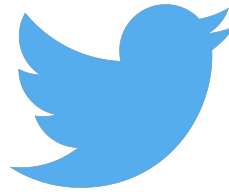# CryptoSentiment

Anuj Suhag

Parameters impacting token price

Bitcoin Volume

Public Sentiment

The Problem

Can volatility in bitcoin prices be predicted with public sentiment?

Have bitcoin prices started to stabilize ?

Clients - Potential and Current

Potential Clients

1. VIA Inc, Boston
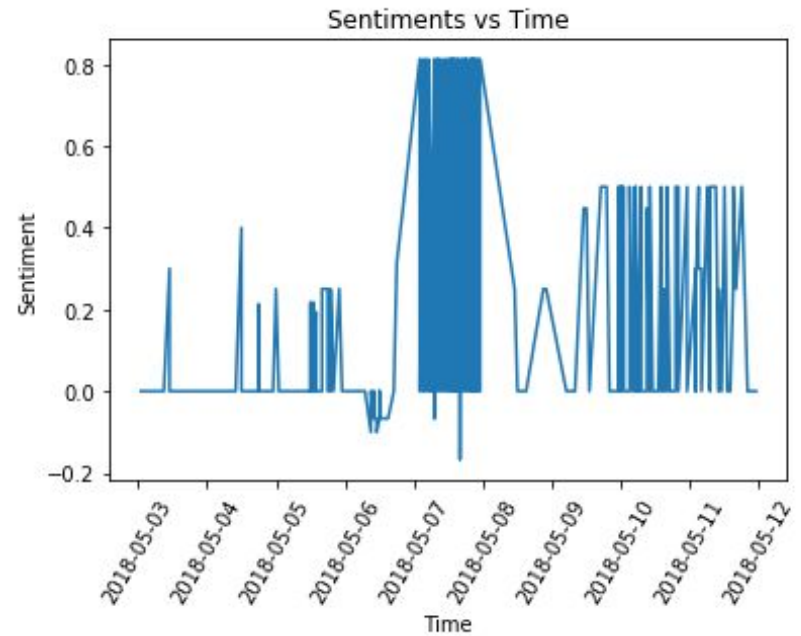2. Intek LLC, Houston

Current Client

1. Centareum, Singapore

Data Acquisition

- Tweets collected from #cryptocurrency with Tweepy
  - April 11th to May12th,2018 for model building
  - May 20th to May 25th,2018 for model testing

- Bitcoin Prices collected from https://coindesk.com/price/

Exploratory Data Analysis

Closing Price vs Time / Sentiments vs Time

- Bitcoin prices are consistently falling since they peaked in June 2017
- Literature reasons a relation between public sentiment and bitcoin prices
- Sentiment polarity varis rapidly and visual investigation is insufficient

Illustration of tweets correctly identified as positive

" With an audience of 10 million per month (120 million per year) at a rate of € 3.0 per user - the profit will be € 360 000…"

" It's exciting to announce that Animation Vision Cash (AVH) is listing on JNB Exchange！Trading will commence at 16:0… "

"⬆ +10 Ethereum price breaks above $800.00 with strong Bullish trend"

Example of tweets with no sentiment

We will be airdropping 2000 $NBT tokens each to 60,000 Subscribers

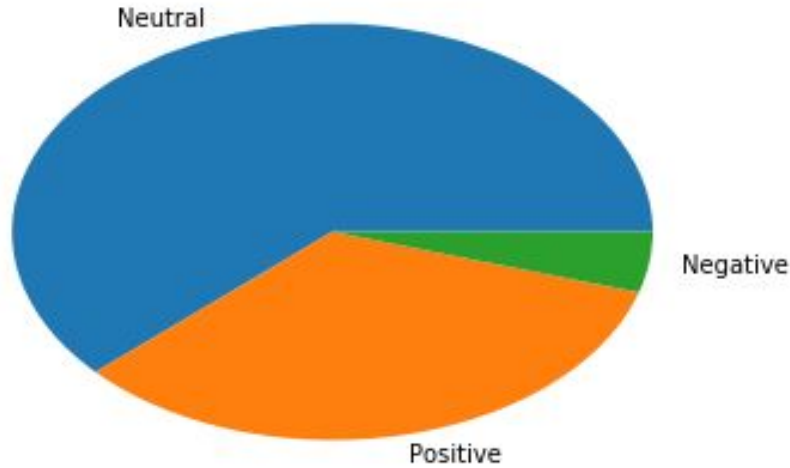#BTC 8442.16$ | 7058.36€ #XRP 0.69$ | 0.58€ #ETH 676.88$ | 565.93€ #LTC 141.00$ | 117.88€ #DASH 400.13$

