

Do Tweets Have an Effect on Bitcoin

Abdullah Alathel

Goal

The goal of the project was to find if tweets sentiment have an effect on bitcoin price, or if there is a relationship between tweets sentiment and bitcoin price.

Data Source

- First dataset is from Kaggle which have been Scrapped from twitter 2016-01-01 to 2019-03-29, Collecting Tweets containing Bitcoin or BTC
 - The dataset originally contained more than 18M tweets, my original goal was to process the whole dataset, but I ran into few problems, also due to the time constraints, I was forced to limit it to the year 2018 only
- Second dataset is from Yahoo Finance and contains bitcoin daily historical data for the period of 2018-01-01 to 2018-12-31

Algorithms

- Twitter-roBERTa-base for Sentiment Analysis
 - pre trained on ~58M tweets and fine tuned for sentiment analysis
 - It requires minimal text preprocessing
- Least Angle Regression for closing price prediction
 - Gave the best scores on cross validation testing

Tools

- Pandas
- Numpy
- Hugging Face transformers
- Pytorch
- Plotly
- AutoViz
- SweetViz
- Fasttext
- PyCaret

Process

- First, tweets text was preprocessed, and spam accounts and tweets were filtered out, also filtered out any non-english tweet
- Concatenated sentiment analysis scores using Twitter-roBERTa-base
- Grouped by day and took mean for sentiment scores and sum of retweets, and merged with bitcoin data
- Trained model on data of first 10 months and used last two for testing
- Used PyCaret to compare and test models

Conclusion

MAE	MSE	RMSE	R2	RMSLE	MAPE
3151.7475	11544013.1256	3397.6482	-8.1333	0.6082	0.7930

The price prediction was far off, and I was not able to find any relation between tweet sentiments and bitcoin price. Blue line is the actual price

