Cryptocurrency Analysis and Price Prediction using LSTM Neural Network and its influence on economy

Submitted in partial fulfillment of the requirements of the course of
DS203 Programming for Data Science
Centre for Machine Intelligence and Data Sciences
Indian Institute of Technology Bombay
2021

Neeraj Nixon
20D070056
Email: 20d070056@iitb.ac.in
Electrical Engineering
IIT Bombay

Shrey Ganatra
20D070074
Email: 20d070074@iitb.ac.in
Electrical Engineering
IIT Bombay

Utkarsh Jindal
200070086
Email: 200070086@iitb.ac.in
Electrical Engineering
IIT Bombay

***Abstract*—The goal of this study is to create a Prediction Model for 11 cryptocurrencies i.e. Binance Coin, Bitcoin, Cardano, Dogecoin, Ethereum, HEX, Polkadot, Solana, Tether, USD Coin, and XRP using LSTM Neural Network and its analysis and also discuss the impact of cryptocurrecies on economy.The final plot of almost all major cryptocurrencies show rise in value during the time period from Dec 2020 to April 2021 except for Tether, USD Coin, and XRP which showed a large variation from this trend.Similarly volume traded was also high during the time period from Dec 2020 to April 2021 for all the cryptocurrencies except Solana.**

***Index Terms*—- Cryptocurrency, Price Prediction**

I. Introduction

Since the invention of Bitcoin in 2008, many other cryptocurrencies have been invented. There exists more than 7500 cryptocurrencies as of November 2021. With the growth in number of cryptocurrencies, their popularity has also increased. But the fear of investment and purchase still exists due to its unstable nature. Recently the news of Indian government banning the private cyrptocurrecies has affected the market. We have also observed the changes in prices of cryptocurrencies by tweets of few individuals. So, it is very difficult for a common man to trust the market and buy cryptocurrency. Many countries have started imposing regulations and legalities for the purchase and trade of cryptocurrencies to reduce illegal activities.

Despite such difficulties, Cryptocurrencies continue to gain popularity. One must wonder why? Cryptocurrencies are virtual/digital currency which use decentralized control as opposed to a central bank digital currency. This feature attracts a lot of people towards it as no government can control the price of cryptocurrency.

We attempt to analzye some popular cryptocurrencies like Bitcoin, Binance Coin, Ethreum, Dogecoin, Polkadot, Cardano, HEX, Solana, Tether, USD Coin and XRP from the time-period of 2015-2021. We have split the dataset into train and validation to predict the future prices of these cryptocurrencies using LSTM (Long Short-Term Memory). Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. This projects predicts the prices of some popular cryptocurrencies to help interested people understand the future of market and invest wisely.

II. Background and Prior Work

This report is written for the course DS203-Programming for Data Science, at IIT Bombay. One must have good knowledge of some programming language. We have used Python for this project. Apart from this, one must have understanding of LSTM which is artificial RNN. One can also predict prices using other Machine Learning algorithms.

Some of the works related to cryptocurrency price prediction are listed below:

1) **Forecasting and trading cryptocurrencies with machine learning under changing market conditions**[Online]
https://jfin-swufe.springeropen.com/articles/10.1186/s40854-020-00217-x

2) **Bitcoin Price Prediction and Analysis Using Deep Learning Models**[Online]
https://www.researchgate.net/publication/345262720_Bitcoin_Price_Prediction_and_Analysis_Using_Deep_Learning_Models

III. Data and Methodology

The project aims to analyse the price of cryptocurrencies and predict the prices using the deep learning model LSTM (Long Short-Term Memory). We also attempt to understand the impact of major cryptocurrencies on others. We have applied the same model on all the cryptocurrencies under consideration: Bitcoin, Binance Coin, Ethereum, Dogecoin, Cardano, HEX, Polkadot, Solana, Tether USD Coin, and XRP. We took the data of the associated prices from over a substantially long interval (5 years), which also includes the COVID-19 period. A broad range of data is necessary for the prediction of prices.

