

CRYPTO CURRENCY PRICE ANALYSIS AND PREDICTION USING TIME SERIES DATA

MINOR PROJECT SYNOPSIS

BACHELOR OF TECHNOLOGY

Information Technology

SUBMITTED BY

AJIT KUMAR

(2004884)

AYUSH SHEKHAR

(2004896)

CHANDAN KUMAR

(2004899)

HARSH MISHRA

(2004924)



GURU NANAK DEV ENGINEERING COLLEGE

(An Autonomous College Under UGC ACT)

1. INTRODUCTION

Although machine learning has been successful in predicting stock market prices through a host of different time series models, its application in predicting cryptocurrency prices has been quite restrictive. The reason behind this is obvious as prices of cryptocurrencies depend on a lot of factors like technological progress, internal competition, pressure on the markets to deliver, economic problems, security issues, political factor etc. Their high volatility leads to the great potential of high profit if intelligent investing strategies are taken. Unfortunately, due to their lack of indexes, cryptocurrencies are relatively unpredictable compared to traditional financial predictions like stock market prediction.

It works by using special gates to allow each Long -Short-Term-Memory (LSTM) layer to take information from both previous layers and the current layer. The data goes through multiple gates (like forget gate, input gate, etc.) and various activation functions (like the tanh function, relu function) and is passed through the LSTM cells. The main advantage of this is that it allows each LSTM cell to remember patterns for a certain amount of time. The thing to be noted is that LSTM can remember important information and at the same time forget irrelevant information.

This work includes various concepts of trends and its uses which shows the value of Bitcoin varies everyday, it would be very interesting for investors to forecast the Bitcoin value but at the same time making it difficult to predict. Also we will be predicting the Bitcoin price accurately taking into consideration various parameters that affect the Bitcoin value, to reduce the time complexity here in this paper we use an algorithm linked to artificial intelligence named LASSO (least absolute shrinkage selection operator). Paper like uses Gated Recurrent Unit (GRU) technique to establish the prediction accurately which describes excellent predictions depending on the mean absolute percentage error (MAPE). Results obtained from these models show that the gated recurrent unit (GRU) performed better in prediction for all types of cryptocurrency.

2. OBJECTIVES

These are certain objectives of this work :-

1. To identify the the most significant factors that affect the value of cryptocurrencies.
2. To enhance the accuracy of cryptocurrency price prediction and develop a more transparent and trustworthy cryptocurrency market.
3. To reduce the cryptocurrency scams and eliminate the need of brokers.
4. To provide a platform for users to access real-time information about the cryptocurrency market, including price trends enabling them to make timely investment decisions.

3. FEASIBILITY STUDY

Crypto currency price analysis involves predicting the future prices of various cryptocurrencies based on historical price data, market trends, and other relevant factors. The goal of this analysis is to provide insights into potential future price movements to help investors make informed decisions.

The feasibility of developing a crypto currency price analysis model depends on several factors, including the availability of historical price data, the accuracy of the data, and the effectiveness of the predictive models used.

Data Availability and Accuracy:

The availability and accuracy of historical price data for cryptocurrencies are critical for the development of a reliable price analysis model. Fortunately, there are several platforms that provide free and paid access to historical price data for various cryptocurrencies. However, the accuracy and completeness of this data can vary, and it may be challenging to obtain comprehensive data for all relevant cryptocurrencies.

Predictive Models:

Various predictive models can be used to analyze historical price data and make predictions about future price movements. Machine learning techniques such as regression analysis, time series analysis, and deep learning algorithms can be used to develop effective predictive models. However, the accuracy of these models depends on the quality of the data used, the features selected, and the modeling techniques employed.

Market Trends and Other Factors:

In addition to historical price data and predictive models, market trends, news events, and other factors can influence crypto currency prices. Therefore, it is essential to consider these factors when developing a price analysis model.

Overall, the feasibility of developing a reliable crypto currency price analysis model depends on the availability and accuracy of historical price data, the effectiveness of the predictive models used, and the consideration of market trends and other relevant factors. With the right

approach and resources, it is possible to develop a successful crypto currency price analysis model.

4 Methodology

For developing the system certain methodologies have been used. They are as follows: Overview of methodology utilized in this project work is described by below figure .

1. Datasets Collection - Data Collection is done using Yahoo Finance of different digital crypto currencies with the updated score of the values with the highest , lowest , avg and volume of prices . We have collected data of three crypto currencies(BTC , ETH , USDC).

