

**UNIVERSITY COLLEGE OF
ENGINEERING AND TECHNOLOGY**

***VINOBA BHAVE UNIVERSITY
Hazaribag***



***BITCOIN PRICE PREDICTION USING
MACHINE LEARNING TECHNIQUE***

A PROJECT REPORT

Submitted by:-

- ANISHA KUMARI [1911102]
- RAHUL KUMAR [1911113]
- SAURABH KR. DEEPAK [1911115]
- SHIVAM ANAND [1911116]
- JAULY MISHRA [1911127]

Under the guidance of

MR. VIKASH KISHORE

(Asst. Professor, Department of Information Technology)

ACKNOWLEDGEMENTS

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend our sincere thanks to all of them.

We Would like to express my deepest gratitude to our supervisor, Mr. Vikash Kishore (Asst. Professor), for his valuable guidance, consistent encouragement, personal caring, timely help and pro-viding us with an excellent atmosphere for doing research. All through the work, in spite of his busy schedule, he has extended cheerful and cordial support to me for completing this research.

We would also like to express our gratitude towards our parents & friends for their kind co-operation and encouragement which help us in completion of this project.

We would like to express our special gratitude and thanks to college professors for giving us such attention and time.

My thanks and appreciations also go to my colleague in developing the project and people who have willingly helped me out with their abilities

DECLARATION

We certify that,

- a) The work contained in this report is original and has been done by we under the guidance of my mentor.
- b) We have followed the guidelines provided by the Institute in preparing the thesis.
- c) We have confirmed to the norms and guidelines in the Ethical Code of Conduct of the Institute.
- d) Wherever we have used materials (figures and text) from other sources, we have given due credit to them by citing them in the text of the thesis and giving their details in the references.

Date:- _____

Anisha Kumari (1911102)
Rahul Kumar (1911113)
Saurabh Kr Deepak (1911115)
Shivam Anand (1911116)
Jaully Kumari (1911127)

CERTIFICATE

This is to certify that the Project Report entitled, “**Bitcoin Price Prediction**” submitted by Anisha Kumari(1911102), Rahul Kumar(1911113), Saurabh Kr. Deepak(1911115), Shivam Anand(1911116), Jaully Kumari(1911127) to **University College of Engineering and Technology, Hazaribag**, India, is a record of bonafide Project work carried out them under our supervision and guidance and is a worth of considerance for the award of the degree of Bachelor of Technology in Information Technology of the Institute.

Supervisor

External

HOD I.T
UCET VBU Hazaribag

ABSTRACT

With the advent of technological marvels like global digitization, the prediction of the bitcoin has entered a technologically advanced era, revamping the old model of trading. With the ceaseless increase in market capitalization, bitcoin trading has become a center of investment for many financial investors.

Many analysts and researchers have developed tools and techniques that predict bitcoin price movements and help investors in proper decision-making. Advanced trading models enable researchers to predict the market using non- traditional textual data from social platforms. The application of advanced machine learning approaches such as text data analytics and ensemble methods have greatly increased the prediction accuracies. Meanwhile, the analysis and prediction of bitcoin markets continue to be one of the most challenging research areas due to dynamic, erratic, and chaotic data.

This study explains the systematics of machine learning- based approaches for bitcoin prediction based on the deployment of a generic framework. Findings from the last decade (2013–2023) were critically analyzed, having been retrieved from online digital libraries and databases like ACM digital library and Scopus.

Furthermore, an extensive comparative analysis was carried out to identify the direction of significance. The study would be helpful for emerging researchers to understand the basics and advancements of this emerging area, and thus carry-on further research in promising directions.

Keywords: Machine Learning; Auto Regressive Integrated Moving Average Model; Long Short Term Memory Model

TABLE OF CONTENTS :

1. Introduction
2. Scope of the Project
3. Library Used
4. Overview Of Methods
5. LSTM Model
6. LSTM Vs RNN
7. Overview Of LSTM Methodology
8. Implementation Of LSTM On Bitcoins Data.
 - a) Reading Data
 - b) Exploring Dataset
 - c) Data Pre-Processing
 - d) Splitting Data For Training
 - e) Implementation Of Our LSTM Model
 - f) Visualization
 - g) Inference
9. Shortcomings of LSTM
10. ARIMA Model
11. Why ARIMA is better than LSTM
12. Conclusion
13. Future Enhancement
14. References

INTRODUCTION

Bitcoin, the pioneering cryptocurrency, has garnered significant attention and popularity since its inception. With its decentralized nature and potential for high returns, Bitcoin has become a preferred investment option for individuals and institutions alike. However, the volatile nature of the cryptocurrency market poses a considerable challenge for investors in predicting the price movements of Bitcoin accurately. In response to this challenge, the application of machine learning techniques has gained prominence as a promising approach to forecast Bitcoin's price.

The objective of this project is to develop a robust machine learning model capable of predicting Bitcoin's price based on historical data and relevant features. By analyzing patterns, trends, and relationships within the data, we aim to provide valuable insights into the future price movements of Bitcoin, enabling investors to make more informed decisions.

