



A Systematic Approach to Analyzing Syndication of Viral News Stories

Samuel Teplov

Gaurav Deshpande

Alon Peer

Faculty Advisor : Dr. Nicolas Christin

PROBLEM STATEMENT

- Misinformation has become a normal occurrence in our interconnected society
- Media outlets further exacerbate this problem by blindly syndicating articles
- Tracing the propagation path of syndicated articles is a very tedious task

OBJECTIVE

- To create a system which automatically:
 - ⇒ Identifies related articles
 - ⇒ Clusters them together
 - ⇒ Performs analysis on the associated metadata.

CONTRIBUTIONS

- 🔍 Semi-automated search of related articles
- 🏗️ Classification of articles
- 📄 Identification of propagation path
- 📊 Analysis of various metadata features

FUTURE WORK

- Fully automated process
- Incorporate existing models on fake news and clickbait
- Make codebase open-source.

METHODOLOGY

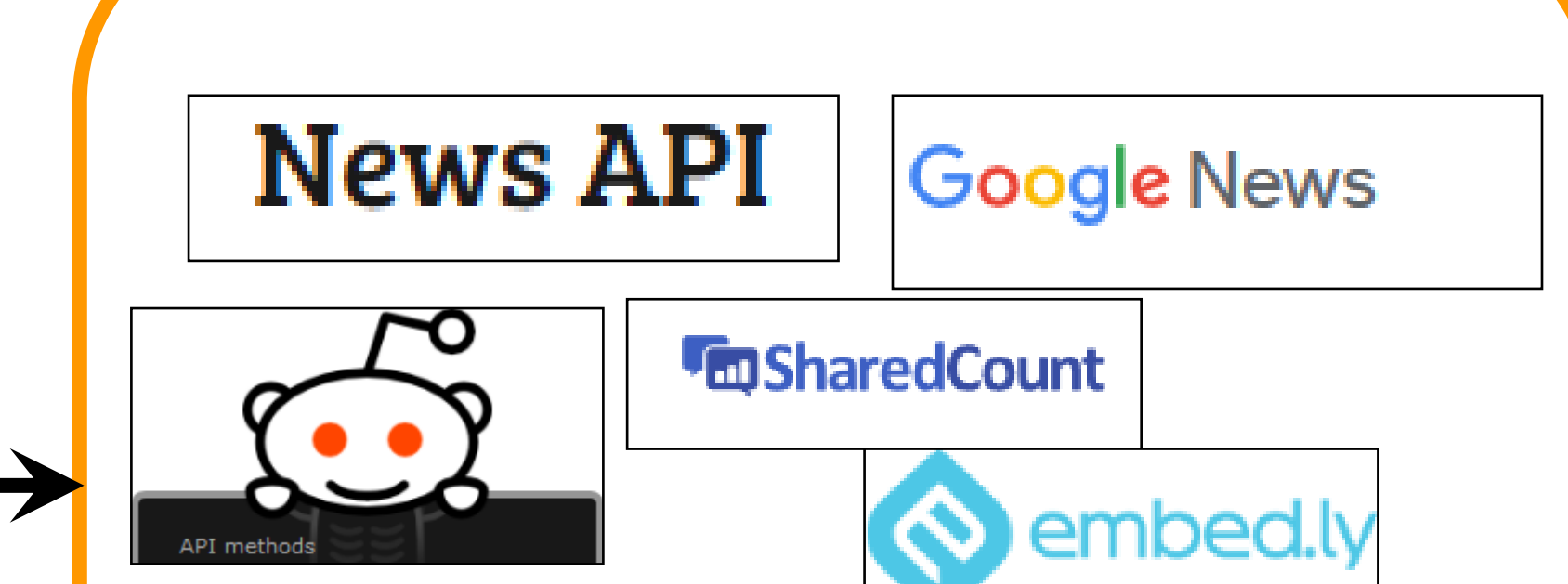
Stage 1 - Data Collection



Identification of Viral News

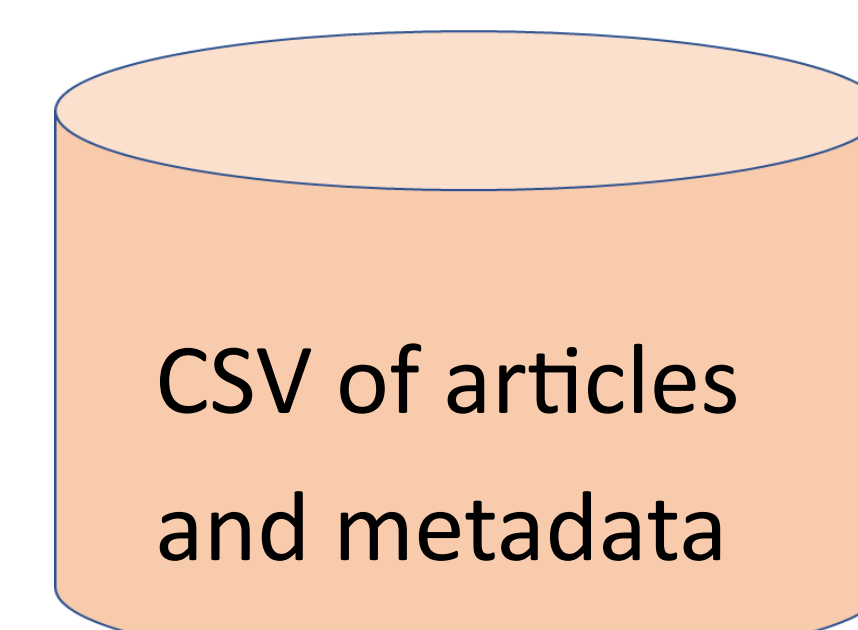


Gathering Similar Articles

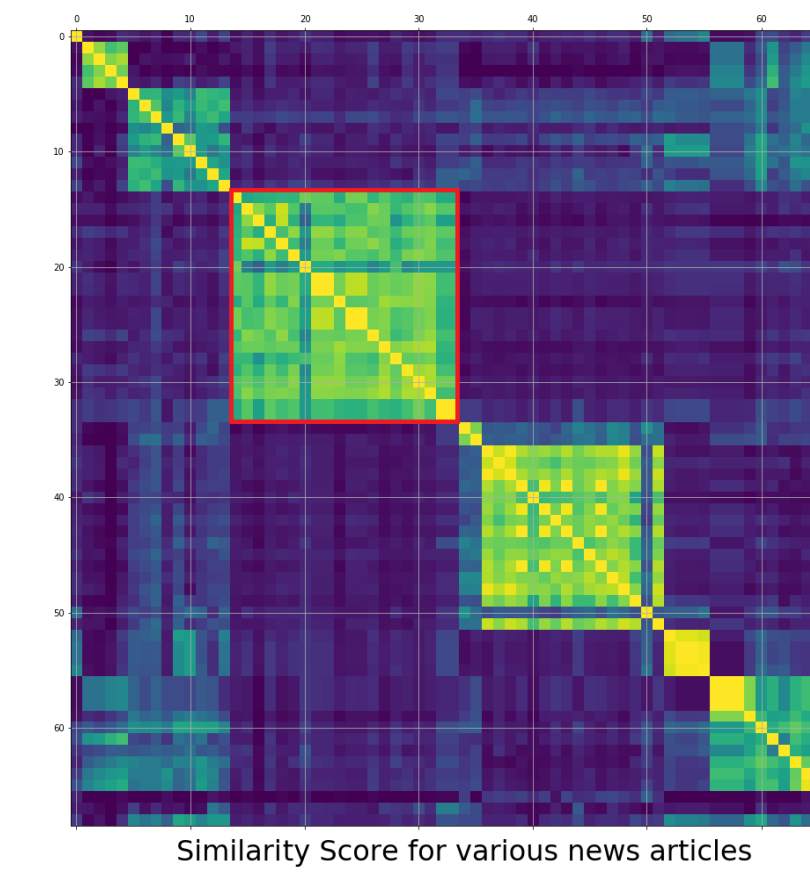


Web Scraping

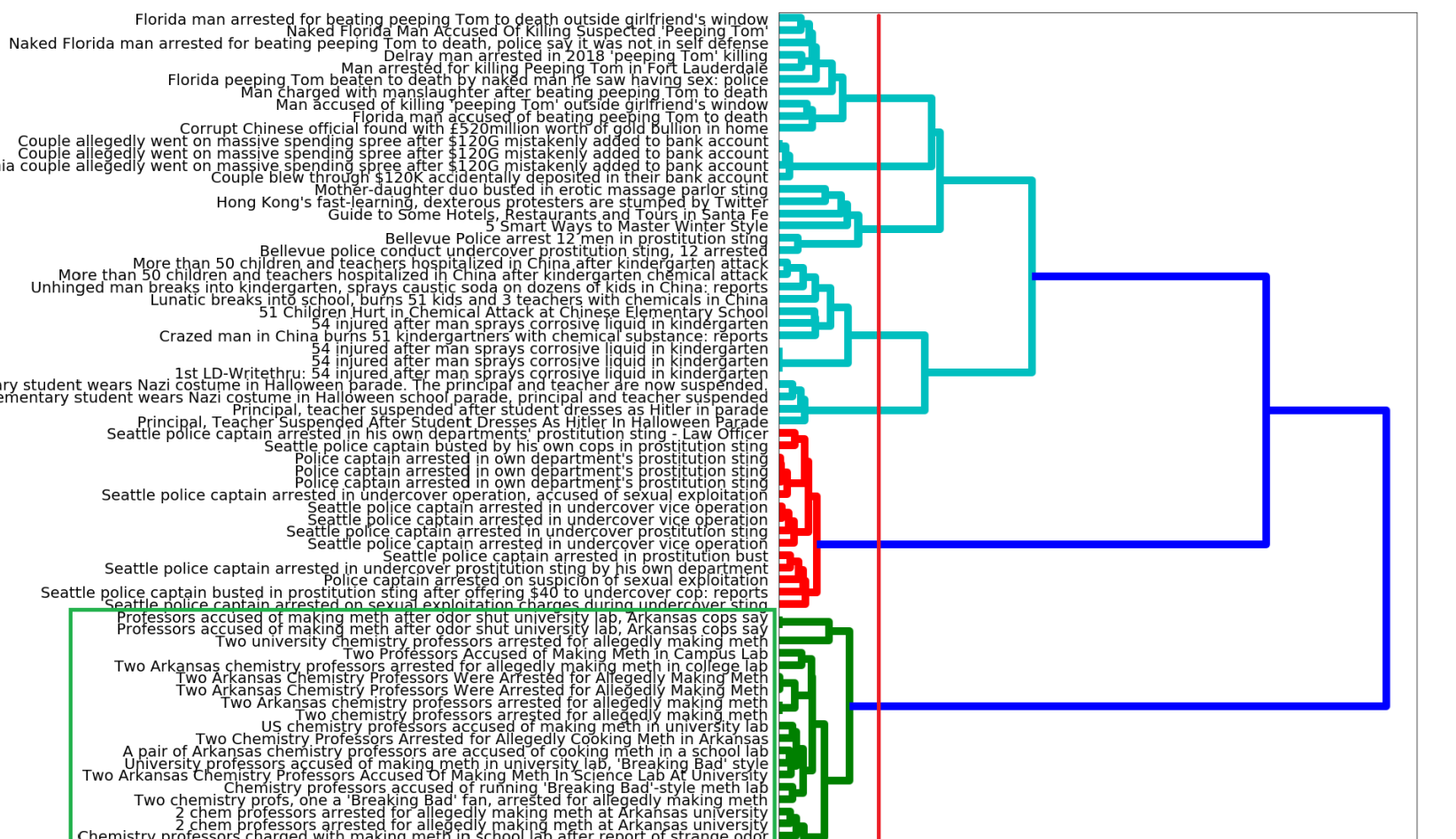
Stage 2 - Applying Machine Learning Techniques



