Social networks sentiment analysis for cryptocurrencies

Using discriminative aggregation for price prediction

MS&E 448 Midterm Presentation
Boris Beltinoff and Emile Clastres
April 27, 2021

Overview

- **Project outline**

    - Project goal: predict cryptocurrencies price based on social networks sentiment

    - Technology used: Twitter scraper not based on API

    - Preliminary analysis: Tendencies of tweets with respect to number of followers and tweet frequency

    - Future analysis: ML/RL on tweets based on language and timing, joined by data from cryptocurrencies exchanges

- **Preliminary data visualization**

    - Word Cloud: helps visualize the text content of relevant tweets

    - Tendencies of tweets with respect to number of followers and tweet frequency

    - Word Clouds for tweets from top 10 users by number of followers and top 10 users by tweet frequency

Boris Beltinoff and Emile Clastres

Project Outline: Project Goal

◈ We want to predict price of one of or several cryptocurrencies in order to trade them effectively using social media sentiment.

◈ We also want to classify/rate users when it comes to their crypto tweeting activity: prefer "high-quality" users

◈ We will be using Twitter scraping, data from cryptocurrencies exchanges, Machine Learning and Reinforcement Learning in order to achieve this goal.

Boris Beltinoff and Emile Clastres

Project Outline: Technology Used

◈ TweetScraper can get tweets from Twitter Search. It is built on Scrapy without using Twitter's APIs. The crawled data is not as clean as the one obtained by the APIs, but the benefits are you can get rid of the API's rate limits and restrictions. Ideally, you can get all the data from Twitter Search. (https://github.com/jonbakerfish/TweetScraper/blob/master/README.md)

◈ Uses Twitter Advanced Search, which allows robust filtering parameters.

◈ Search results frequently need to be further filtered. They are google-like in the sense that they include tweets outside of the given parameters.
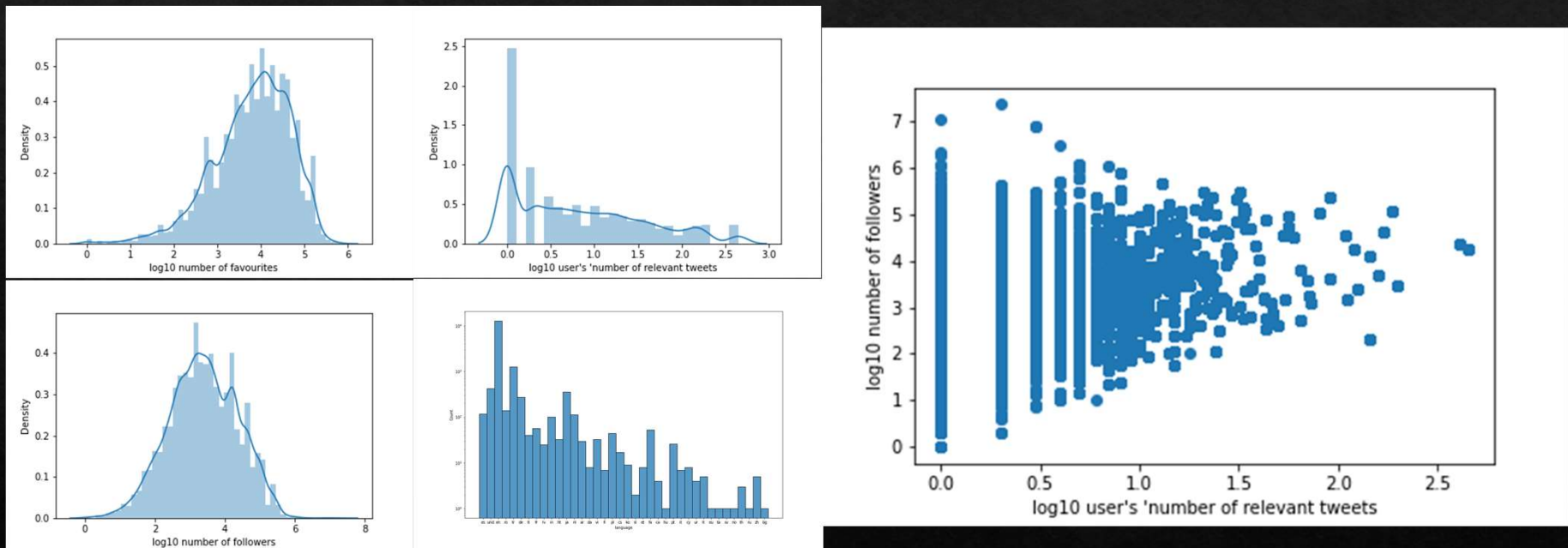
Boris Beltinoff and Emile Clastres

Project Outline: Preliminary Analysis

◈ Rudimentary analysis of most frequent words used in tweets about LTC (Litecoin) as one example of Twitter cryptocurrencies activity.

