

Using Natural Language Processing and Neural Networks on Articles to Predict Change in Cryptocurrency Value

Preben Brudvik Olsen
MSc in Big Data

The Objectives

News articles are one of the most influential sources upon which stock investors base their buying and selling decisions [1].

A study found that news articles have their strongest predictive power on stocks within 20 minutes after being published [2].

The purpose of this project is to see if the two theories described above translates to the cryptocurrency market.

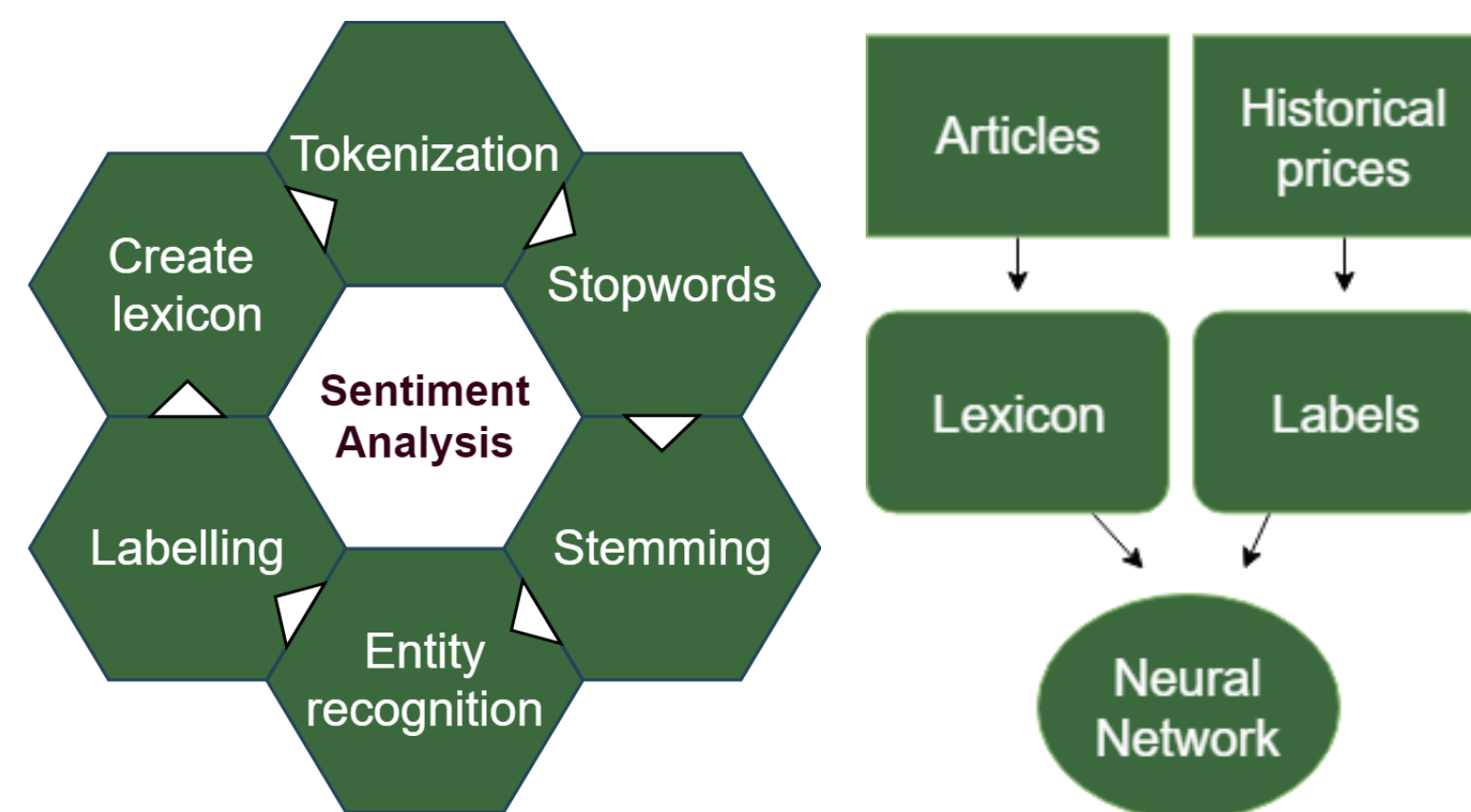


The Process

Using Bitcoin.com, a main news-source for cryptocurrency, nearly 2.000 articles were scraped.

The values for each given cryptocurrency at publishing time and 20 minutes after was needed. Increases in value were labelled «positive» and decreases were labelled «negative».

A lexicon was created from the articles and combined with the respective sentiment labels was fed into a deep neural network to learn how to classify unseen articles as positive or negative.



The Obstacles

Achieving 100 % accuracy in sentiment analysis on large datasets is currently impossible. Newspapers are not always first to provide the news (news often spread through forums in the cryptocurrency environment), and other newspapers might publish the article before the one used here.

The biggest obstacle in this project is that the cryptocurrency market is highly volatile. Overall, there are several reasons why an article might be labelled incorrectly.

The Sources

[1] <https://www.jstor.org/stable/2352710>

[2] <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.5103>