

Analysing the trend and predicting the future of blockchain technology using twitter data

Aim: Analysis of blockchain data on twitter using *Tweepy* (Python library).

The aim of this project is to examine the blockchain discussions on twitter and analyzing the trend of blockchain and demonstrating the future of blockchain technology in multiple industrial applications. In order to identify and further the development of the blockchain technology, this project reviews the Twitter data on the blockchain using various Natural Language Processing (NLP) tools to extract knowledge from unstructured text data.

1. Extract real-time data from twitter
2. Pre-processing the data (lemmatization)
3. Clustering, Labelling, Topic modelling (LDA or NMF)
4. Cluster Map visualization
5. Visualization of trend of different domains

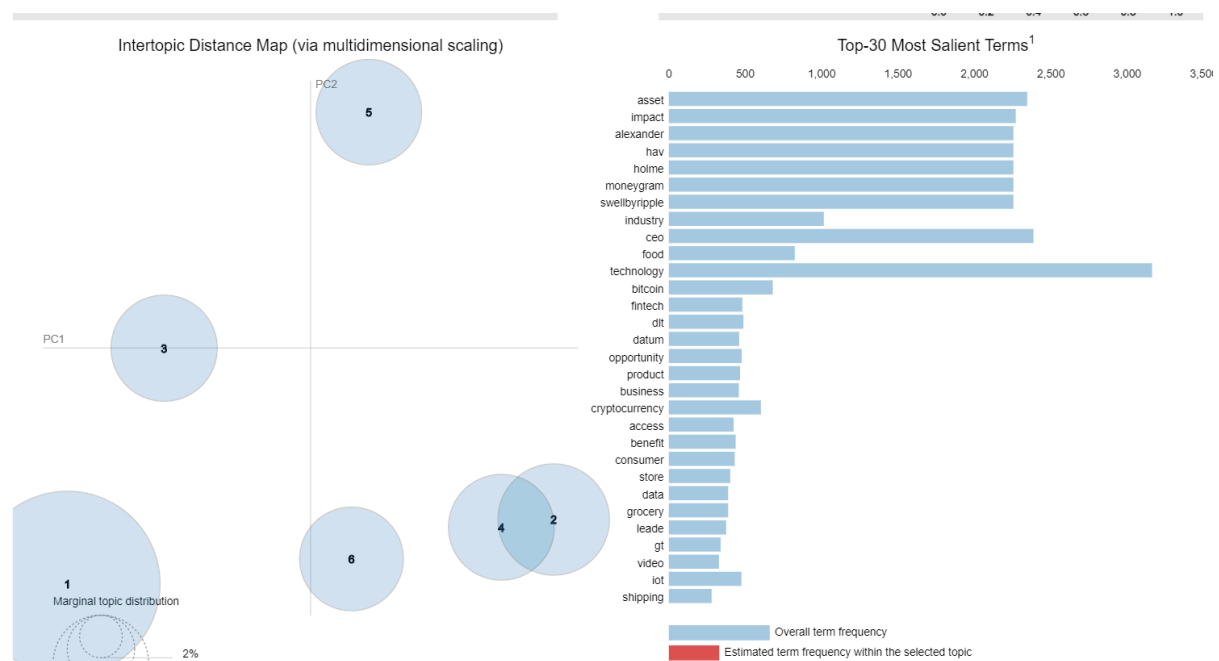


Fig 1. Overall term frequency and estimated term frequency within the selected topic