◈ Distribution of tweets with respect to the user's number of followers, tweet frequency, amount of likes, and languages used to tweet about LTC
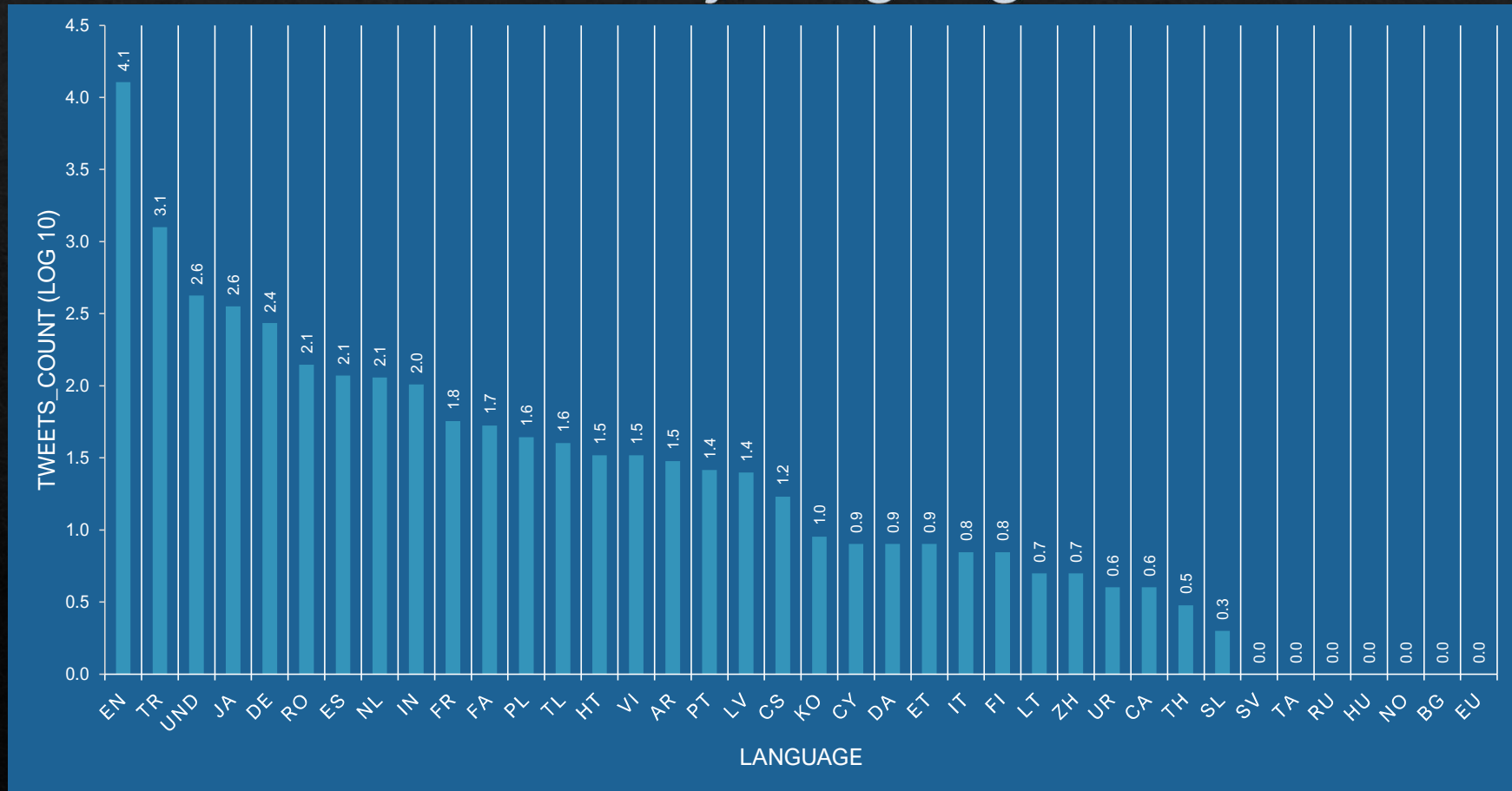
Boris Beltinoff and Emile Clastres

Project Outline: Future Analysis

◈ We will join the sentiment data with data from cryptocurrencies exchanges to identify trends in price movements preceded by Twitter activity.

◈ Plan to do ML/RL on tweets based on language and timing.

◈ Will consider several geographical, timing, and cultural factors.

◈ Interesting to see if meme cryptocurrencies have more immediate price reactions to social media activity, e.g. Bitcoin, Dogecoin

Boris Beltinoff and Emile Clastres

Tweets distribution

Boris Beltinoff and Emile Clastres

Tweets distribution by language

Aggregate word cloud

Boris Beltinoff and Emile Clastres

Word clouds for top 10 users by followers (left) and frequency (right)

Boris Beltinoff and Emile Clastres