Binance Coin: The data was taken over the period from 13 Sept, 2017 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" which represent the date, opening value, highest value, lowest value, closing value, adjusted closing price of the cryptocurrency and the volume of cryptocurrency. The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Bitcoin: The data was taken over the period from 13 Sept, 2015 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Cardano: The data was taken over the period from 1 Oct, 2017 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume". The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Dogecoin: The data was taken over the period from 13 Sept, 2015 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume". The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Ethereum: The data was taken over the period from 13 Sept, 2015 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

HEX: The data was taken over the period from 22 Sept, 2020 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Polkadot: The data was taken over the period from 22 Sept, 2020 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume". The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Solana: The data was taken over the period from 22 Sept, 2020 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Tether: The data was taken over the period from 13 Sept, 2015 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

USD Coin: The data was taken over the period from 13 Sept, 2019 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

XRP: The data was taken over the period from 13 Sept, 2015 to 13 Sept, 2021. The columns are "Date", "Open", "High", "Low", "Close", "Adj Close"and "Volume" . The date is of string type, volume of 64 bit integers and the rest are 64 bit floating points rounded off to 6 decimal places.

Fig. 1. **TABLE 1**. Bitcoin Dataset(First 5 values).

	Date	Open	High	Low	Close	Adj Close	Volume
0	2015-09-13	235.242004	235.934998	229.332001	230.511993	230.511993	1.847880e+07
1	2015-09-14	230.608994	232.440002	227.960999	230.643997	230.643997	2.099780e+07
2	2015-09-15	230.492004	259.182007	229.822006	230.304001	230.304001	1.917780e+07
3	2015-09-16	230.250000	231.214996	227.401993	229.091003	229.091003	2.014420e+07
4	2015-09-17	229.076004	230.285004	228.925995	229.809998	229.809998	1.893540e+07

The initial part of the report deals with the analysis of the prices of cryptocurrencies, for which the data was taken from Kaggle.

IV. Experiments and Results

Pre-Processing

The dataset used here has many NaN values. Since we are creating a predicting model which require only the 'Date' and 'Close' values we remove all rows that contain a NaN value in either columns.

Analysis

For the first part of analysis the closing value v/s date is plotted to find how the price vary with time and compare the plots of all the cryptocurrencies.

For Binance coin it is seen that the price increases steeply during the time period between Dec 2020 to April 2021 and falls around May 2021 only to rise again in the July-August 2021 time period.

A similar trend is observed in the case of Bitcoin, Cardano, Dogecoin, Ethereum, Polkadot, and Solana with only slite variation in the time period.It is also to be noted that these 6 cryptocurrencies are also the most popular ones.

XRP has reached its highest value at the end of 2017 and its price fell steeply at the beginning of 2018 only to follow the above mentioned trend.

Meanwhile Tether and USD coin do not seem to follow this trend at all, rather remain almost constant throughout the time period.

Fig. 2. Binance Coin (Close Price v/s Date).

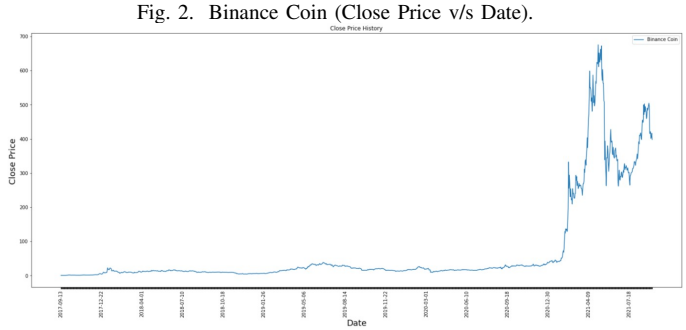


Fig. 3. Bitcoin (Close Price v/s Date).

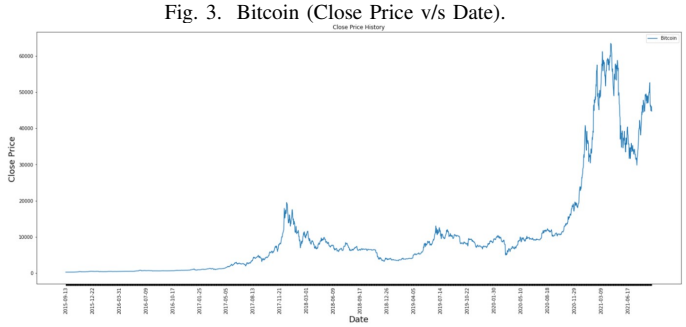


Fig. 4. Tether (Close Price v/s Date).

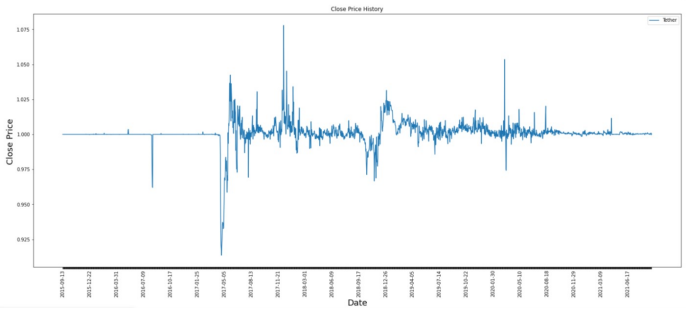


Fig. 5. Bitcoin, Ethereum and Tether (Volume v/s Date).

