have first to create a model using the training data where weights are initialized on them. The weights are checked and tried to minimize the error based on the bias obtained using the iterative number of epochs(we used 10 epochs for our experiment) through the neural networks model. The output is again checked till we reach an exact or nearer to the expected output is obtained.

For situations involving sequences and time series, LSTMs offer a highly promising solution. The difficulty of training them, though, is the one drawback that can be thought of with it. Even basic models require a lot of system resources and training time. But that's only a hardware limitation! The following figure summarizes the overall functioning and pipeline of the prediction of the bitcoin prices from the beginning to the end of the process till we get the optimal solution.

The process starts from collection, pre-processing, model generation, and compilation, to evaluation and deployment after we get the expected or near-expected solution. Our main focus was on the quality of data along with quantity for better and more precise output. Figure summarizes the overall pipeline of bitcoin price prediction using LSTM.

LSTM is a sort of recurrent neural network (RNN) architecture that was developed to overcome the vanishing or exploding gradient problem that plagued standard RNNs. It captures and stores long-term dependencies in sequential data using gated units and a memory cell. The gates regulate the flow of information, allowing LSTMs to retain and use knowledge over long periods of time, making them useful for tasks such as language modeling, time series prediction, and more.
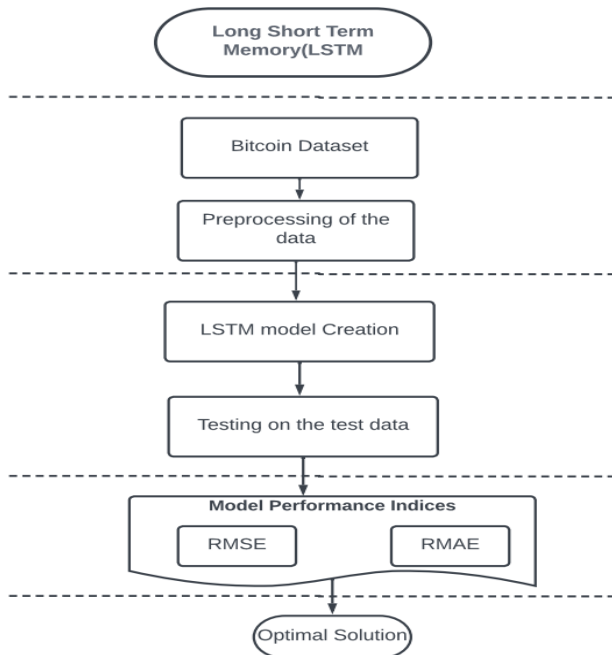
Figure 2. Our Overall Prediction Model

The "Long Short-Term Memory" of the LSTM architecture is only used by the RNNs' short-term memory, which can store, process, and analyse thousands of timestamps.A typical LSTM unit consists of a cell, an input gate, an output gate, and a forget gate. Due to the control, these three cells exert over the flow of data into and out of the cell, the cell successfully preserves values throughout a range of time periods. Remember gates choose whether information from the previous state to reject or extract by assigning a value between 0 and 1 to a former state in reference to a current input. A value of 1 denotes preservation or keeping, while numerous 0s denote discarding.
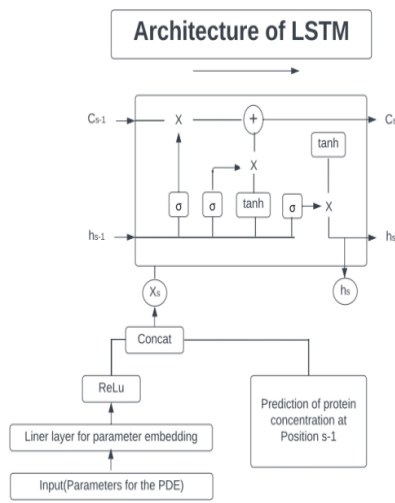

Figure 3. LSTM Architecture

Very Similar to forget gates, input gates also use a function that helps to analyse which new bits/units of information will be stored in the existing form. each and every piece of information is provided in the current state a value between 0 and 1, output gates are well capable to decide which information are output while taking into consideration of both past and current states/form. The LSTM network may collect useful, long-term dependencies to provide predictions, ,By selectively providing

relevant data in both present and future time steps from the current state.

