Todd Gavin

DSCI510

18 November 2022

**Ideas:**

* How weather affects global food supplies?
* How sentiment on crypto twitter influences crypto currency prices?
* Forest Fire Prediction
* How do the trades of major crypto wallets influence crypto currency prices?
  + Can the trades of major crypto wallets be used as an indicator for major crypto currency price actions?
* How do the trades of U.S. politicians influence stock market prices?
  + Can the trades of U.S. politicians be used as an indicator for major stock market price actions?
* How does crypto twitter sentiment affect the prices of major NFT collections?

**APIs:**

* Etherscan API – get transactions for a particulate wallet address
* CoinMarketCap API – get current and historical cryptocurrency price data

**DSCI 510: Principles of Programming for Data Science**

Fall 2022 Homework 2  
Plan for Final Project **Due: 11/18/2022 23:59 PM PT**

* Please submit a PDF file answering the following 3 questions.
* Please be concise and informative.
* Use of bullet points is encouraged.
* No more than 1 page.

1. **What’s the name of your final project? Please describe it as a research question and provide a short description [5 points]**

**Can the timing and volume Ethereum trades of major cryptocurrency wallets be used as an indicator for major price actions of the Ethereum cryptocurrency?**

It is assumed that major cryptocurrency wallets have alpha information that gives them an “edge” or “advantage” when determining the future price of the Ethereum cryptocurrency. Therefore, they may have the ability to perform Ethereum trades with a greater success rate of being profitable.

This project aims to analyze how the timing and volume of Ethereum trades (buys and sells) from major cryptocurrency wallets correlates with major price actions (up or down movement) of the Ethereum cryptocurrency. For example, these trades can be considered a “strong” indicator if a major cryptocurrency wallet buys Ethereum, and a larger volume of it, a specific time, prior to a significant increase in the price of the Ethereum cryptocurrency. Therefore, “buying” should indicate an increase in Ethereum’s price, and “selling” should indicate a decrease in Ethereum’s price.

1. **What data sources are available? Could you find multiple data sources? How are you going to collect them? How many data samples are you going to collect? [5 points]**

**Ethereum Transaction Data (JSON):**

* To collect data on the Ethereum trade transaction of major cryptocurrency wallets, I will be using the Etherscan API.
* The Etherscan API can return to me all the transactions for a specific time and cryptocurrency by account (input wallet address).
* I plan to collect the 100 most recent Ethereum trade transactions from 10 major cryptocurrency wallets. (1000 transactions total)

**Current and Historical Ethereum Price (JSON):**

* To collect data on the current and historical price data of Ethereum, I will use the CoinMarketCap API.
* For a specific cryptocurrency (in this case, it’s Ethereum), I will retrieve its price data based on the timing on the timing of trades from the major cryptocurrency wallets.
* I will retrieve the price data on 7-day intervals for 10 years (520 data points).

1. **What kind of analyses or visualizations do you want to do? [5 points]**

* Linear Regression Model of Trade Transaction Timing, and Ethereum price movement. (Independent Variable: Date and Time)
* Bar Graph of the success rate of profitability of based on the transactions from the major cryptocurrency wallets.
* Line Graph of when and where major cryptocurrency wallets bought into and sold out of the Ethereum cryptocurrency.

Rubric:

1. The project name should:
   1. Reads like a research project. 1points.
   2. Have a short description. 2 points.
   3. Indicates the data source. 1point.
   4. Indicates the analysis method. 1 point.

2. The data sources should:

1. Specify the general nature of data, e.g. weather, hotel price, airfare, etc. 1 point.
2. Specify exact data sources, e.g. a link to the dataset, a link to an API, or a link to a website to parse. 1 point.
3. Have 2 or more data sources. 1point.
4. Includes a short description and your plan. 1 point.
5. The data source should be large and recent. It should contain at least 100 samples (recommends 1000 samples or more), and is no older than 10 years from now. 1 point.

3. The analyses and visualizations should:

1. Feasible.Itshouldbeananalysisthatcouldbedonewithyourdatasource.2 points.
2. Meaningful. It should be something that is not obvious, non-trivial, not in common sense. 2 points.
3. Impactful. It should be something having an impact and is interesting to the audience. 1 point.