The project will follow a comprehensive methodology involving data collection, preprocessing, model selection, training, evaluation, and prediction. Historical Bitcoin price data, along with pertinent features such as trading volume, market sentiment, and macroeconomic indicators, will be collected and processed. Various machine learning algorithms, including linear regression, decision trees, random forests, support vector machines (SVM), and deep learning models like recurrent neural networks (RNN) or long short-term memory (LSTM), will be explored and trained using the preprocessed data.

The trained models will be evaluated using appropriate metrics to assess their predictive accuracy, and predictions will be visualized through graphs and charts for better interpretation. A comparative analysis of the models will be conducted, considering factors such as accuracy, computational efficiency, and interpretability.

The anticipated outcomes of this project include a reliable machine learning model that can forecast Bitcoin's price movements and contribute to the understanding of the cryptocurrency market. However, it is crucial to acknowledge that the inherent volatility and uncertainty of the cryptocurrency market make accurate predictions challenging. Therefore, the developed model should be utilized as a tool to aid decision-making rather than as a definitive indicator of future performance.

SCOPE OF THE PROJECT

The scope of the project titled "Bitcoin Price Prediction" is to develop an advanced and optimized system for forecasting the price of Bitcoin. The project aims to utilize various techniques, such as machine learning algorithms, statistical models, and technical analysis, to improve the accuracy of predictions and assist investors in making informed decisions.

The project will involve collecting and analyzing large amounts of historical data related to Bitcoin, including price movements, trading volumes, market trends, and other relevant factors. This data will serve as the foundation for training and testing various prediction models.

Machine learning algorithms will be implemented to identify patterns and correlations within the data. Techniques like regression analysis, time series analysis, and deep learning models can be used to capture complex relationships and predict future price movements.

To optimize the prediction system, feature engineering techniques will be applied to select the most relevant and influential variables. These may include factors like trading volume, market sentiment, social media trends, economic indicators, and news sentiment analysis. By incorporating these features, the model can capture a broader range of market dynamics and improve its forecasting capabilities.

To evaluate the performance of the prediction models, backtesting and cross-validation techniques will be employed. Historical data will be divided into training and testing sets to assess the accuracy and reliability of the models. The project aims to achieve a high level of accuracy, measured by metrics like mean absolute error (MAE), mean squared error (MSE), and directional accuracy.

Finally, the project will culminate in the development of a user-friendly interface or application that allows users to access the Bitcoin price predictions. This interface can provide real-time updates, visualizations, and customizable settings to cater to individual preferences.

LIBRARIES USED

1.TensorFlow:This library was developed by Google in collaboration with the Brain Team. It is an open-source library used for high-level computations. It is also used in machine learning and deep learning algorithms. It contains a large number of tensor operations. Researchers also use this Python library to solve complex computations in Mathematics and Physics.

2. Matplotlib: This library is responsible for plotting numerical data. And that's why it is used in data analysis. It is also an open-source library and plots high-defined figures like pie charts, histograms, scatterplots, graphs, etc.

3. Pandas: Pandas are an important library for data scientists. It is an open-source machine learning library that provides flexible high-level data structures and a variety of analysis tools. It eases data analysis, data manipulation, and cleaning of data. Pandas support operations like Sorting, Re-indexing, Iteration, Concatenation, Conversion of data, Visualizations, Aggregations, etc.

4. Numpy:The name "Numpy" stands for "Numerical Python". It is the commonly used library. It is a popular machine learning library that supports large matrices and multi-dimensional data. It consists of in-built mathematical functions for easy computations. Even libraries like TensorFlow use Numpy internally to perform several operations on tensors. Array Interface is one of the key features of this library.

Basic Terminologies Used:

1. Opening Price:

- i. The opening price is **the price at which a security first trades when an exchange opens for the day**. An opening price is not identical to the previous day's closing price. There are several day-trading strategies based on the opening price of a market or security.

2. Closing Price:

- i. The closing price is **the last price at which a security traded during the regular trading day**. A security's closing price is the

standard benchmark used by investors to track its performance over time.

3. Volume:

Volume **measures the number of shares traded in a bitcoin or contracts traded in futures or options.** Volume can indicate market strength, as rising markets on increasing volume are typically viewed as strong and healthy. When prices fall on increasing volume, the trend is gathering strength to the downside.

Overview of the methods used:

1) Data Collection:

Data collection is a very basic module and the initial step towards the project. It generally deals with the collection of the right dataset. The dataset that is to be used in the market prediction has to be used to be filtered based on various aspects. Data collection also complements to enhance the dataset by adding more data that are external. Our data mainly consists of the previous year bitcoin prices. Initially, we will be analyzing the Kaggle dataset and according to the accuracy, we will be using the model with the data to analyze the predictions accurately.

A. Preprocessing :

Data pre-processing is a part of data mining, which involves transforming raw data into a more coherent format. Raw data is usually, inconsistent or incomplete and usually contains many errors. The data pre-processing involves checking out for missing values, looking for categorical values, splitting the data-set into training and test set and finally do a feature scaling to limit the range of variables so that they can be compared on common environments.

