

# CMPS 287 Project

Team K Members:

Hiba Lubbad 201902178

Hamza El Sheikh 201904899

Dany Tarabay 201903983

26 March 2021

## Phase One

### 1 Topic

Cryptocurrency Price Prediction using real-time data, official reports, influential figures, and market ambiance.

### 2 Cryptocurrencies

Due to time limitations and the complex nature of the project, we will narrow our focus to two cryptocurrencies: Bitcoin and Ethereum. These are the two largest cryptocurrencies in terms of market capitalization representing over **1.225 trillion dollars** in combined value.

### 3 Problem

The cryptocurrency market is an unstable market due to the rapid changes and some affective factors. News is one of the factors that can change the whole market, for example, Wall street statements, Elon Musk tweets, European Central Bank statements, and other public figures. Another important factor is the adoption of these currencies, or in other words the public ambiance towards this technology.

### 4 Impact/Significance

Achieving a good accuracy for stock market prediction will have a significant impact on the decisions of cryptocurrency investors leading to potential increased profits. Additionally, our model can be used by the government in policymaking to regulate the market and financial researchers towards more informed studies

of the cryptocurrency market behavior. Finally, low and middle-class people are more involved in investing due to the increased adoption of on-demand investing applications like Robinhood, which lowers the barriers for such communities to invest. However, these communities are more vulnerable to market changes, increasing the importance of our project.

## 5 Resources

- Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis
- Investigating the Problem of Cryptocurrency Price Prediction: A Deep Learning Approach
- Stock Forecasting using Machine Learning
- Short-term stock market price trend prediction using a comprehensive deep learning system
- Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models

## Phase Two

### 6 Determinants that have an impact on the BTC price

According to the following research paper titled “Forecasting the price of Bitcoin using deep learning” by Mingxi Liu et al.[1], the authors have presented research that correlated 40 determinants to the price of bitcoin. For example, they cited the findings of Alstyne and Marshall (2014) that the Dow Jones index, the euro-dollar exchange rate, and the oil price have significant impacts on the price of Bitcoin. We base our collection of datasets based on a chosen 27 determinants that we found most relevant and eliminated factors such as “Exchange rate of US dollar to Norwegian krona”. For each of the determinants below, we collected daily data from 2015 till the present that we will merge to create a finalized dataset for our project. The determinants are listed below.

### 7 Dataset

**Bitcoin info (8):** Date - Bitcoin price - Open - Close - High - Low - Trading Volume - Market Cap

**Mining info (3):** Mining difficulty - Mining award - Average Hashrate

**Public attention (4):** Google Trends - Twitter tweets - Baidu index - Reddit posts

**Macroeconomics Environment (12):** Brent Crude futures - CBOE volatility index - NYSE Copper Futures - Dow Jones industrial average - New York Stock Exchange Gold futures - HuShen 300 index - Korea composite index - Singapore Straits Times index - Standard & Poor’s 500 index - FTSE 100 index - Tokyo’s Nikkel index - US dollar index - WTI oil future

*We’re going to have a dataset of 27 columns.*

### 8 Twitter Data

Our long-term goal is to scrape through Google trends, Twitter, Top 10 news headlines (Optional), and Baidu. Each row should be iterated over an algorithm (TextBlob, Haven on demand, Tweepy) that generates the Subjectivity, Objectivity, Positivity, Negativity, and neutral ratios. However, for our current plan, we will start by scraping Twitter data.

→ After Getting these ratios we have to add them into the dataset as a new indicator that may impact the price of BTC. Then we have to divide our dataset into training, testing, and validation datasets.

## 9 Model Selection

According to the Literature referenced below, the models having the highest accuracy and widely used to predict Bitcoin prices are LSTM and ARIMA. After finalizing our dataset, we look forward to verify and choose one of those models for our project.

## 10 References and Citations

- Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis
- Investigating the Problem of Cryptocurrency Price Prediction: A Deep Learning Approach
- Stock Forecasting using Machine Learning
- Short-term stock market price trend prediction using a comprehensive deep learning system
- Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models
- Mingxi Liu, Finance Research Letters

## 11 Dataset Links

- STI Index
- GSPC index
- BTC KRW Index
- CSI 300 Index
- VIX Index
- Baidu Historical Price
- F2pool info ( Mining info )
- Bitcoin Market Capitalization
- BTC Historical Data

- FTSE 100 Index
- S&P 500 Index
- Brent Crude Oil Prices
- Dow Jones Industrial Index
- Gold Price
- S&P 500 Futures Index
- Tokyo's Nikkei Index
- US Dollar Index
- WTI oil futures