Attributes of datasets :-

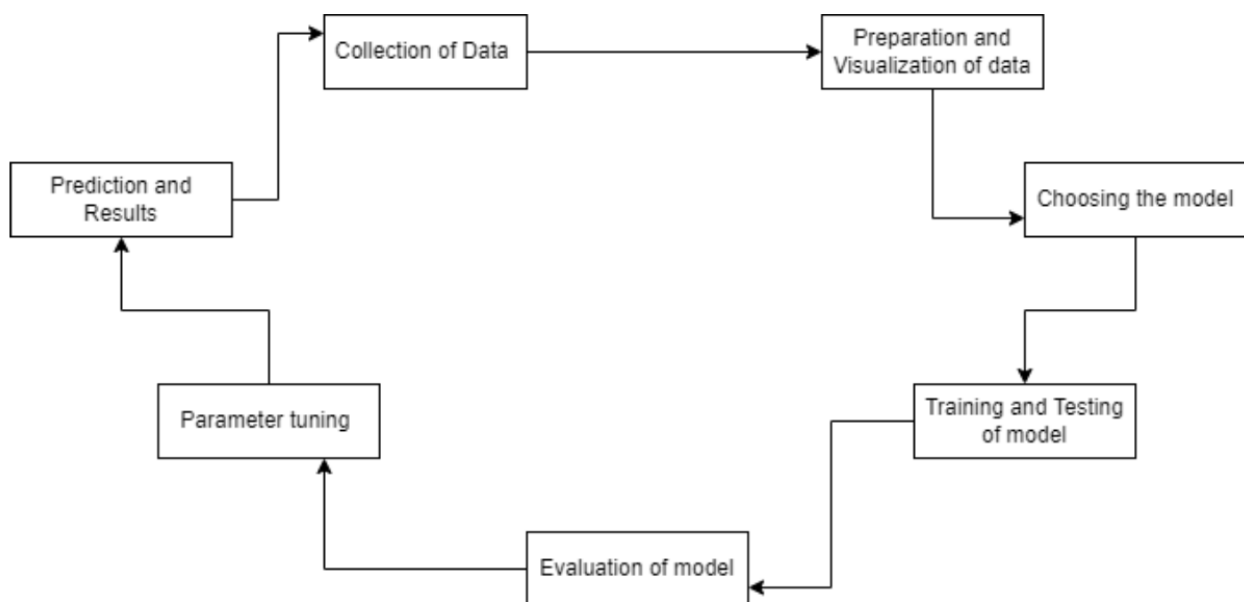
i) Date - Refers to the date on which dataset is taken with the values of the prices .

ii) Open - Refers to the opening prices of the crypto currencies .

iii) High - Refers to the highest price of the digital currency for the date mentioned in the dataset.

iv) Low - Refers to the lowest price of the crypto currency for the date mentioned in the dataset.

v) Close - Describes the tentative closing price of the crypto currencies **vi) Adj Close** - Describes the final closing price of the crypto currencies. **vii) Volume** - Refers to the final volume of price data of crypto currency



2. Diving Data Chunks - The data sets used contains .csv file and hence divided in chunks to get input for forecasting day - by - day values.

3.Importing Libraries -

Scikit-learn - Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. The library is built upon the SciPy (Scientific Python). The library is focused on modeling data. Clustering is one of the model provided by scikit-learn .

Seaborn - Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

Numpy - NumPy, which stands for Numerical Python, is a library consisting of multidimensional array objects and a collection of routines for processing those arrays. Using NumPy, mathematical and logical operations on arrays can be performed.

Pandas - Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data. Pandas allows us to analyze big data and make conclusions based on statistical theories.

Matplotlib - Matplotlib is one of the most popular Python packages used for data visualization. It is a cross-platform library for making 2D plots from data in arrays. It provides an object-oriented API that helps in embedding plots in applications using Python GUI toolkits.

4. Reading input - Input in the format of financial data thus shows the highest and lowest trends .

5. Training and Testing - Training and testing of data sets are being done in the ratio of 80:20.

6. Prediction - Prediction is done in the format of comparing with real valued currency (USD) by initializing a time forecasting series model .

5. FACILITIES REQUIRED FOR PROPOSED WORK

A) Languages Used:

● Python: The entirety of the code written for this project as in Python. Being a full-edged programming language, Python is a great tool to implement algorithms for production use. There are several Python packages for basic data analysis and machine learning. Python is an interpreted high-level general-purpose programming language. Its design philosophy emphasizes code readability with its use of significant indentation. Its language constructs as well as its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

B) Technologies used:

We would be using the following technology stack for this project:

1. Machine learning
2. Deep Learning
3. Neural Networks

B) Hardware Requirements

1. Operating System
2. Processor
3. Memory
4. Internet connection