B. Training the Machine

Training the machine is similar to feeding the data to the algorithm to touch up the test data. The training sets are used to tune and fit the models. The test sets are untouched, as a model should not be judged based on unseen data. The training of the model includes cross-validation where we get a well-grounded approximate performance of the model using the training data. Tuning models are meant to specifically tune the hyper parameters like the number of trees in a random forest. We perform the entire cross-validation loop on each set of hyper parameter values

Finally, we will calculate a cross-validated score, for individual sets of hyper parameters. Then, we select the best hyper parameters. The idea behind the training of the model is that we some initial values with the dataset and then optimize the parameters which we want to in the model. This is kept on repetition until we get the optimal values. Thus, we take the predictions from the trained model on the inputs from the test dataset. Hence, it is divided in the ratio of 74:26 where 74% is for the training set and the rest 26% for a testing set of the data.

C. Data Scoring :

Data scoring refers to the process of assigning scores or weights to different data points or features based on their relevance and significance in predicting Bitcoin's price. This scoring mechanism helps prioritize and select the most influential variables for the prediction models. By assigning higher scores to the most informative features, the models can focus on capturing the key factors driving Bitcoin's price movements. Data scoring improves the accuracy and efficiency of the prediction system by reducing noise and emphasizing the most impactful data points during the analysis and modeling-stages.

D. Proposed system :

In this proposed system, we were able to train the machine from the various data points from the past to make a future prediction. We took data from the previous year Bitcoin to train the model. We majorly used two machine-learning libraries to solve the problem. The first one was numpy, which was used to clean and manipulate the data, and getting it into a form ready for analysis. The other was scikit-learn, which was used for real analysis and prediction. The data set we used was from the previous years bitcoin markets collected from the public database available online, 74 % of data was used to train the machine and the rest 26 % to test the data.

The basic approach of the supervised learning model is to learn the patterns and relationships in the data from the training set and then reproduce them for the test data. We used the python pandas library for data processing which combined different datasets into a data frame. The tuned up data frame allowed us to prepare the data for feature extraction.

The data frame features were date and the closing price for a particular day. We used all these features to train the machine on random forest model and predicted the object variable, which is the price for a given day. We also quantified the accuracy by using the predictions for the test set and the actual values. The proposed system touches different areas of research including data pre-processing, random forest, and soon..

Implementation & Code:

Bitcoin market forecasting is an attractive application of linear regression. Modern machine learning packages like scikit-learn make implementing these analyses possible in a few lines of code. Sounds like an easy way to make money, right? Well, *don't cash in your 401k just yet*. As easy as these analyses are to implement, selecting features with ample enough predictive power to turn a profit is more of an art than science. In training our model, we'll take a look at how to easily add common technical indicators to our data to use as features in training our model.

LSTM Explained :

Now, let's understand 'What is LSTM?' First, you must be wondering 'What does LSTM stand for?' LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a variety of *recurrent neural networks (RNNs)* that are capable of learning long-term dependencies, especially in sequence prediction problems. LSTM has feedback connections, i.e., it is capable of processing the entire sequence of data, apart from single data points such as images. This finds application in speech recognition, machine translation, etc. LSTM is a special kind of RNN, which shows outstanding performance on a large variety of problems.

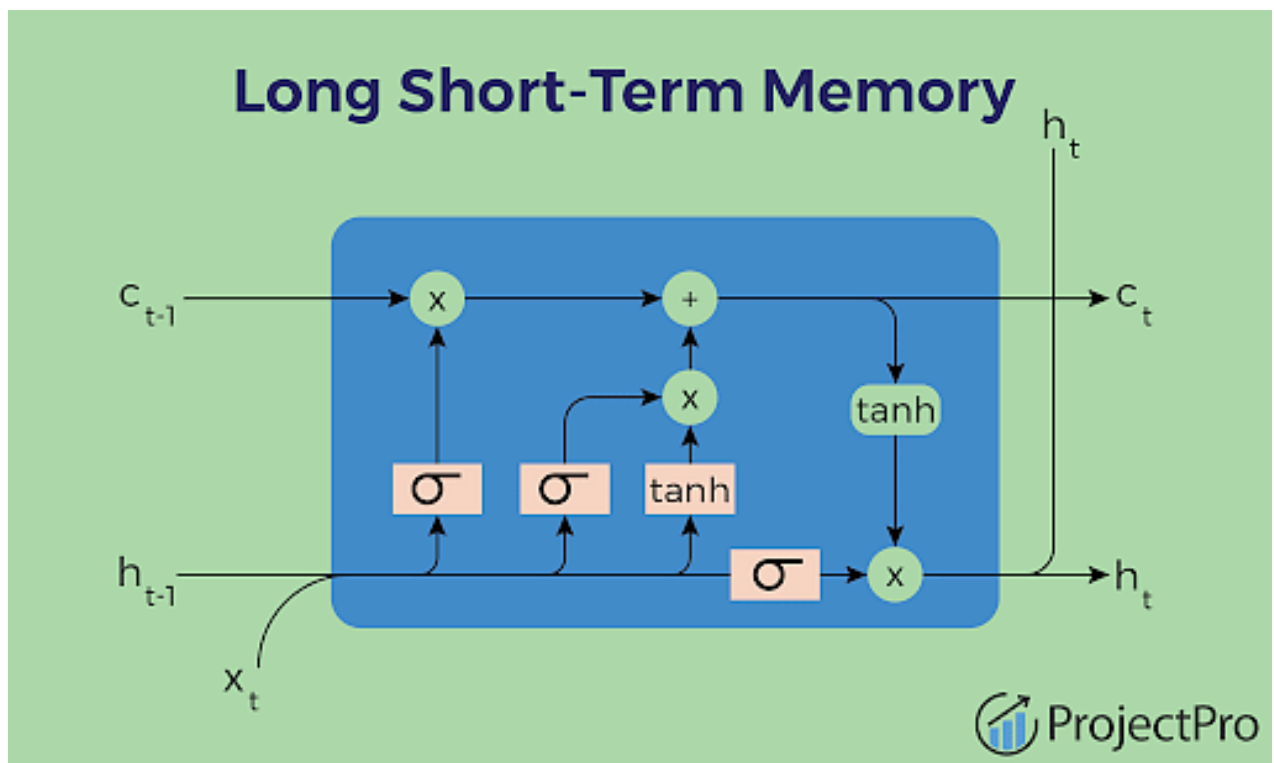
Next in this extensive blog on 'What is LSTM?', let us discuss the logic behind it.

