



Shawn Anderson, Ka Hang Jacky Lok, Vijayavimohitha Sridhar

LITECOIN

ETHEREUN

AUTOMATICALLY LINK

ETH

MARKETRETWEET

TOKEN GREAT SUCCESSFUL

CASH

WORLD

STELLAR

CHANGE MONEY

NEO

WALLET

MP DAY PRICE

Motivation

Cryptocurrencies are programmable digital assets which have perceived value due to designed scarcity. The first cryptocurrency, Bitcoin was created in 2009. Today, there are thousands of cryptocurrencies with a combined market cap of over \$250 billion dollars USD.

In 2017, the world saw a peak in the 'hype cycle' of cryptocurrency technology. As data scientists, we do not believe in hype, we can only trust the data.

This project is an open-ended investigation into what insights can be gained from collecting and analyzing data from the web in pursuit of understanding the reality of the cryptocurrency ecosystem.

Goals

- 1. Visualize the cryptocurrency ecosystem
- 2. Identify factors which effect the value of cryptocurrencies and the opinions of market participants
- 3. Explore the possibility of predicting the future value of cryptocurrencies based on past data

Approach

- Implementation of scraper modules for price, Twitter, GitHub, and Wikipedia News data
- Data cleaning, filtering, imputation, and integration
- Visualization and statistical analysis on data, using insights to further refine prior steps
- Formatting of price data into observation, target pairs for training and testing
- Deep learning architecture design with Keras

- Language Processing

- > Our Train/Validation/Test sets are of sizes 34457 / 8614 / 1396 , the
- Training is done on GPU (Nvidia Geforce GTX 1050)
- \succ CNN trained for 40K epochs in 2 hours and 8 minutes with 11,539 trainable parameters
- trainable parameters
- \succ LSTM trained for 10K epochs in 2 hours and 28 minutes with 4,193





Visualizing and Forecasting The Cryptocurrency Ecosystem



Price Forecasting

We trained two models using price data with CNN and LSTM

 \succ Given 60 days of price history for a coin, our models will output a real value between 0 and 1. A value over 0.5 is a prediction that the price will rise tomorrow, under 0.5, the price will fall.

test set consists of 15 days of most recent history

 \succ CNN achieves 58% validation accuracy for binary next day prediction > LSTM achieves 56% validation accuracy for binary next day prediction Our models recommendations for April 6th 2018:





- for use in production.
- Portfolio recommendations
- Long Term Coin Evaluations
- Trend adjusted analysis

Conclusion and Future Work

Data visualization has provided insights into the nature of cryptocurrencies and their relationship to news, GitHub and Twitter. \checkmark We have shown CNN to be a reliable architecture for cryptocurrency price forecasting. With additional data and tuning, we see a potential

Future work Linear and Stochastic Forecasting with Transfer Learning

https://github.com/LinuxIsCool/733Project