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
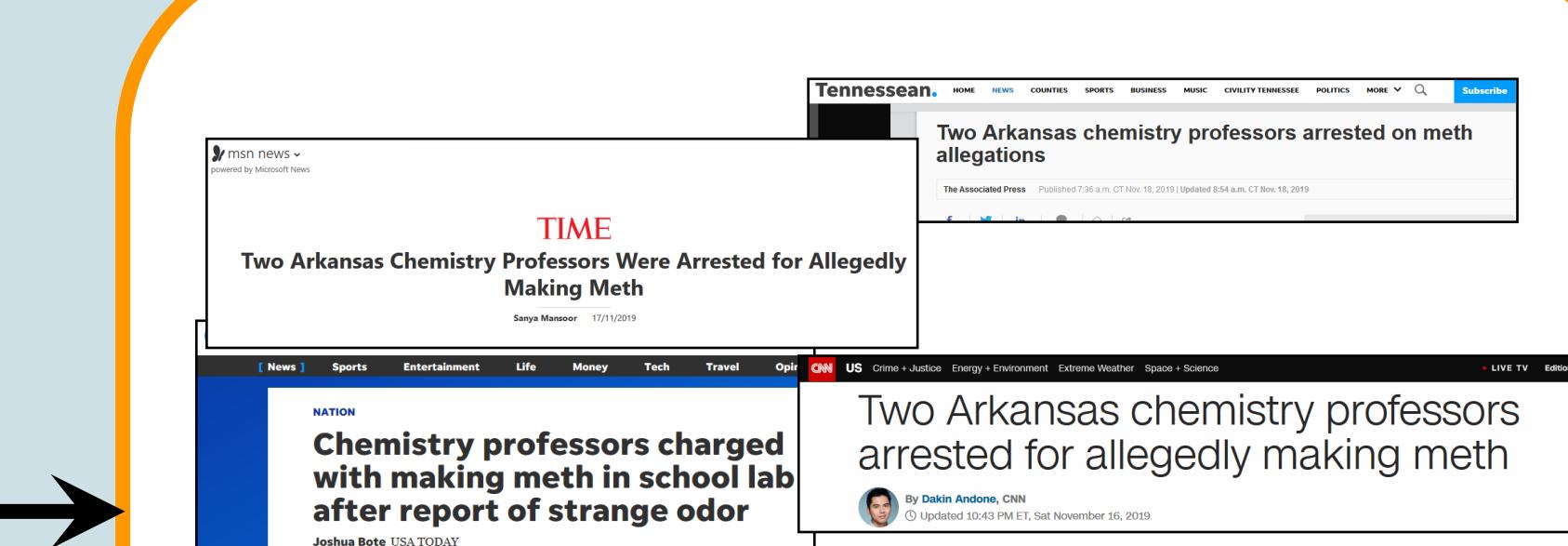
- Fully automate process