The Logic Behind LSTM :

The central role of an LSTM model is held by a memory cell known as a 'cell state' that maintains its state over time. The cell state is the horizontal line that runs through the top of the below diagram. It can be visualized as a conveyor belt through which information just flows, unchanged. Information can be added to or removed from the cell state in LSTM and is regulated by gates. These gates optionally let the information flow in and out of the cell. It contains a pointwise multiplication operation and a sigmoid neural net layer that assist the mechanism.

The sigmoid layer gives out numbers between zero and one, where zero means ‘nothing should be let through,’ and one means ‘everything should be let through.’

Further in this ‘What is LSTM?’ blog, you will learn about the various differences between LSTM and RNN.



LSTM vs RNN

Consider, you have the task of modifying certain information in a calendar. To do this, an RNN completely changes the existing data by applying a function. Whereas, LSTM makes small modifications on the data by simple addition or multiplication that flow through cell states. This is how LSTM forgets and remembers things selectively, which makes it an improvement over RNNs.

Now consider, you want to process data with periodic patterns in it, such as predicting the sales of colored powder that peaks at the time of Holi in India. A good strategy is to look back at the sales records of the previous year. So, you need to know what data needs to be forgotten and what needs to be stored for later reference. Else, you need to have a really good memory. Recurrent neural networks seem to be doing a good job at this, theoretically. However, they have two downsides, exploding gradient and vanishing gradient, that make them redundant.

Here, LSTM introduces memory units, called cell states, to solve this problem. The designed cells may be seen as differentiable memory.

To Sum up!

LSTM networks are indeed an improvement over RNNs as they can achieve whatever RNNs might achieve with much better finesse. As intimidating as it can be, LSTMs do provide better results and are truly a big step in Deep Learning. With more such technologies coming up, you can expect to get more accurate predictions and have a better understanding of what choices to make.

Overview of LSTM methodology :

Long short term memory (LSTM) is a model that increases the memory of recurrent neural networks. Recurrent neural networks hold short term memory in that they allow earlier determining information to be employed in the current neural networks. For immediate tasks, the earlier data is used. We may not possess a list of all of the earlier information for the neural node. In RNNs, LSTMs are very widely used in Neural networks. Their effectiveness should be implemented to multiple sequence modelling problems in many application domains like video, NLP, geospatial, and time-series. One of the main issues with RNN is the vanishing gradient problem, and it emerges due to the repeated use of the same parameters, in RNN blocks, at each step. We must try to use different parameters to overcome this problem at each timestep. We try to find a balance in such a situation. We bring novel parameters at each step while generalizing variable-length sequences and keeping the overall amount of learnable parameters constant. We introduce gated RNN cells like LSTM and GRU.

Gated cells hold internal variables, which are Gates. This value of each gate at each time step depends on the information at that time step, including early states. The value of the gate then becomes multiplied by the different variables of interest to influence them. Time-series data can trace progress over milliseconds, days, and years.

Early, our perspective of time-series data meant more static; the everyday highs and lows under temperature, the opening and closing amount of the bitcoin market. Now we will go to the coding part. We will implement LSTM on the bitcoin price dataset.

Implementation of LSTM on BITCOIN PRICE PREDICTION:

Step 1 : Reading Data:

series data is a series of data values gathered over time interims, allowing us to trace differences over time.

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
sns.set_style('whitegrid')
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')

from keras.models import Sequential
from keras.callbacks import EarlyStopping
from keras.layers import Dense,LSTM,Dropout

from sklearn.preprocessing import MinMaxScaler
```

```
In [2]: data_dir='BTC_USD.csv'
df= pd.read_csv(r"C:\Users\DELL\Dropbox\My PC (DESKTOP-S92PGJG)\Desktop\project\BTC_USD.csv")
```

Step 2 -Exploring Dataset:

The dataset contains 14 columns associated with time series like the date and the different variables like close, high, low and volume. We will use opening and closing values for our experimentation of time series with LSTM.

We have performed a few feature extraction here like dropped extra columns like date , adj. Now we can be using matplotlib to visualize the available data and see how our price values in data are being displayed. The green colour was used to visualize the 17 open variable for the price-date graph, and for the closing variable, we used red colour.