4. EXPERIMENTAL FINDINGS AND RESULTS

Long Short-Term Memory(LSTM) models are particularly effective in forecasting the price of bitcoin because they can manage non-linear correlations between input variables and output objectives and capture long-term dependencies in the data. The primary function required for the analysis of the time series bitcoin data is data collection. We have collected two lakhs and sixty-four bitcoin data. The dataset comprises 3 months of data from 1st January to 20th March of 2023. The dataset has Datetime, Symbol, Open, High, Low, and Close attributes and we worked on the Close column for the prediction and analysis. We created the LSTM model by using the training data which comprises 175,045 rows, and the model is fitted and tested on 75019 testing datasets. After doing this, we got the loss function history of 0.0014 value after applying 10 epochs with batch size 100. It is visualized as:
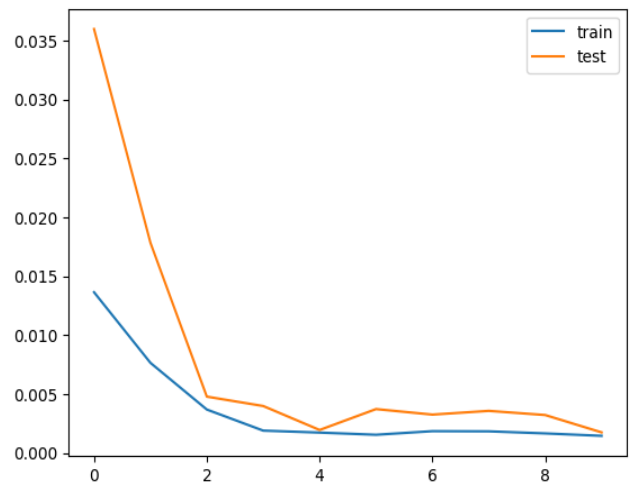

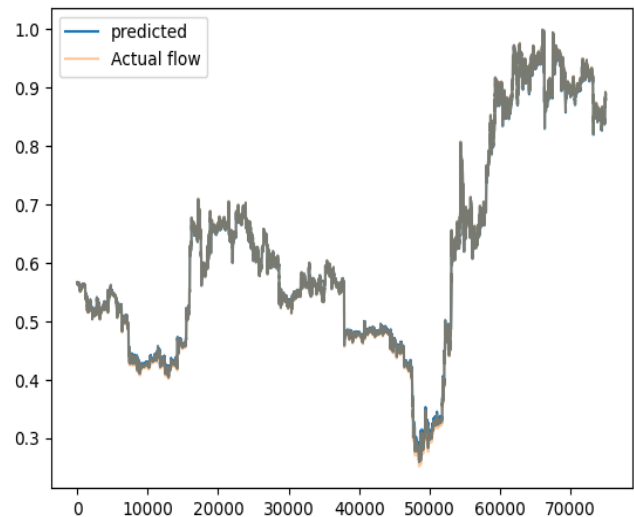Figure 4. Loss function history of the Model

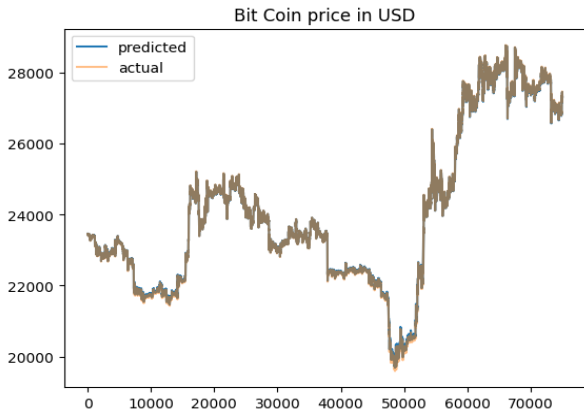Figure 5. Prediction on normalized data

Figure 6. Prediction on real non-normalized data

Table 1. RMSE and RMAE evaluation

	Root Mean Square Error	Root Mean absolute Error
0	29.592783	4.615849

Table 1 demonstrates how the RMSE and RMAE evaluation indicators are calculated based on the LSTM model prediction results.

It is found that 52.78% is the highest percentage of accuracy of the LSTM model, found till 2019 and our model has the accuracy increased to 61.34% of the accuracy for the prediction of bitcoin price. Similarly,

```
--------------------
Accuracy of the Model
--------------------
Accuracy of LSTM model is: 61.34 %
```

Figure 7. Accuracy of the model

```
Error percentage of 30th March data = 2.3057653933357067
Accuracy percentage is: 97.6942346066643
=====================================================
=====================================================
Error percentage of 31st March data = 3.83204484579233
Accuracy percentage is: 96.16795515420768
```

Figure 8. Accuracy of the new data prediction

The price of bitcoin is extremely volatile and its prediction with 100% accuracy is never the possible thing to happen. But, its precision and accuracy can be improved. In our experimental analysis, we gained an average of 96.88% accuracy. However, we gained 97.694 % and 96.027% of accuracy on prediction of 30th and 31st March data respectively, which is around 35% increase in accuracy percentage for real data, in which the data tested was never on the entire dataset. Furthermore, we have taken the dataset of 3 months for experimental analysis. It's because the price of bitcoin was extremely fluctuating due to COVID-19 and many other factors which reached $ 68770 from $ 320.19 which is around a 21377% increment in January 2021. This is the reason we took the data for the last three months from January to March 2023.