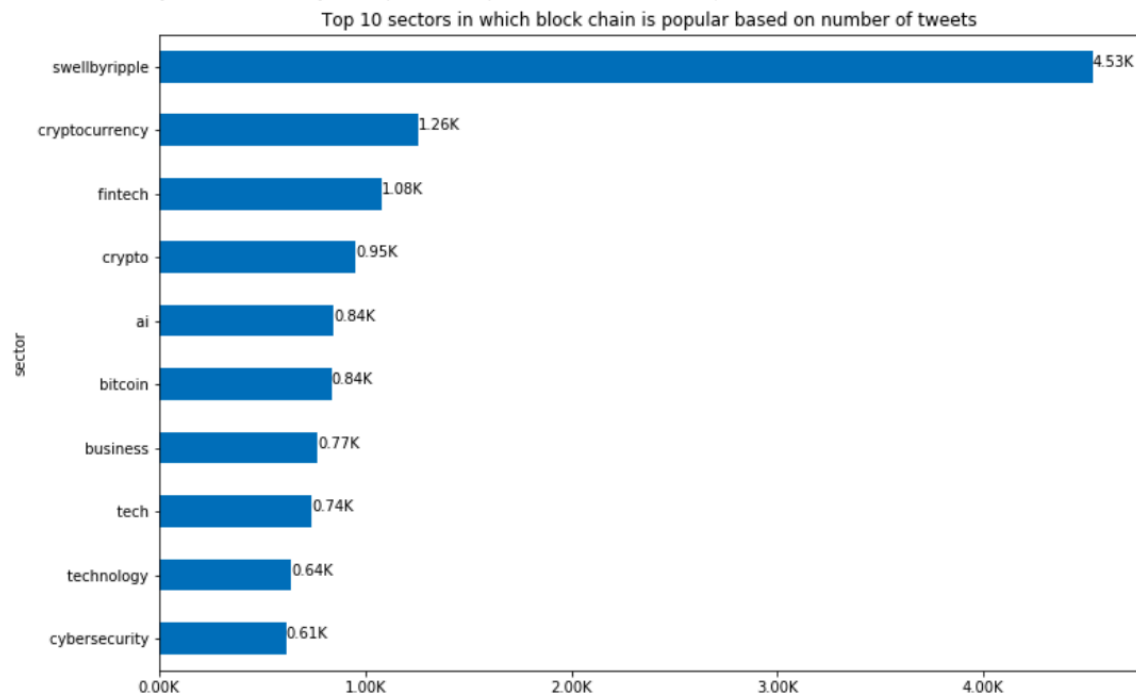


Fig 2. Number of tweets per sector

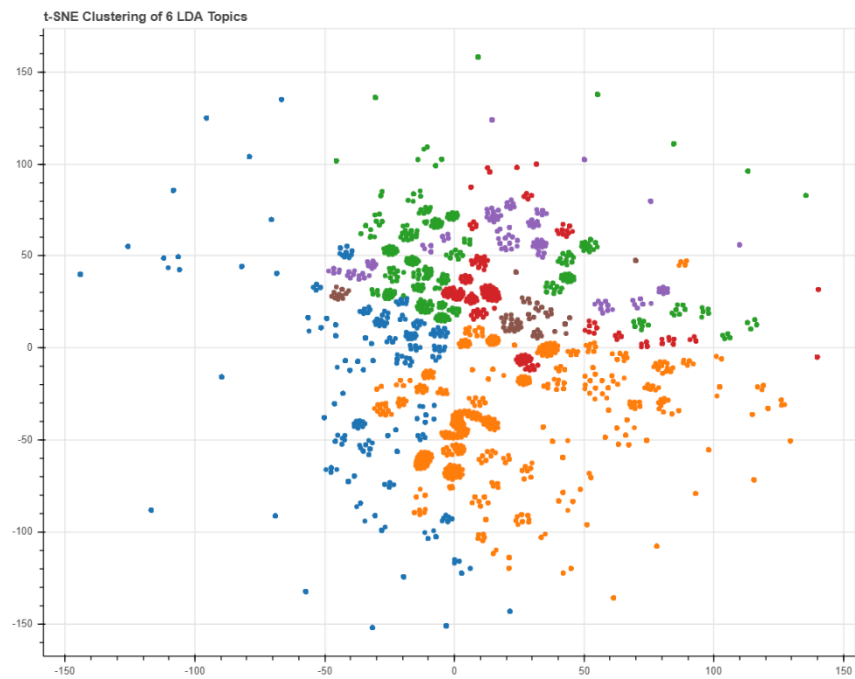
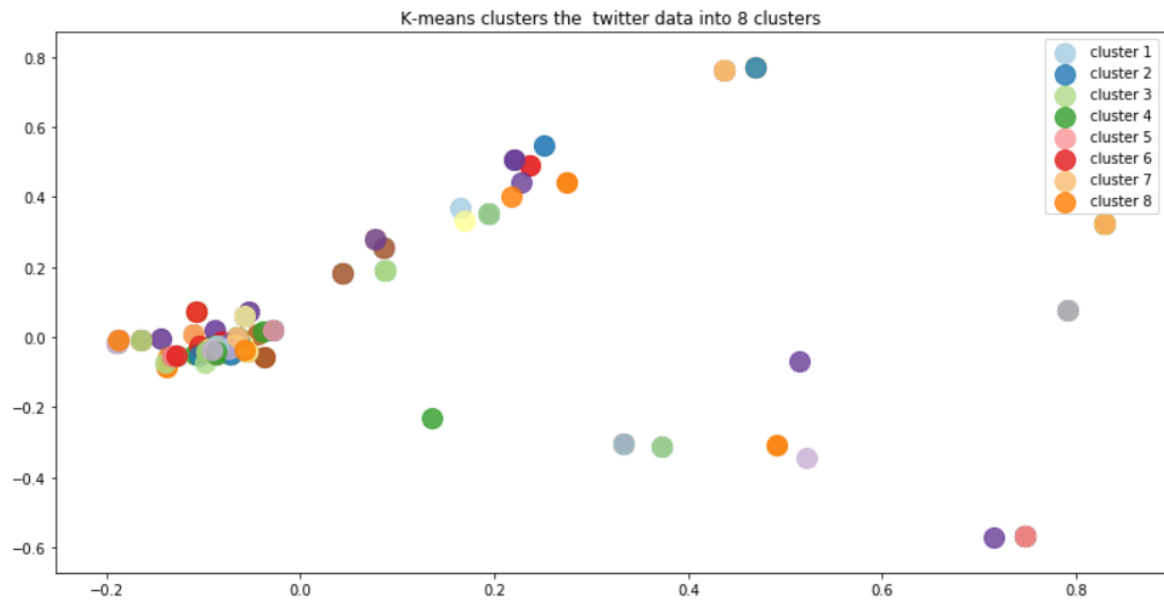


