ORIE 5741 Project Proposal

Benyuan He and Litong Liu

PROBLEM

Cryptocurrency industry is a relatively new and rapidly growing sector that encompasses a wide range of digital assets and related technologies. At its core, the crypto industry is built around blockchain technology, which is a decentralized ledger system that allows for secure, transparent, and tamper-proof transactions.

Cryptocurrencies are often used as a medium of exchange for goods and services, and their value is determined by market demand. Since this nascent market doesn't have many large institutions involved, theoretically, it is easier to make profit from it. In this project, we want to explore the cryptocurrency market and try to find some potential ways to take returns from it.

DATA

For this project, we consider the cryptocurrency market, and our data comes from mainly two sources:

- Historical data from cryptocurrency exchange, such as BINANCE: From BINANCE, the dataset contains 7 explanatory variables: datetime, open, high, low, close, volumn, and turnover.
- On-chain data: The historical data that are recorded on the ledger on the nodes of Ethereum chain. From one node of the Ethereum chain, we can get the following 4 explanatory variables:
  1) Transaction data: The sender and receiver addresses, the amount of cryptocurrency or digital asset being transferred, the transaction timestamp, and any additional data included in the transaction.
  2) Block data: The block height, hash, timestamp, and number of transactions included in the block.
  3) Wallet data: The balance of a particular wallet, as well as its transaction history.
  4) Token data: Token balances, transfers, and transactions for a particular token on the blockchain.

POTENTIAL METHODOLOGY

There are two potential approaches to our problems:

1) The first approach is that we can establish some regression model to predict the market value based on features like active addresses of Ethereum, block height, and etc. This allows us to utilize the fitted model to value Ethereum. The comparison between the market value and valuation of Ethereum can indicate whether it is overvalued or undervalued, providing a reference for investment decisions. The details of investment strategy will be explored in our project report.
2) The second approach is an event-driven investment strategy using on-chain data. Initially, we would do data analysis and employ hypothesis testing to justify whether there is relationship between the wallet transaction records of large institutions and the past price trends. If they do exhibit some correlation, we would utilize the past three years' on-chain data (which contains the transaction data from large institutions) to fit a regression model to predict the future price rises and falls.