

# Proposal

## Members

- Haruki Moriguchi (# 260665818) : *haruki.moriguchi@mail.mcgill.ca*
- Edwin H. Ng (# 260732345) : *edwin.ng@mail.mcgill.ca*

## Predicting Bitcoin Price based on Sentiment Analysis on Social Media

In recent years, Bitcoin has gained enormous popularity. After receiving many media coverages in 2017, the price went up drastically from 1,000 to 20,000 CAD, from which it has since gone down. In fact, this sudden increase in price is not surprising, since behavioral economics states that there are correlations between the public sentiment and the financial market. Fortunately, with the advent of social media, the information about public feelings has become abundant, where Twitter has received a lot of attention from researchers. The aim of this project is thus to test the hypothesis that, in addition to past historical prices, the public sentiment also influences the stock market. We will focus particularly to the case of the Bitcoin prices. The plan of the project is as follows. A corpus of tweets will be pre-processed and represented using word2vec model rather than N-gram model, since word2vec allows for sustainability in word meaning across different contexts. We will then build a public sentiment classifier from these tweets by training different classifiers that will classify tweets as being “negative” or “positive”. The best sentiment analysis model will then be applied to tweets that are related to Bitcoin or cryptocurrencies. Once sentiment analysis on Bitcoin tweets has been assessed, we are then going to also extract its past historical price movements. The idea is that the future movement of a stock price should mainly be able to reflect its past tendencies, but that the public opinion will also influence its future paths. Therefore, we will then train a final machine learning model (SVM, neural networks, etc, and compare results) that will classify the Bitcoin price movement as being “up” or “down” at each time frame (minutes, hours, or days) based on both its past historical movements and on its public opinion in that time frame. That is, the public opinion obtained by the sentiment analysis described above will be fed in to this new model. It is interesting to note that Bitcoin prices may also have a delay response time to the public opinion at a certain time frame. For example, a tweet on day 1 might only impact the price on day 2. It is then interesting to also test multiple possible reaction times from the public sentiment to the price, and to tune that hyperparameter.