#Cryptos: #BTC 8544.57$ | 7145.82€ #XRP 0.69$ | 0.58€ #ETH 677.00$ | 566.18€ #LTC 141.64$ | 118.46€ #DASH 401.22$

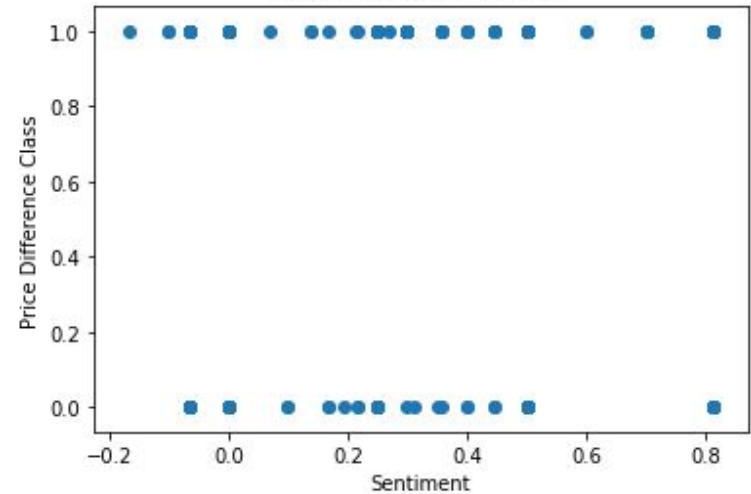Following are two examples of tweets that were mislabeled with 0 polarity by WordBlob.

"4 hours left! The registrations haven't stopped. If you haven't signed up yet hurry up and do it……"

"cryptocurrency Price Analysis for #Bitsend #BSD : ⬇ ☐ Last Hour Change : -1.05 %......"

Pie chart of Sentiment split

Sentiment vs Price Class

- Price Class :
  - Delta Price > 0 : Price Class = 1
  - Delta Price <=0 : Price Class = 0
- Majority of sentiment polarity is neutral followed by positive and negative being the least observed polarity.
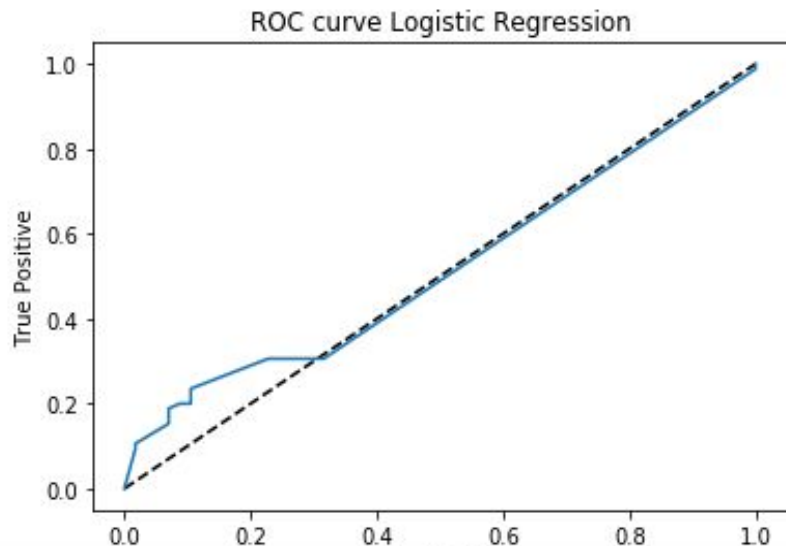- Sentiment vs Price Class indicates no observable pattern

Machine Learning

- Technique selected : Logistic Regression
  - Class Imbalance exists. Use weighted_class='balanced'

- Target : Price class
  - Model Making using data from April 11th - May 12th
  - Model Testing using data from May 20th - May 25th