```
In [3]: df.head()
```

```
Out[3]:
```

	Timestamp	Open	High	Low	Close	Volume_(BTC)	Volume_(Currency)	Weighted_Price
0	1325317920	4.39	4.39	4.39	4.39	0.455581	2.0	4.39
1	1325317980	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	1325318040	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	1325318100	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	1325318160	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
In [4]: df.tail()
```

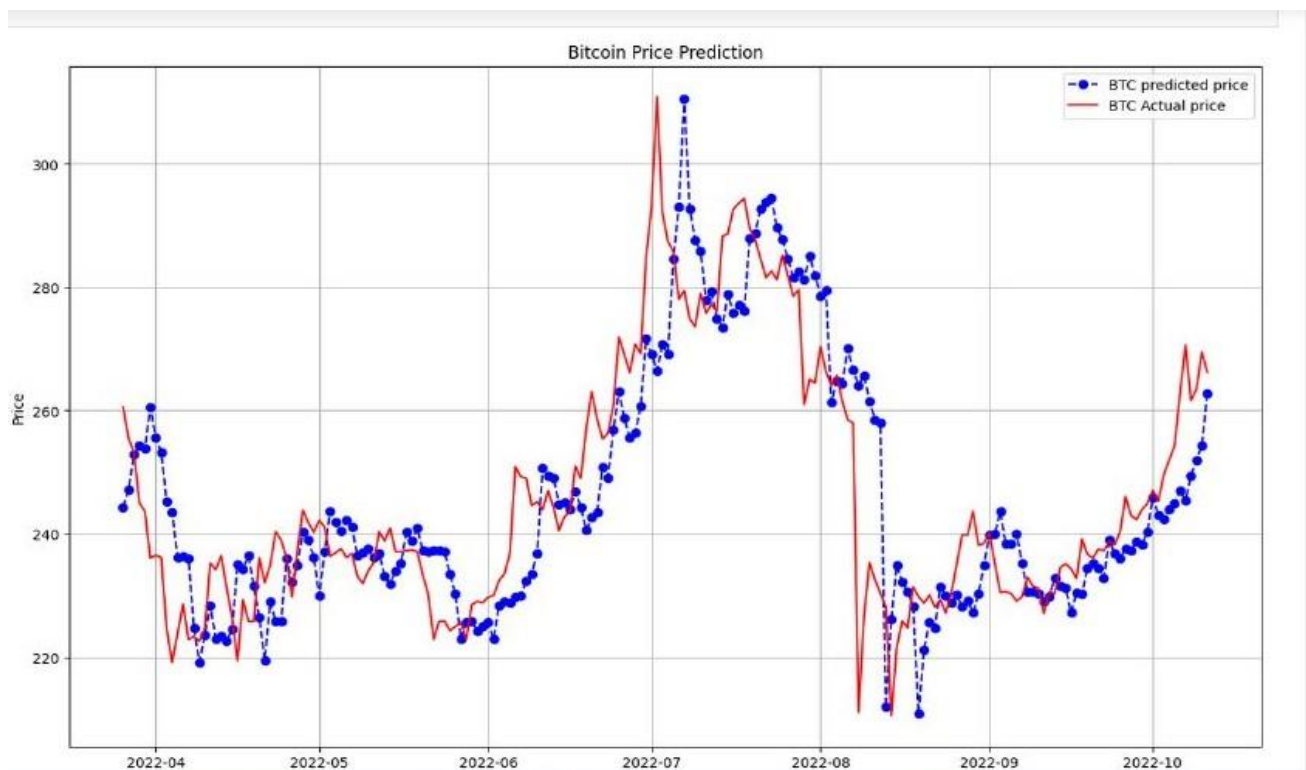
```
Out[4]:
```

	Timestamp	Open	High	Low	Close	Volume_(BTC)	Volume_(Currency)	Weighted_Price
1048570	1398232120	734.60	734.60	730.00	734.55	1.789687	1313.640757	734.005698
1048571	1398232180	734.55	734.55	730.71	730.71	0.110236	80.802051	732.991499
1048572	1398232240	734.40	734.40	730.51	730.51	0.554788	407.247985	734.063488
1048573	1398232300	730.51	733.63	730.51	731.10	0.620446	453.777190	731.372707
1048574	1398232360	733.00	734.00	733.00	734.00	9.214205	6757.415227	733.369303

Graphical representation of data.

