Project Proposal
Bitcoin Price Prediction Using Twitter Comments and Commodity Price

Xinyu Feng (xf55) Yiran Hao(yh494) Tiange Hou(th527)

October 2021

According to recent news, Morgan Stanley doubled its exposure to Bitcoin through Grayscale shares, while JPMorgan CEO still wouldn't buy it though he predicted Bitcoin price to rise by 10 times. Investors hold various attitudes towards Bitcoin, whether optimistically believing it an investment opportunity or pessimistically believing it a market bubble.

1 Problem Statement

We are wondering how the prices of Bitcoin are influenced by following factors:

1. **Commodity Market Performance**

   Bitcoin is regularly referred to as new gold, digital gold or gold 2.0. Similar to Bitcoin, commodies like gold and crude oil derive much of its value from limited supply and growing consumer demand during periods of increased inflation and government regulation fears due to their hedging properties. Therefore, we analyzes the dynamic connectedness among bitcoin, gold, and crude oil volatilities and returns, synthesizing an effective factor to predict the trend of Bitcoin price.

2. **Online Social Media Comment**

   JP Morgan launched "Volfefe Index" back in 2019, which is an stock market index to reflect the influence of US former President Trump tweets on US market.

   Such index inspired us to think about the how the comments or insights from social media platform could influence the market of Bitcoin, since the cryptocurrency market depends on people's thoughts, comments and perception. Therefore, we would uses Twitter comments to analyze and forecast the movement and prices of the Bitcoin market.

2 Description of Dataset

1. **Bitcoin Price Dataset**

   We obtained historical data of Bitcoin's price from October 1st, 2018 to September 30th, 2021 at Yahoo Finance including features on volatility, trading volume, daily close price, etc., as well as Statista, including Bitcoins' supply share (Bitcoins' supply vs. Total supply), etc.

2. **Commodities - Gold & Crude Oil Price Dataset**

   This data set consists of daily US dollar price series of gold and crude oil ranging from 1 October 2018 to 30 September 2021. All the series are readily available at Yahoo Finance. However, since gold and crude oil do not have weekend values, Bitcoin's price values for Saturdays and Sundays were removed from the data set.

3. **Twitter Comment Dataset**

   We obtained a data set from kaggle, which contained tweets with hashtag #Bitcoin or #btc. We plan to use features like the followers number, the date of the tweet post, whether they are re-tweeted, and the overall active user to find its relation with the movement of the Bitcoin market.

3   Reference

1. https://www.kaggle.com/kaushiksuresh147/bitcoin-tweets

2. https://www.marketwatch.com/investing/cryptocurrency/btcusd/download-data

3. Wołk, Krzysztof. (2019). Advanced social media sentiment analysis for short-term cryptocurrency price prediction. Expert Systems. 37. 10.1111/exsy.12493.

4. https://cointelegraph.com/magazine/2021/10/02/morgan-stanley-gbtc-alibaba-crypto-mining-gear-halt-scenario-6m-btc-price-hodlers-digest-sept-26-oct-2