Fig 3. Bokeh Plot



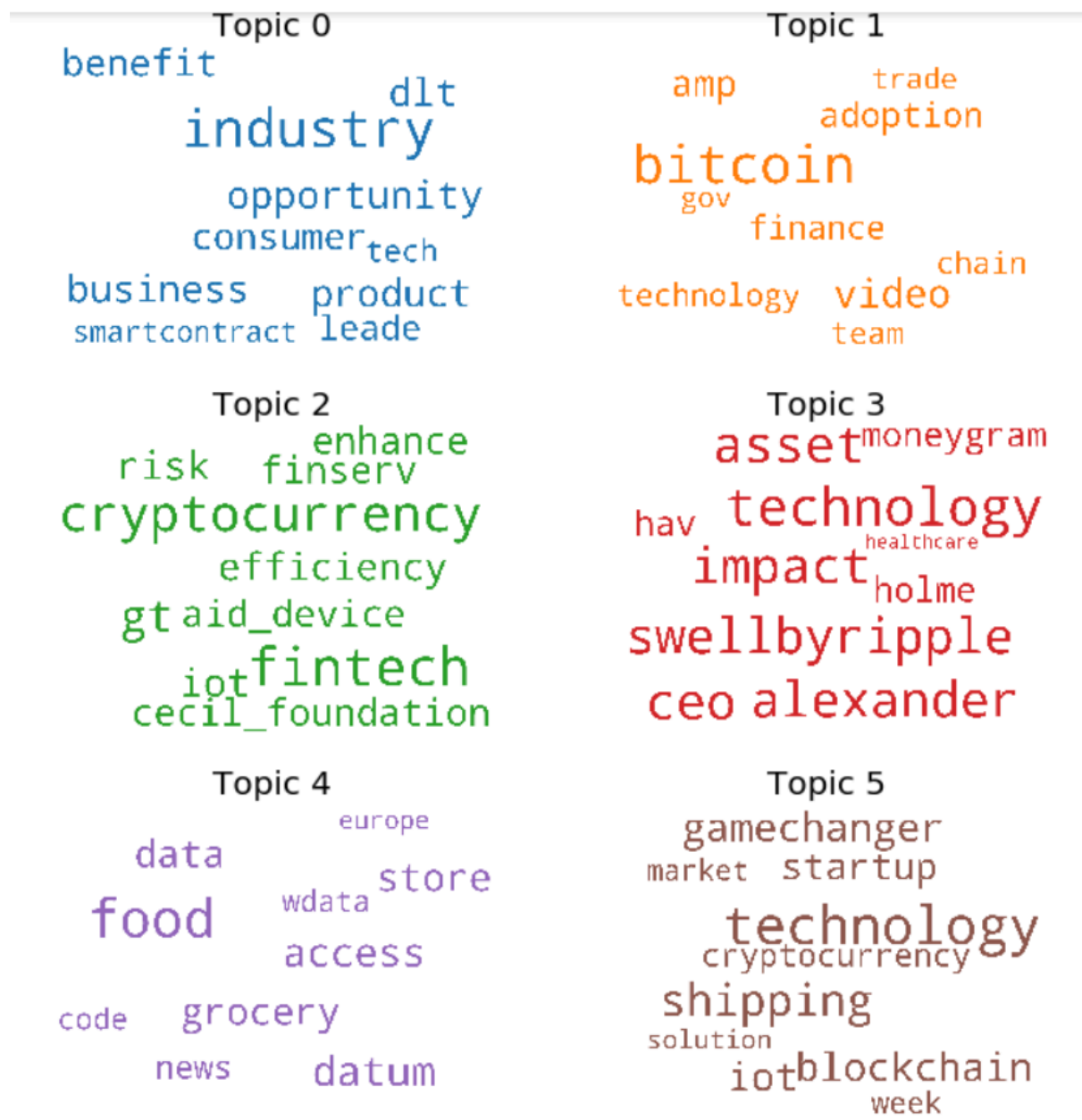


Fig 5. Word cloud topic modelling example