- Model hyperparameter tuning using GridSearchCV

Model Performance

Best hyperparameters for logistic regression {'C': 0.001, 'penalty': 'l2'}:

Score of Logistic Regression 0.49696969696969695:

Tuned hyperparameters

Mean Accuracy of the model

Model Performance

	Class 0 Predicted	Class 1 Predicted
Class 0 Actual	39	18
Class 1 Actual	59	26

```
Area under the ROC curve 0.517234262125903

Classification Report :
             precision    recall   f1-score    support

          0       0.40      0.68       0.50         57
          1       0.59      0.31       0.40         85

avg / total       0.51      0.46       0.44        142
```
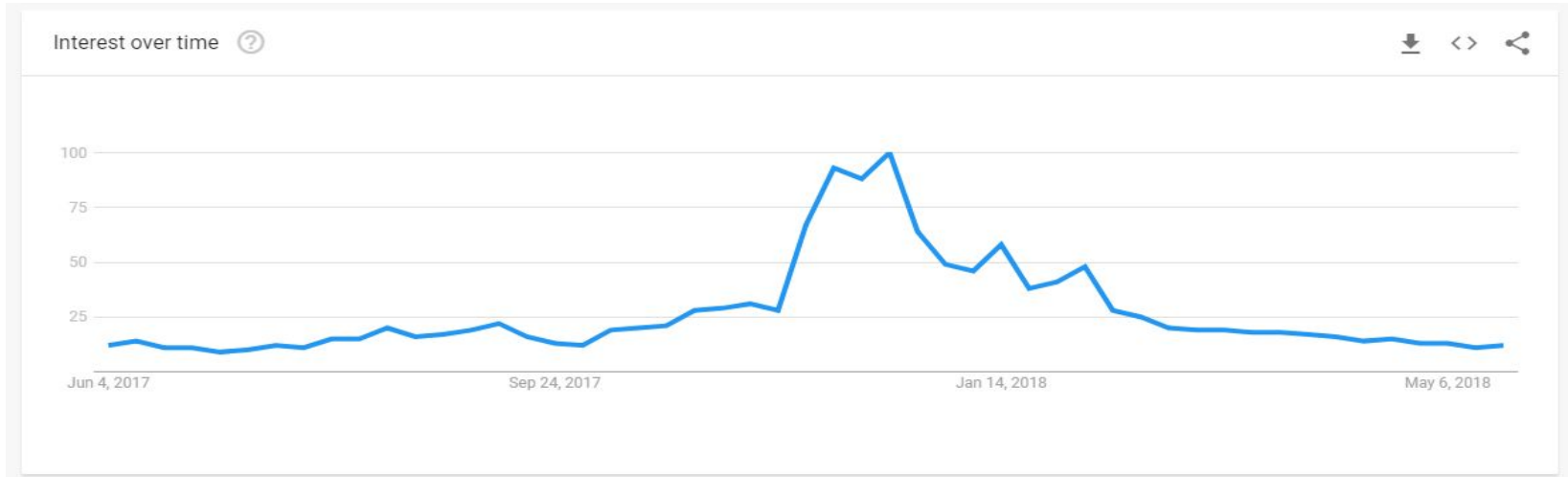
Results and Discussions

Following are the results of blind test performed on the model using tweets from May 20th to May 25th.

- The model correctly labeled 49.4% of sentiments with price class 1
- 50.6% of sentiments were incorrectly labeled as price class 1

"...There are allegedly several **very large holders of Bitcoin** (and other cryptocurrencies). The concentration of wealth gives them the incentive to push prices higher."

"...**Miners**, some of which may have large positions, are also motivated to see prices rise as it should mean more money for them." - May 28th, " I would be shocked if bitcoin prices weren't manipulated", Peter Tchir, Forbes

- For the month of May 2018, there is no visible correlation between public sentiment and bitcoin prices. The results are inline with the finding of major market manipulation by high net worth individuals and token miners, indicating markets driven by sentiments of few parties of interest and not the emotions of the mass.

- Initial investigation, when compared to prominent literature review, suggests the general polarity of sentiments and volume of tweets has gone down since early 2018. Indicating a clear decline in public's interest in crypto tokens suggesting market stabilization