5. CONCLUSIONS
In the studies we conducted, we found the accuracy and precision of our LSTM model to predict the bitcoin price to be 61.34% and

61% respectively. The LSTM model is a very useful model for predicting time series or sequential data such as Bitcoin price prediction because the LSTM model is better at recognizing longer-term relationships. It is a fact that the integrity along with quality of training data and the size of the dataset population play a vital role in predicting the Bitcoin price. In this study, we have used the dataset population size of over 2.5 lakhs, which is a minute-wise bitcoin price of the last three months i.e from 1st January to 29th March 2023. This dataset size is not that big so for our future studies, we plan to increase our data set population which will be of minute-wise data from 2014 to till date. In future, we have also planned to include other parameters such as social media and sentiment analysis which will help to discover new findings to help in the studies in the field of Bitcoin and its trading which will help the youth to understand this field better about when to invest and when to buy/sell the bitcoin.

6. REFERENCES
[1] E. Sin and L. Wang, "Bitcoin price prediction using ensembles of neural networks, "2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 2017

[2] S. Velankar, S. Valecha, and S. Maji, "Bitcoin price prediction using machine learning," 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, Korea (South), 2018.

[3] Amin Azari, "Bitcoin Price Prediction: An ARIMA Approach", 2019 Cornell University.

[4] K. Rathan, S. V. Sai and T. S. Manikanta, "Crypto-Currency price prediction using Decision Tree and Regression techniques," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019.

[5] Ji, Suhwan, Jongmin Kim, and Hyeonseung Im. "A comparative study of bitcoin price prediction using deep learning." Mathematics 7, no. 10 (2019): 898.

[6] Kwon, Do-Hyung, Ju-Bong Kim, Ju-Sung Heo, Chan-Myung Kim, and Youn-Hee Han. "Time series classification of cryptocurrency price trend based on a recurrent LSTM neural network." Journal of Information Processing Systems 15, no. 3 (2019): 694-706.

[7] R. Albariqi and E. Winarko, "Prediction of Bitcoin Price Change using Neural Networks," 2020 International Conference on Smart Technology and Applications (ICoSTA), Surabaya, Indonesia, 2020.

[8] Hari Krishnan Andi, "An Accurate Bitcoin Price Prediction using logistic regression with LSTM Machine Learning model",Journal of Soft Computing Paradigm (JSCP) (2021).

[9] Andi, Hari Krishnan. "An accurate bitcoin price prediction using logistic regression with LSTM machine learning model." Journal of Soft Computing Paradigm 3, no. 3 (2021): 205-217.

[10] Nimkova, P. and Moreira, R., Improved Bitcoin Price Prediction based on COVID-19 data, 2023.

[11] Taryanto, R. and Afifah, N. Crypto Price Prediction as an Investment Opportunity: An Empirical Study of Three

Global Cryptocurrencies.Asian Journal of Economics, Business, and Accounting, 2023.

[12] Seighaly, M., Koosha, E. and Abbasi, E. Measuring the Accuracy and Precision of Random Forest, Long Short-Term Memory, and Recurrent Neural Network Models in Predicting the Top and Bottom of Bitcoin price. Journal of Mathematics and Modeling in Finance, 2023.

[13] Wu, E. Application of Sentiment Analysis and Machine Learning Techniques to Predict Daily Cryptocurrency Price Returns, 2023.

[14] Petrovic, A., Jovanovic, L., Zivkovic, M., Bacanin, N., Budimirovic, N. and Marjanovic, M. Forecasting Bitcoin Price by Tuned Long Short Term Memory Model. In 1st International Conference on Innovation in Information Technology and Business (ICIITB 2022).

[15] Latif, N., Selvam, J.D., Kapse, M., Sharma, V. and Mahajan, V., Comparative Performance of LSTM and ARIMA for the Short-Term Prediction of Bitcoin Prices. Australasian Accounting, Business, and Finance Journal, 2023.

[16] Yamin, M.A. and Chaudhry, M. Cryptocurrency market trend and direction prediction using Machine Learning: A Comprehensive Survey, 2023.

[17] Kamolov, S. High-frequency Bitcoin price prediction: statistical learning approach. Annals of Mathematics and Computer Science, 2023.

[18] Latif, Navmeen, Joseph Durai Selvam, Manohar Kapse, Vinod Sharma, and Vaishali Mahajan, "Comparative Performance of LSTM and ARIMA for the Short-Term Prediction of Bitcoin Prices," Australasian Accounting, Business and Finance Journal 17, no. 1 (2023): 256-276.

[19] Yamin, Muhammad Abubakar, and Maham Chaudhry. "Cryptocurrency market trend and direction prediction using Machine Learning: A Comprehensive Survey." (2023).

[20] Raju, S. M., & Tarif, A. M. (2020). Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis. ArXiv. /abs/2006.14473

[21] Azari, A. (2019). Bitcoin Price Prediction: An ARIMA Approach. ArXiv. /abs/1904.05315.

[22] Li, X., & Wu, W. (2020). A Blockchain Transaction Graph based Machine Learning Method for Bitcoin Price Prediction. ArXiv. /abs/2008.09667