PROJECT PROPOSAL

08.04.2022

İbrahim Ege Oral - 29299
Mustafa Ozan Yaman - 28909
Ömer Can Öztürk - 29248
Deniz Çetin - 28896

Project Name: Bitcoin Price Prediction

We will attempt to tackle the problem of trying to predict the price of Bitcoin using the graphical data analysis techniques and machine learning algorithms. The outcome for the project would be to approximate the future prices of bitcoin as close as possible. For this project we are planning on utilizing the Kaggle dataset and we will be extracting features from additional datasets and combine them with the Kaggle dataset and perform a large-scale data aggregation in order to predict the future prices. Such datasets that we will be using can be seen below.

- https://www.kaggle.com/c/bitcoin/data
- https://www.kaggle.com/mczielinski/bitcoin-historical-data
- https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory
- https://www.kaggle.com/prasoonkottarathil/btcinusd

We will be checking the data and visualizing it with the help of data visualization programs in order to deeply explore the data, communicate between different variables and how they affect each other which means we will be analyzing the relationships between different entities and factors. We will be using graphs and charts to compare and analyze the data we gathered from the initial data sets. Such as analyzing the highest and lowest price of bitcoin of all time with respect to volume and analyzing how high and low prices change with respect to time. After analyzing the given data sets, we will be including 5-6 additional new features from additional data sets and visually explain them one by one. Depending on the data set's attributes we will also use proper methods to treat the missing values (if any exists) in the data set to get a much more complete and unbiased analysis of our data set.

During the statistical and hypothesis testing phase, we will also use experimental methods to observe relationships with specific factors affecting Bitcoin prices relative to our data set. For example we will have control and treatment groups while examining the relations between the prices and that certain attribute to deeply understand whether this attribute directly, indirectly, heavily or weakly affects the Bitcoin prices in the market. We will also benefit from blocking usage for possible groups that may be affected differently by a variety of factors. For instance if we encounter a situation where the Bitcoin prices might be substantially influenced by a certain incident, we will create blocks for values during the incident and before/after the incident so that our data set values will not be generalized and will be more precise.

After having tested the data, we will compare our datasets with each other to see if there is correlation between them such as conducting statistical tests to check how the volume and close features contribute to opening price of the next day and more. Since our hypothesis

includes the correlation of the values in our data sets, we will be able to get the most accurate data to use in the machine learning step of the project. We will be applying machine learning algorithms such as KNN and Linear regression etc. Then, we will be analyzing the results of our tests and algorithms in graphs. When it comes to the machine learning part, first, we will make sure to use clean data meaning there won't be any incorrect, corrupted, duplicate, or incomplete data in our data set. At the end of our Project, we will have predicted trends of all features (open, high, low, close, adj close, volume) with our machine learning algorithm done by utilizing the steps explained above for the future time span from 09 march 2020 to 09 september 2022.

We will also be using previous research and data analysis in order to guide our own research and analysis such as A Research On Bitcoin Price Prediction Using Machine Learning Algorithms by Lekkala Sreekanth Reddy, Dr.P. Sriramya and similar studies on the field such as the ones linked below:

- https://www.ijstr.org/final-print/apr2020/A-Research-On-Bitcoin-Price-Prediction-Using-Machine-Learning-Algorithms.pdf
- https://www.researchgate.net/publication/339143532_Bitcoin_Price_Prediction_Based_on_Deep_Learning_Methods
- https://ieeexplore.ieee.org/document/8323676