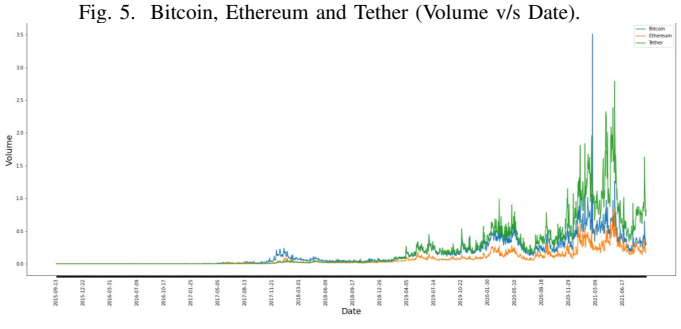
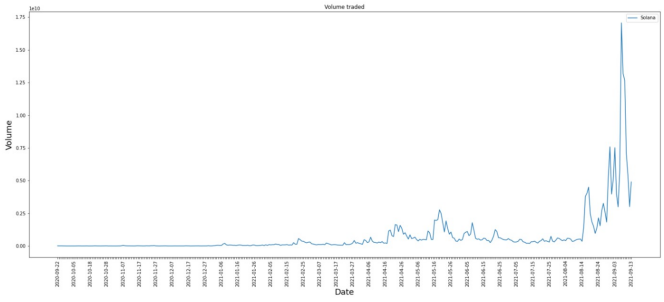


Fig. 6. Solana (Volume v/s Date).



For the second part of the analysis the Volume v/s Date is plotted to find the variations in volume traded over time and the plots are compared for all cryptocurrencies.

For Binance coin, Bitcoin, Cardano, Dogecoin, Ethereum, HEX, Polkadot, Tether, USD Coin,and XRP it is seen that the volume traded increases steeply during the time period between Dec 2020 to April 2021 and falls around May 2021 only to rise again in the July-August 2021 time period.

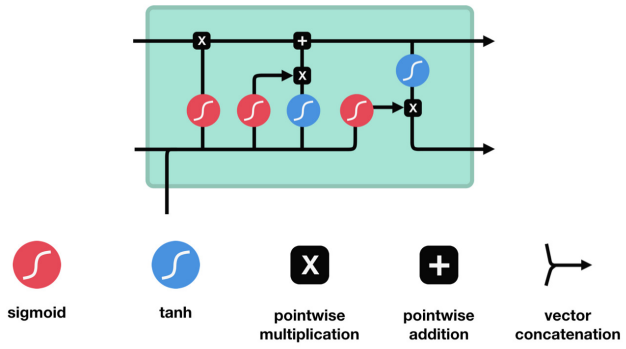
Meanwhile Solana tends to have high volume trading during August and September of 2021.

Prediction Model

Predicting future values from past values are really important as in this case it helps people to foresee what the future price of a particular cryptocurrency is and make rational choices when it comes to investment.Here we are predicting the prices of Binance coin, Bitcoin, Cardano, Ethereum, Dogecoin, Polkadot, solana, Tether, HEX, USD Coin, and XRP using the LSTM Neuarl Network model.

LSTM Neural Network

Fig. 7. LSTM



LSTM 's was created as the solution to shortterm memory. They have internal mechanisms called gates that can regulate the flow of information.

The core concept of LSTM's are the cell state, and it's various gates. The cell state act as a transport highway that transfers relative information all the way down the sequence chain. You can think of it as the "memory" of the network. The cell state, in theory, can carry relevant information throughout the processing of the sequence. So even information from the earlier time steps can make it's way to later time steps,

reducing the effects of short-term memory. As the cell state goes on its journey, information get's added or removed to the cell state via gates. The gates are different neural networks that decide which information is allowed on the cell state. The gates can learn what information is relevant to keep or forget during training.

The Forget gate decides what is relevant to keep from prior steps. The input gate decides what information is relevant to add from the current step. The output gate determines what the next hidden state should be.