Dataset Aggregation & Pre-Processing

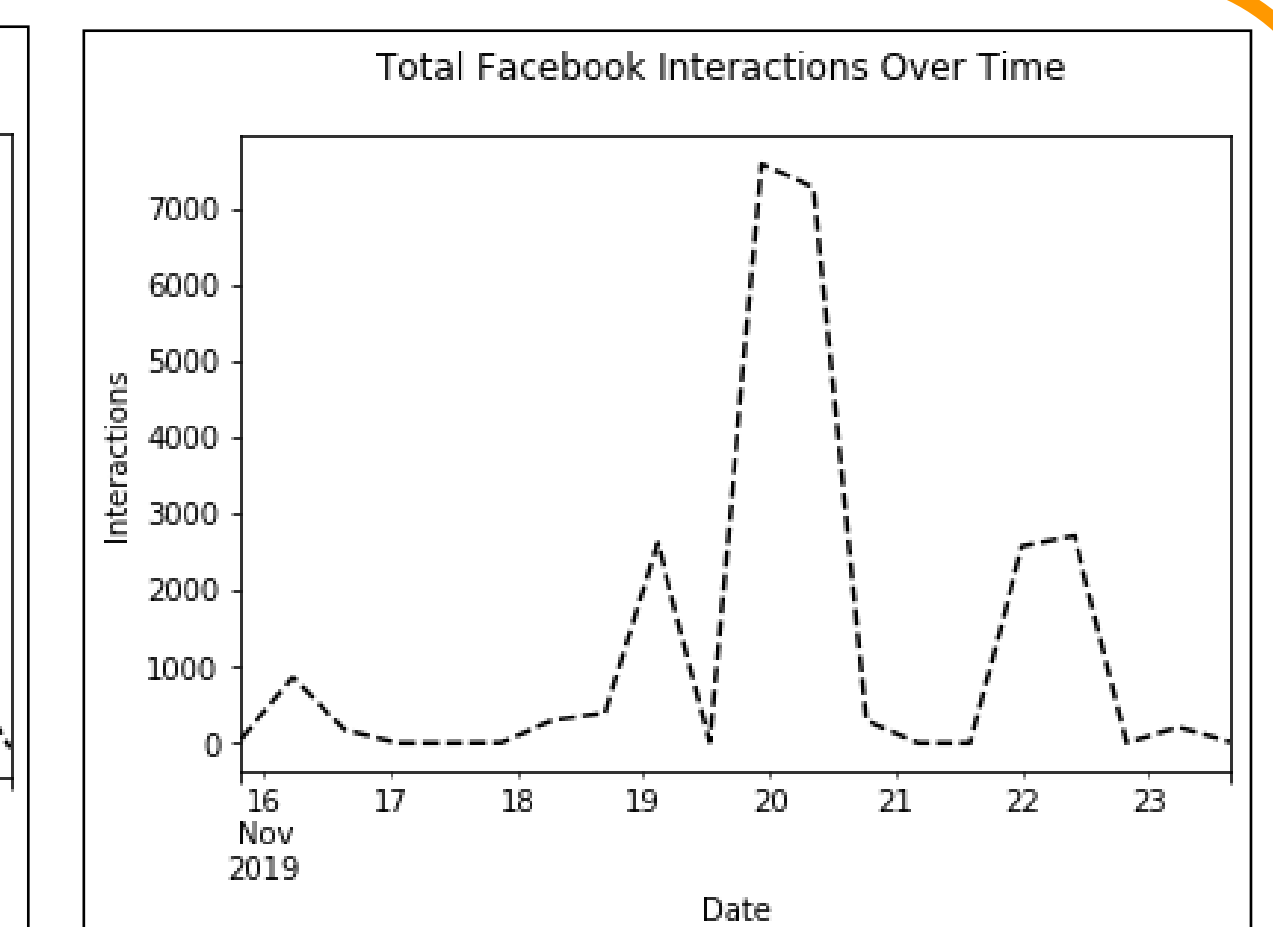
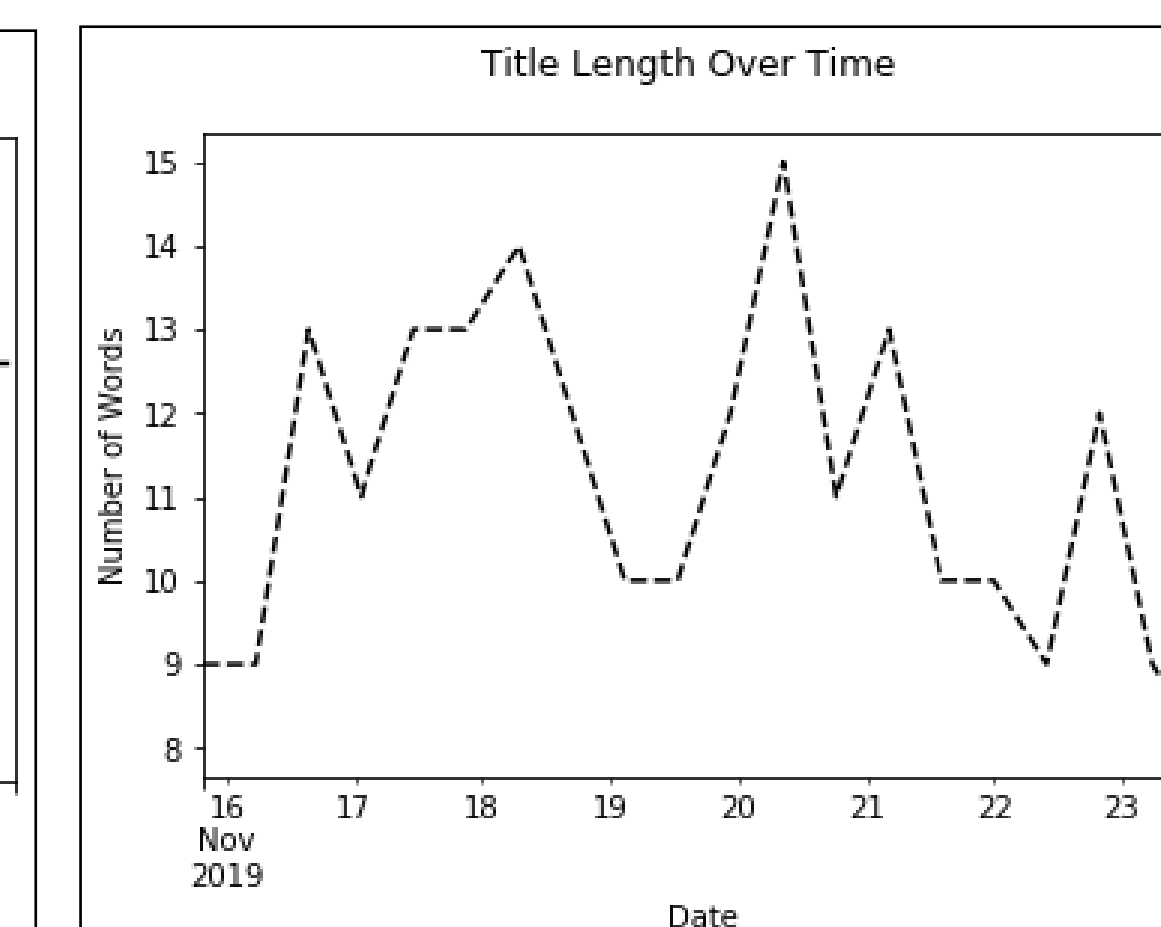
