Group 12 Project Proposal - Cryptocurrency Price Prediction and Investment Portfolio Optimization

Members:
Jingbo Zhang - jingboz@andrew.cmu.edu Wei Xiao - weixiao@andrew.cmu.edu
Yi (Elaine) Zhang - yizhang5@andrew.cmu.edu Xintong Zheng - xzheng3@andrew.cmu.edu
Xi Yan - xiyan@andrew.cmu.edu

(1) a clearly defined question to be explored

We want to address two questions:

1) What is the future price of a specific cryptocurrency based on its historical data? We aim to predict the future price for each cryptocurrency based on available data such as time, open(price), close(price), combined with macro economics variables, etc..
2) How to maximize/optimize potential cryptocurrency portfolio revenue? We intend to give recommendations or suggestions for maximizing potential benefits on questions like "when to buy" and "buy what".

(2) why the question is interesting

Cryptocurrencies are active in the trading market, some people become rich overnight because of it, some people go bankrupt in just a few minutes also because of it. So, our project is interesting for two reasons:

1) Our prediction may be conducive to mitigate the investment risk. When the predicted price is higher than current price, a person should buy the stock because the stock will appreciate in the future. On the other hand, when the predicted price is lower, a person should consider selling his or her current position.
2) Our optimization may be helpful in maximizing his or her profits given the investment budget and fixed time frame.

(3) why the question is challenging

Our prediction might not be accurate, due to the high volatility of the cryptocurrencies market and the constantly changing macro economics. During the process of optimization or prediction, solely relying on the performance of cryptos are not enough, we have to also do feature engineering on macro economics factors such as GDP, interests rates, and other quantitative indicators. We have to think about the selection of models since it is not necessarily linear. It is very challenging to find a model that simulates the cryptocurrencies market as most as possible.

(4) what analytical methods can be used to tackle this question, what data sources might be needed

We will utilize time series models (ARMA, MA) and regression models on 100 cryptocurrencies to predict their prices including Open Price, Close Price, Low Price and High Price. We will also compare the accuracy of time series model and machine learning regression model on predicting the prices of various cryptocurrencies.
We will perform an investment portfolio optimization (Gurobi) with constraints (total budget, diversity of choices) to maximize the profit or minimize the risk within a certain time period (day, week, month). We will perform optimization on historical data and predicted data, and compare the combinations.

Data sources: Top 100 Cryptocurrencies Historical Dataset
https://www.kaggle.com/datasets/kaushiksuresh147/top-10-cryptocurrencies-historical-dataset
Interest rates:
https://home.treasury.gov/interest-rates-data-csv-archive
US GDP:
https://fred.stlouisfed.org/series/GDP

(5) a plan of attack with timeline and team task division

Project Timeline:
Project Proposal - Week 7
Data Collection and Model Plan - Week 8
Data Preprocessing and EDA - Week 9
Modeling and Implementation - Week 10~11
Results and Interpretation - Week 12
Report and Presentation - Week 13~14
Task Division: We plan to work on the project together temporarily. Specific tasks division will be determined based on team members preferences, skills and availability. We will try to divide tasks equally in time and energy.