Step 3-Data Pre-processing:

We must pre-process this data before applying bitcoinprice using LSTM. Transform the values in our data with help of the fit_transform function. Min-max scaler is used for scaling the data so that we can bring all the price values to a common scale. We then use 74 % data for training and the rest 26% for testing and assign them to separate variables.



```
[ ] ma200 = df.Close.rolling(100).mean()
ma200
```

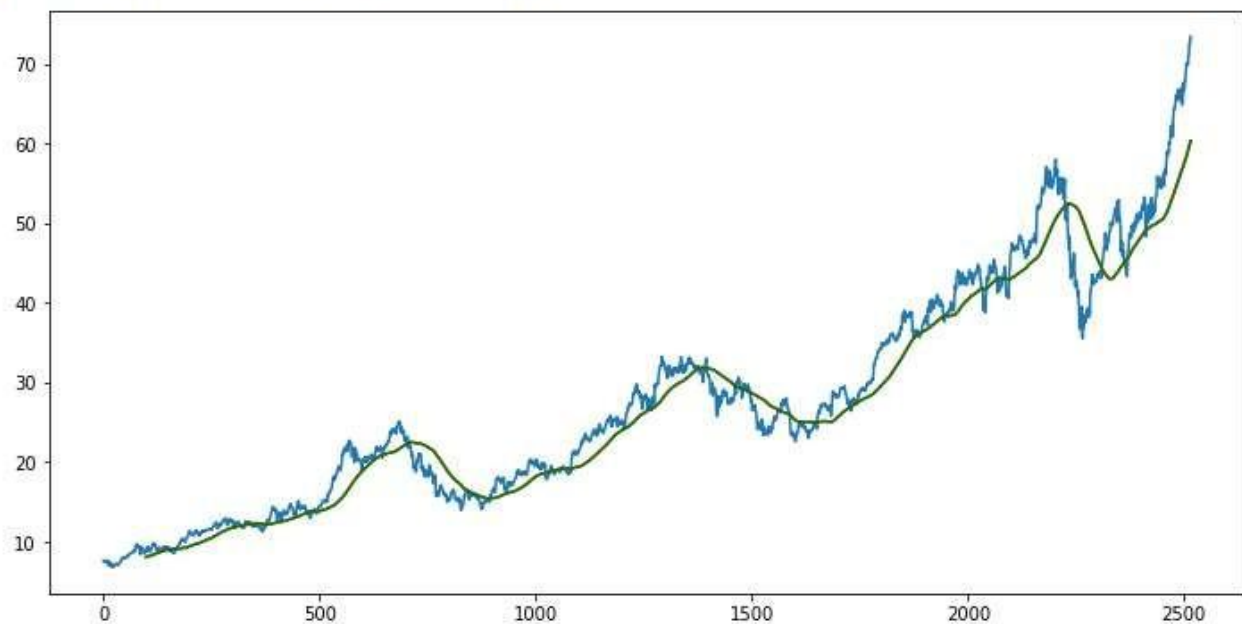
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN
...	
2511	59.401700
2512	59.643125
2513	59.875125
2514	60.106325
2515	60.331875

Name: Close, Length: 2516, dtype: float64

```
[ ] plt.figure(figsize = (12,6))
plt.plot(df.Close)
plt.plot(ma100, 'r')
plt.plot(ma200, 'g')
```

Figure 1: A line plot showing the closing price of a stock over time, with two moving averages (100-day and 200-day) overlaid. The plot is titled 'Figure 1' and has a y-axis labeled 'Close' ranging from 0 to 70. The x-axis represents time, with major ticks at 0, 500, 1000, 1500, 2000, and 2500. The blue line represents the closing price, which shows significant volatility. The red line represents the 100-day moving average, and the green line represents the 200-day moving average. Both moving averages show a general upward trend, with the 200-day moving average being smoother than the 100-day moving average.

[<matplotlib.lines.Line2D at 0x7fd7bb81bf10>]



```
[ ] df.shape
```

(2516, 6)

Step 4-Splitting data for training:

A function is created so that we can create the sequence for training and testing.

```
#Splitting Data into Training and Testing

data_training = pd.DataFrame(df['Close'][0:int(len(df)*0.70)])
data_testing = pd.DataFrame(df['Close'][int (len(df)*0.70):int(len(df))])

print(data_training.shape)
print(data_testing.shape)

(1761, 1)
(755, 1)
```

Step 5- Implementation of our LSTM model:

we will use the Sequential model imported from Keras and required libraries are imported.

```

In [15]: x_train,y_train=np.array(x_train),np.array(y_train)

In [16]: x_train=np.reshape(x_train,(x_train.shape[0],x_train.shape[1],n_cols))

In [17]: x_train.shape,y_train.shape
Out[17]: ((786371, 60, 1), (786371, 1))

In [18]: model=Sequential([
    LSTM(50,return_sequences=True,input_shape=(x_train.shape[1],n_cols)),
    LSTM(64,return_sequences=False),
    Dense(32),
    Dense(16),
    Dense(n_cols)
])
model.compile(optimizer='adam',loss='mse',metrics='mean_absolute_error')

In [19]: model.summary()

Model: "sequential"

```

Step 6 : Visualization:

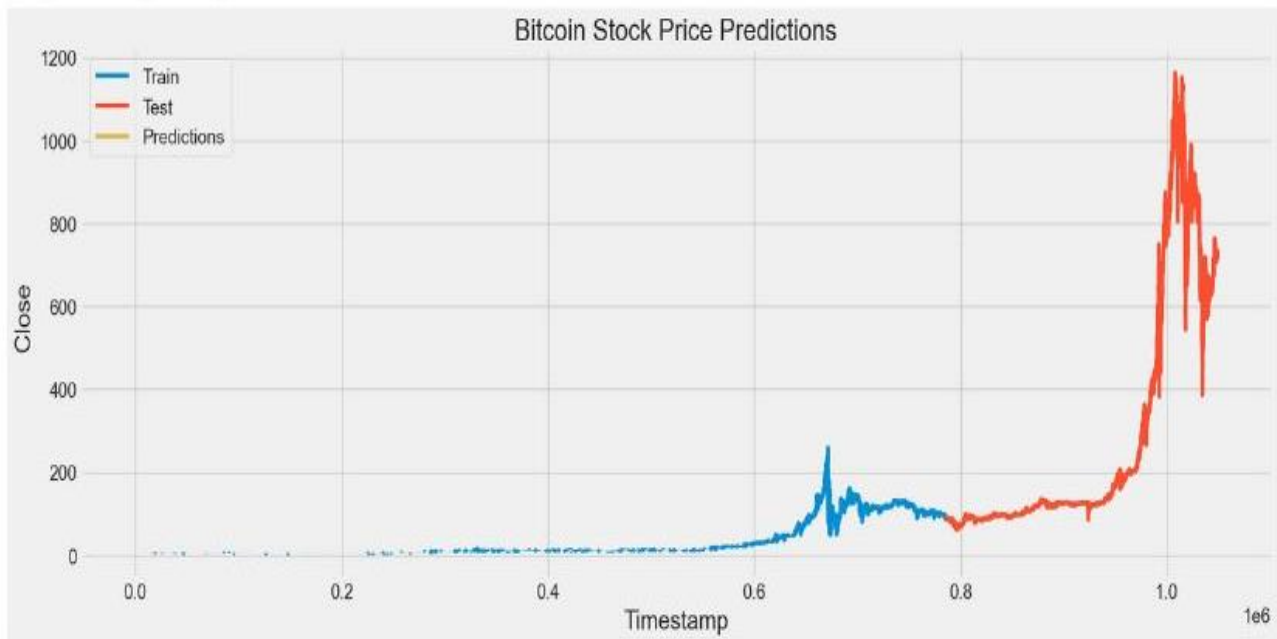
17

After fitting the data with our model we use it for prediction. We must use inverse transformation to get back the original value with the transformed function. Now we can use this data to visualize the prediction

Step 7: Inference :

Here we have used LSTM(Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that can learn and remember long-term dependencies in sequential data. It is particularly effective in capturing and inferring patterns, trends, and dependencies within time series data, making it suitable for tasks such as language translation, and speech recognition.

17



Shortcomings of LSTM:

1. Long training time: LSTM models can have longer training times compared to simpler models due to their complex architecture and computational requirements.
2. Overfitting: LSTM models are prone to overfitting, especially when the dataset is small or the model is overly complex. Regularization techniques and careful model selection are necessary to mitigate this issue.
3. Difficulty in interpretation: LSTM models are often considered as black boxes, making it challenging to interpret the internal workings and understand the reasoning behind their predictions.
4. Limited handling of long-term dependencies: Despite being designed to capture long-term dependencies, LSTMs can still struggle to capture extremely long-term dependencies in time series data.
5. Sensitivity to hyperparameters: LSTMs require careful tuning of hyperparameters, such as the number of LSTM layers, the number of hidden units, and the learning rate, to achieve optimal performance. Improper selection of hyperparameters can lead to suboptimal results.

ARIMA MODEL :

Introduction to ARIMA Model:

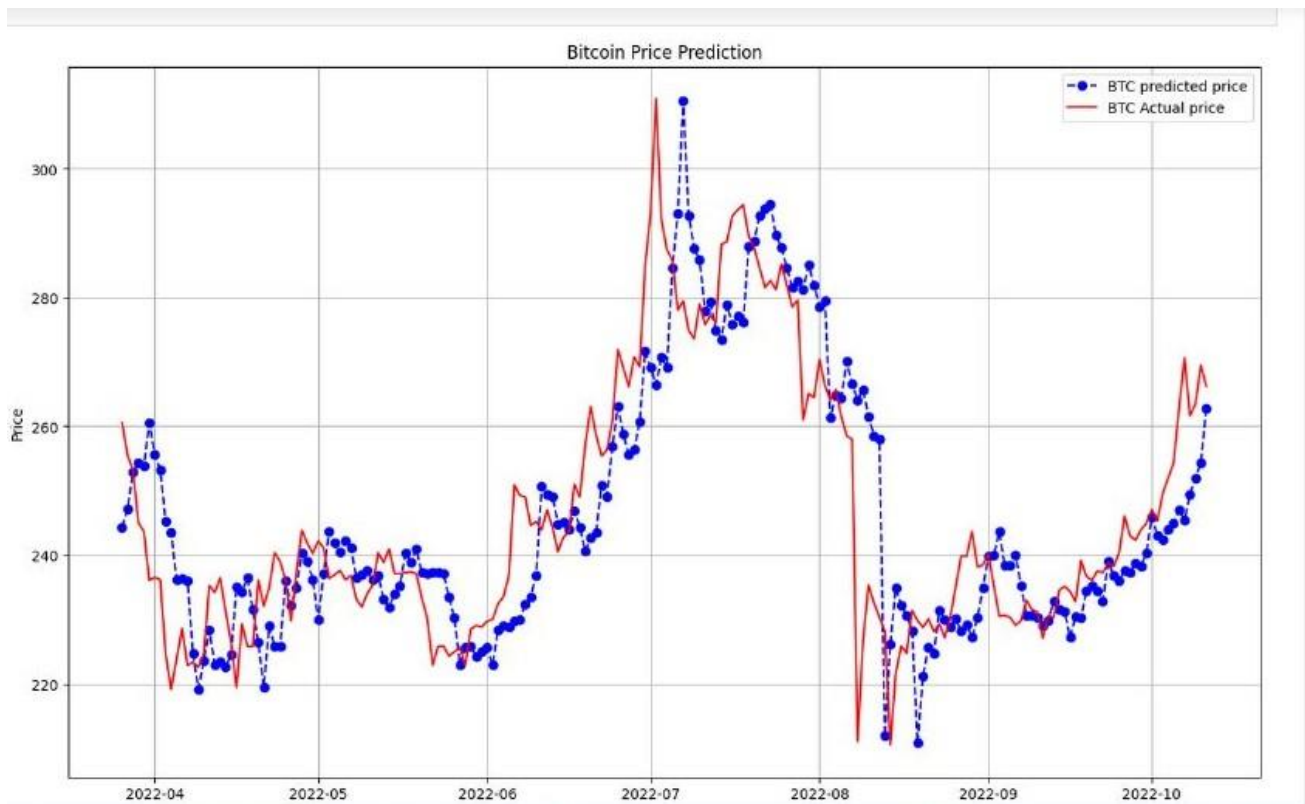
The ARIMA (AutoRegressive Integrated Moving Average) model is a popular and widely used time series forecasting technique. It is a versatile model that combines the concepts of autoregressive (AR), moving average (MA), and differencing to capture and predict patterns in time series data.