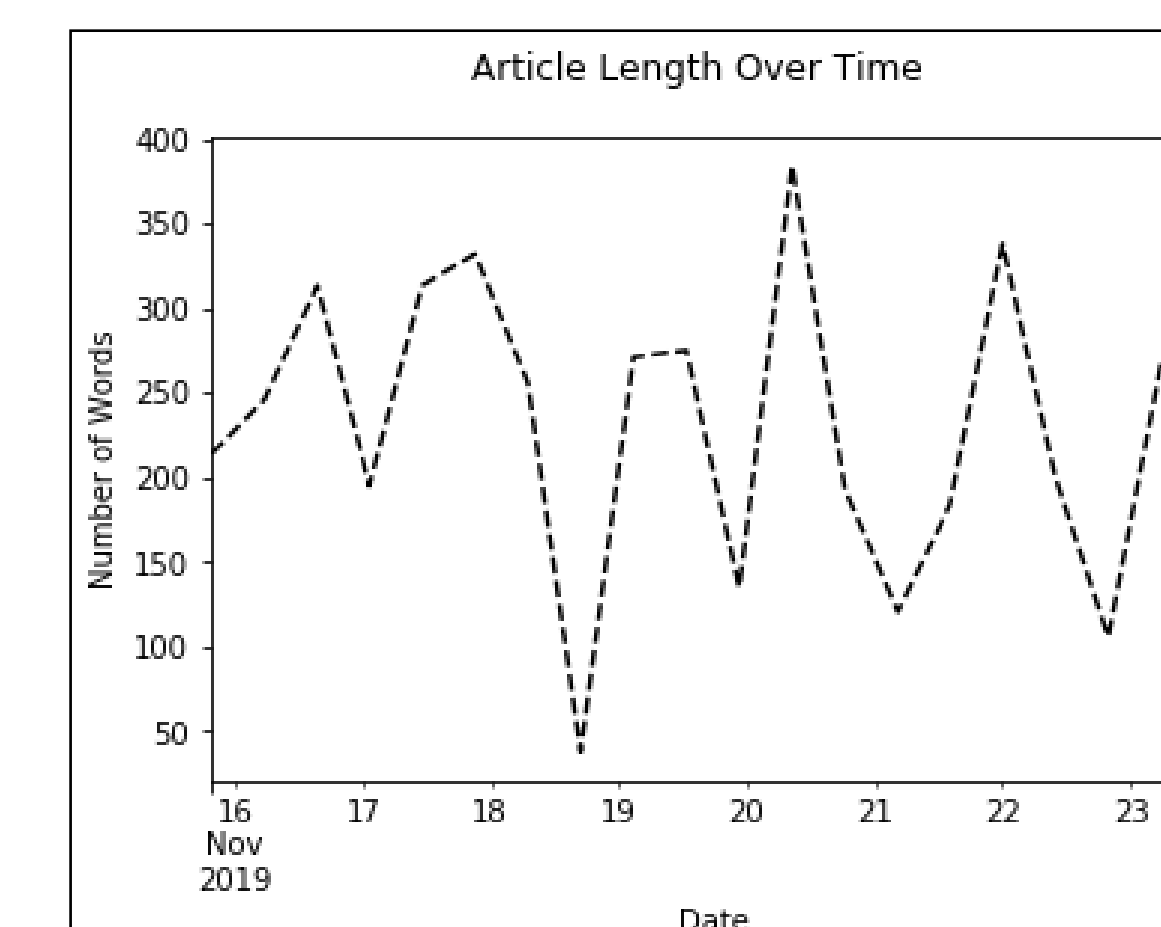
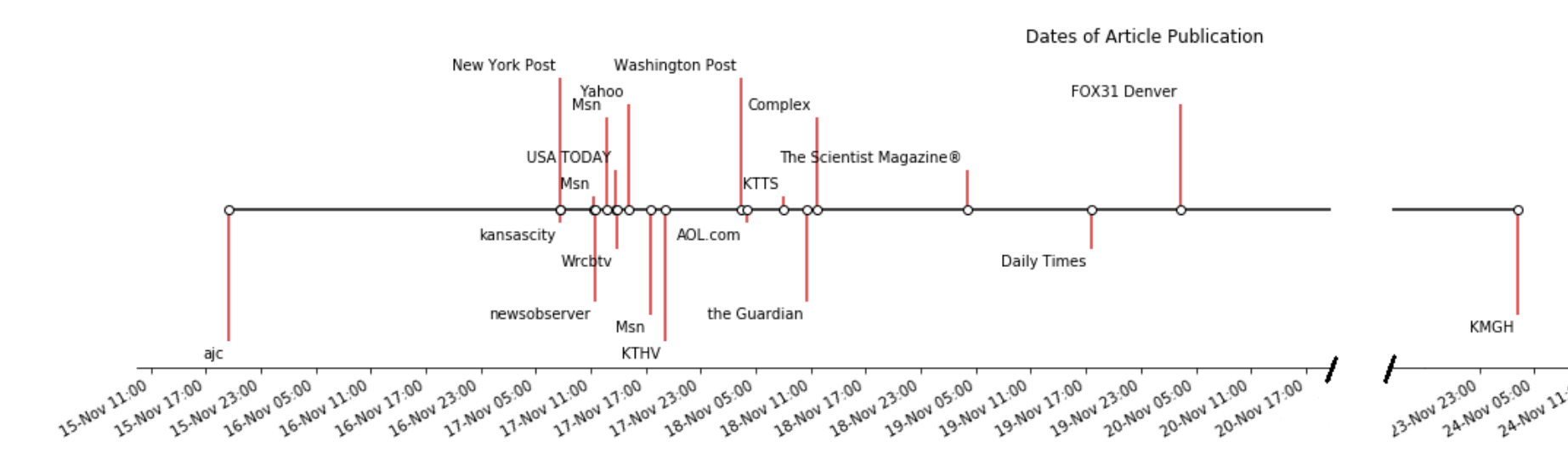


Similarity Score for various news articles



Clustering

Stage 3 - Data Analysis



Analyzing Propagation of Articles