CS 210 Project Proposal Report

Zeynep Yıldız - 29325

Eren Kurt - 29393

Yusuf Kerim Ciğer - 27832

Ataberk Karaman - 28945

Group: 24

07/04/2022

Bitcoin Price Prediction

Table of Contents:

ABSTRACT:

We will use the provided sources from kaggle to get datasets and combine datas and features from the additional sources for a greater exposition of the information, we will read and understand the task and requirements properly. We will first analyze and then visualize the main dataset which is given by our professor in a csv format file and combine datasets by using proper visualization methods such as creating a scatter plot since it is a time series for each and every dataset and give proper transparent graphical representations with maximized data-ink ratio. As a group we will gather information about how bitcoin works and what are the factors that determine its price such as volume, closing price, opening price, all day high and low. We will understand the process behind the purchase and selling of bitcoin and how these different factors affect its momentary price. And eventually with all this knowledge and databases we want to create a model that can predict the bitcoin prices.

INTRODUCTION:

Bitcoin is a cryptocurrency and a kind of electronic cash. It is a peer-to-peer digital currency that may be transmitted from peer to peer (p2p) on the Bitcoin network without the involvement of an intermediary.[1] It maintains track of and encrypts all peer-to-peer transactions. Since the price prediction of bitcoin uses basic factors we will utilize them to dive into data science with programming projects utilizing datasets and features with visualization using the information and knowledge we gained in the CS210 course.

---

1 https://www.researchgate.net/publication/339143532_Bitcoin_Price_Prediction_Based_on_Deep_Learning_Methods

SOLUTIONS AND APPROACHES:

To make this project happen, we will probably use a machine learning algorithm. Therefore, our tools will be Python and some features that Python libraries give to the analyst such as Pandas and Numpy. Firstly, using pandas.read_csv() we will read the CSV format file. Then we will shuffle it to have a more useful dataset since while reading the CSV format files the reading may look correct and useful but it may have been corrupted while reading the file we shuffle the data to avoid using corrupted datasets as correct datasets. After analyzing the dataset, to have a visualized format of datasets, we may use Pandas.DataFrame.plot.scatter() so that we will have a scatter plot in front of us. We are planning to treat outliers seperately to understand the reason they are different and use the proper treatment methods for these outliers.

EXPECTATIONS and PURPOSE:

Our purpose is to create a model that can predict the future prices of Bitcoin with the data we obtained from kaggle. And while doing so we will learn many things about data visualization and treatment. After this project, we will be able to analyze data via Python and some other tools like Pandas, Numpy and Matplotlib. With the help of the programs we mentioned, we are aiming to understand the importance of visualizing data transparently. We will also aim to learn how to use machine learning methods in our model so that the model can predict the momentary bitcoin prices for a short time period more accurately.

ADDITIONAL DOCUMENTS:

https://www.researchgate.net/publication/339143532_Bitcoin_Price_Prediction_Based_on_Deep_Learning_Methods

https://www.sciencedirect.com/science/article/pii/S2405918821000027