The AR component of ARIMA considers the relationship between an observation and a fixed number of lagged observations (autoregressive terms). It captures the dependence of the current value on its previous values, allowing the model to account for temporal patterns and trends.

The MA component models the dependency between an observation and a residual error from a moving average model. It helps capture the influence of past errors on the current observation, enabling the model to account for short-term fluctuations and noise in the data.

The differencing component of ARIMA is used to transform the time series data into a stationary series. Stationarity is a desirable property in time series analysis as it ensures consistent statistical properties over time. Differencing involves taking the difference between consecutive observations, which helps remove trends and seasonality from the data.

By combining these components, the ARIMA model can capture both short-term and long-term patterns in time series data, making it suitable for forecasting future values. The model parameters, including the autoregressive order (p), differencing order (d), and moving average order (q), are determined based on the characteristics of the data and are estimated through techniques like the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).



Why ARIMA is better than LSTM ?

ARIMA (Autoregressive Integrated Moving Average) and LSTM (Long Short-Term Memory) are both popular time series forecasting models, but they have different strengths. ARIMA is often preferred when the time series data exhibits a clear trend or seasonality that can be captured using autoregressive and moving average components. It is computationally efficient and interpretable, making it suitable for simpler time series problems. On the other hand, LSTM is a type of recurrent neural network that excels in capturing complex temporal dependencies and nonlinear patterns in data. It can handle long-term dependencies and is more flexible for modeling intricate relationships. The choice between ARIMA and LSTM depends on the specific characteristics and requirements of the time series being analyzed.

CONCLUSION :

Financial markets provide an excellent platform for investors and traders, who can trade from any gadget that connects to the internet. Over the last few years, people have become more attracted to bitcoin trading. Like any other walk of life, the bitcoin market has also changed due to the advent of technology. Now, people can make their investments grow. Online trading has only changed the way individuals purchase and sell bitcoins. The budgetary markets have advanced rapidly, and have formed an interconnected global marketplace.

We found that the most suitable algorithm for predicting the market price of a bitcoin based on various data points from the historical data is the random forest algorithm. The algorithm will be a great asset for brokers and investors for investing money in the bitcoin market since it is trained on a huge collection of historical data and has been chosen after being tested on a sample data. The project demonstrates the machine learning model to predict the bitcoin value with more accuracy as compared to previously implemented machine learning models.

FUTURE ENHANCEMENT:

Future scope of this project will involve adding more parameters and factors like the financial ratios, multiple instances, etc. The more the parameters are taken into account more will be the accuracy.

The algorithms can also be applied for analyzing the contents of public comments and thus determine patterns/relationships between the customer and the corporate employee. The use of traditional algorithms and data mining techniques can also help predict the corporations' performance .

Anticipating the bitcoin trade worth is extraordinarily in vogue among financial. Generally the specialized examiners and representatives acclimated foresee the bitcoin costs upheld verifiable costs, volumes, esteem designs and in this manner the fundamental patterns. Nowadays the bitcoin worth expectation has turned out to be appallingly muddled than before as bitcoin expenses don't appear to be exclusively influenced gratitude to organization's cash standing anyway furthermore because of socio-prudent state of the nation, political environment and cataclysmic events and so forth.

REFERENCES:

<https://www.kaggle.com/code/meetnagadia/bitcoin-price-prediction-using-lstm#2.-Importing-Library>

<https://www.kaggle.com/code/praneethji/bitcoin-price-prediction-with-arima>

<https://finance.yahoo.com/quote/BTC-USD?p=BTC-USD>