Some of the predicted graphs:

Fig. 8. Binance Coin-LSTM

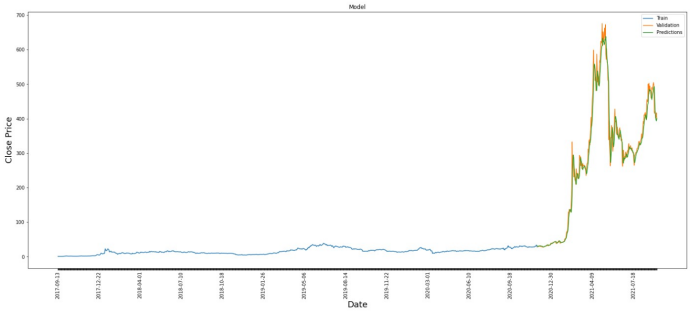
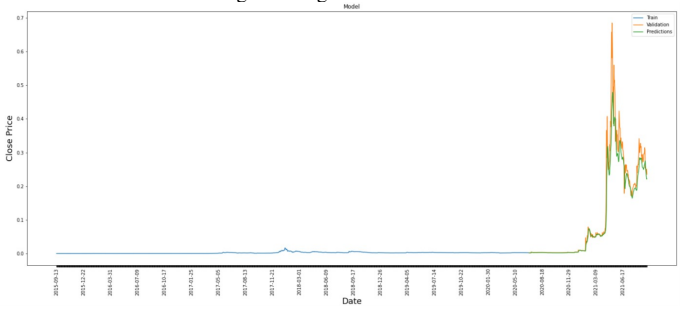


Fig. 9. Dogecoin-LSTM



V. Learning, Conclusions, and Future Work

Impact of Cryptocurrency on economy

The advent of Cryptocurrency is sure to change financial system of the world.

Some argue that cryptocurrency will negatively impact the global economy, particularly the developing ones, as it removes the need for intermediaries like banks. They primarily refer to its newness and lack of trust among a large part of users. They also cite the extreme volatility crypto coins show regularly and the lack of regulating authority. Others say the lack of regulation makes it democratic. They also argue cryptocurrency supports financial inclusion in poor countries at an unmatched rate because of increased transparency during transactions owing to its decentralised ledger system, low cost of the transaction, and its ability to beat inflation if someone decides to hold it as a store of value.

Irrespective of the arguments, El Salvador became the first country in the world in September 2021 to adopt Bitcoin, the largest cryptocurrency, as a legal tender alongside the US dollar. It must have seen some merit in its usage. Also, many other countries are either developing or thinking of developing their own cryptocurrency, called the Central bank Digital Currency (CBDC).

Learning outcome includes Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. The accuracy of the proposed LSTM is investigated by finding the root mean square error (RMSE) to determine its accuracy. The main objective of our study is to forecast the cryptocurrency prices with improved efficiency using a deep learning model and minimize the risks for the investors as well as policy-makers. We can also try using different time series prediction algorithms and compare their accuracy for a better understanding of the future behaviour of various cryptocurrencies.

Contribution of Team members

- **Neeraj Nixon**: Came up with the topic,dataset and did the analysis part of the project.
- **Shrey Ganatra**: Did most of the research coming up with relevant reference materials and did most of the report writing and helped in writing the code for LSTM .
- **Utkarsh Jindal**: Learned about LSTM Neural Network and did the prediction part of the code.

Acknowledgment

We would like to thank our professors, Prof. Amit Sethi, Prof. Manjesh K. Hanawal, Prof. Sunita Sarawagi and Prof. S. Sudarshan for teaching us this course, without which we would not have been able to complete this project. We would also like to thank all the TA's involved with this course, who helped us a lot in doing the various assignments, and completing this course.

Reference

- Dataset for cryptocurrencies[Online] https://www.kaggle.com/ysthehurricane/all-cryptocurrencies-price-20152021

- Forecasting and trading cryptocurrencies with machine learning under changing market conditions[Online] https://jfin-swufe.springeropen.com/articles/10.1186/s40854-020-00217-x
- Bitcoin Price Prediction and Analysis Using Deep Learning Models[Online] https://www.researchgate.net/publication/345262720_Bitcoin_Price_Prediction_and_Analysis_Using_Deep_Learning_Models
- Cryptocurrency Price Analysis-LSTM[Online] https://www.kaggle.com/mikaeelzaidi/cryptocurrency-price-analysis-lstm
- Cryptocurrency Price Prediction using Deep Learning https://towardsdatascience.com/cryptocurrency-price-prediction-using-deep-learning-70cfca50dd3a
- How will cryptocurrency affect the existing global economic order? https://www.ndtv.com/business/can-cryptocurrency-change-the-global-economy-2560228