- Incorporate existing models on fake news detection

- Make code base open-source

METHODOLOGY
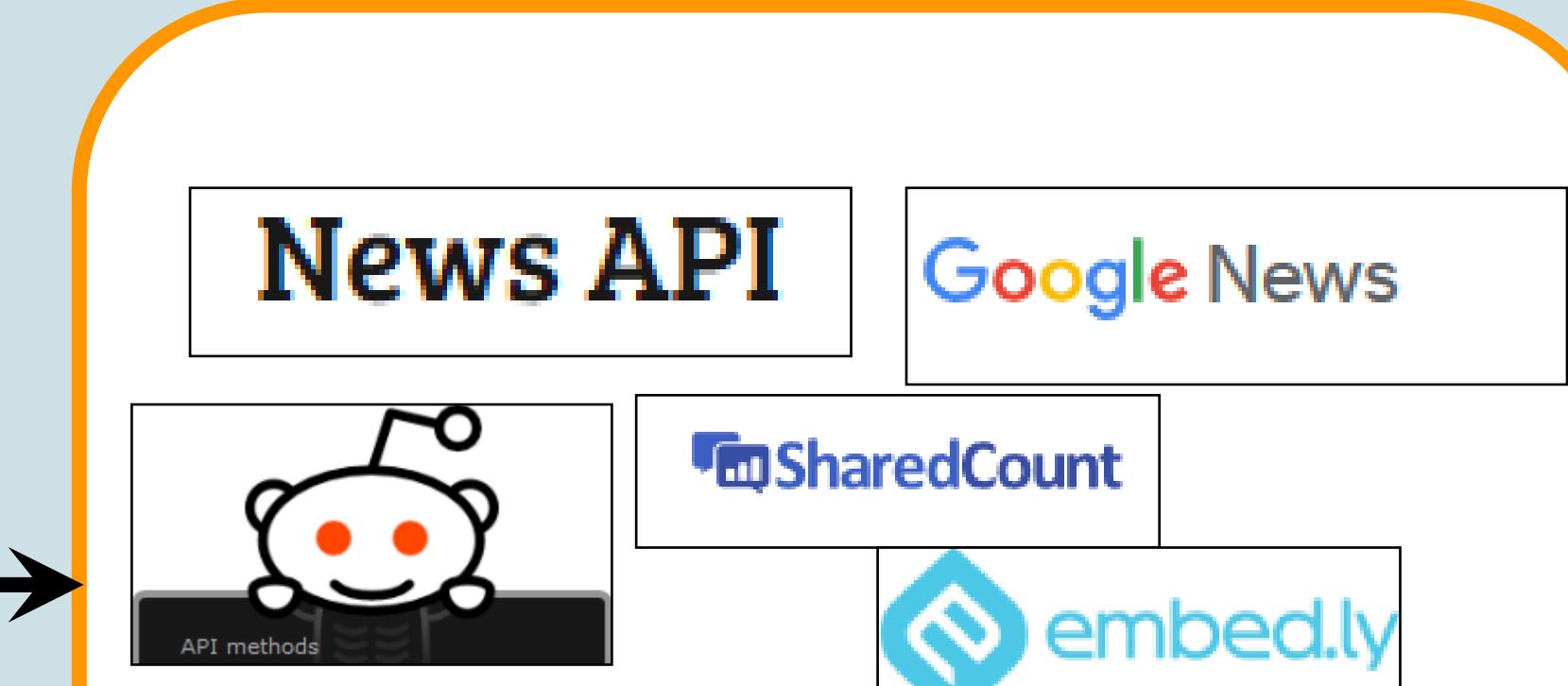
Stage 1 - Data Collection

Identification of Viral News   **Gathering Similar Articles**   **Web Scraping**
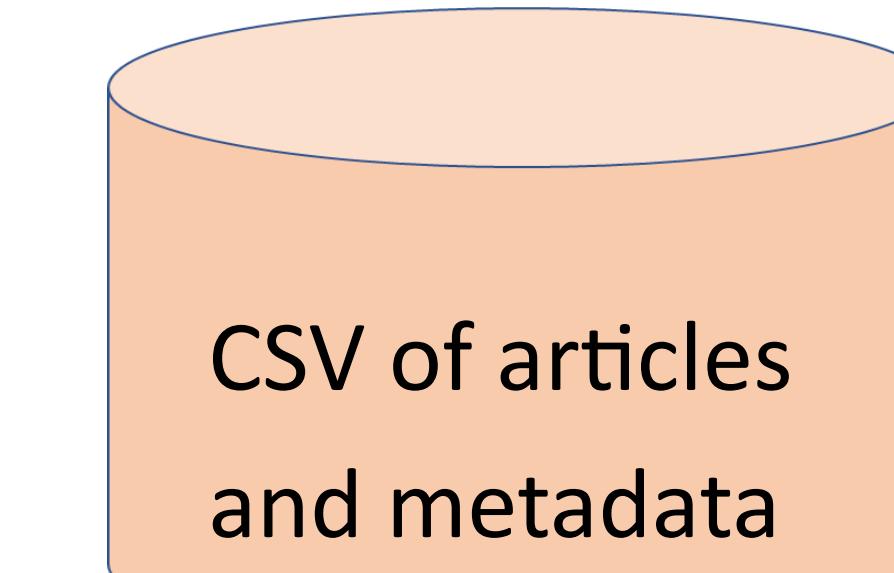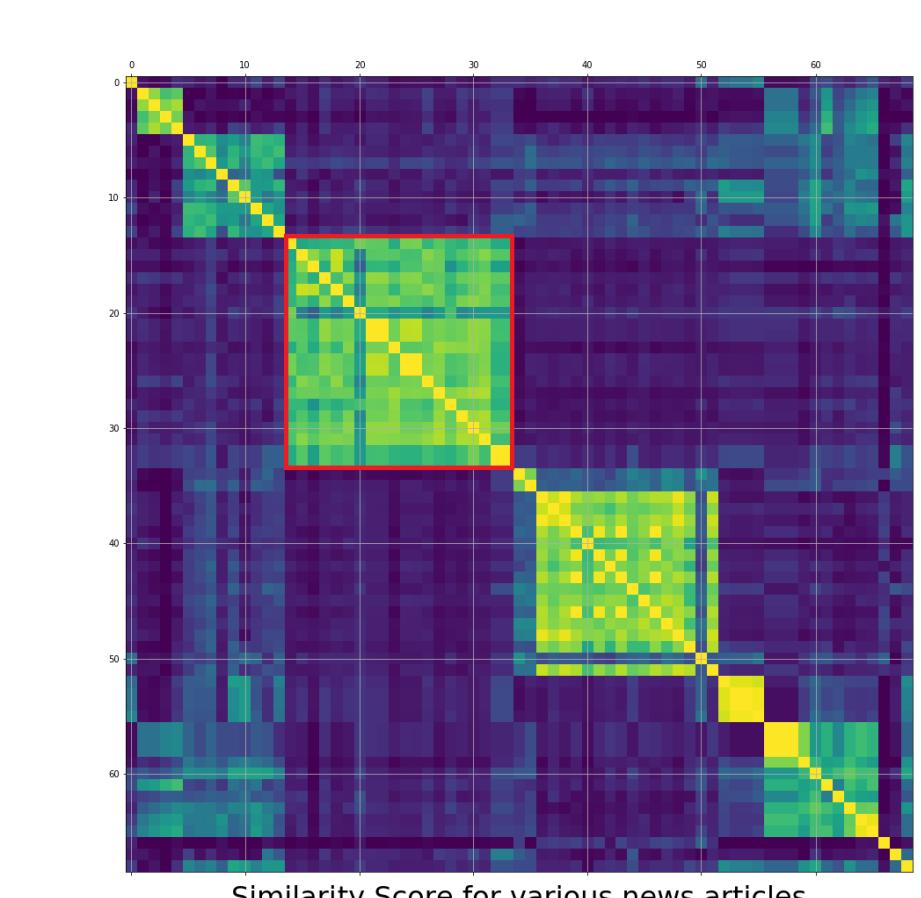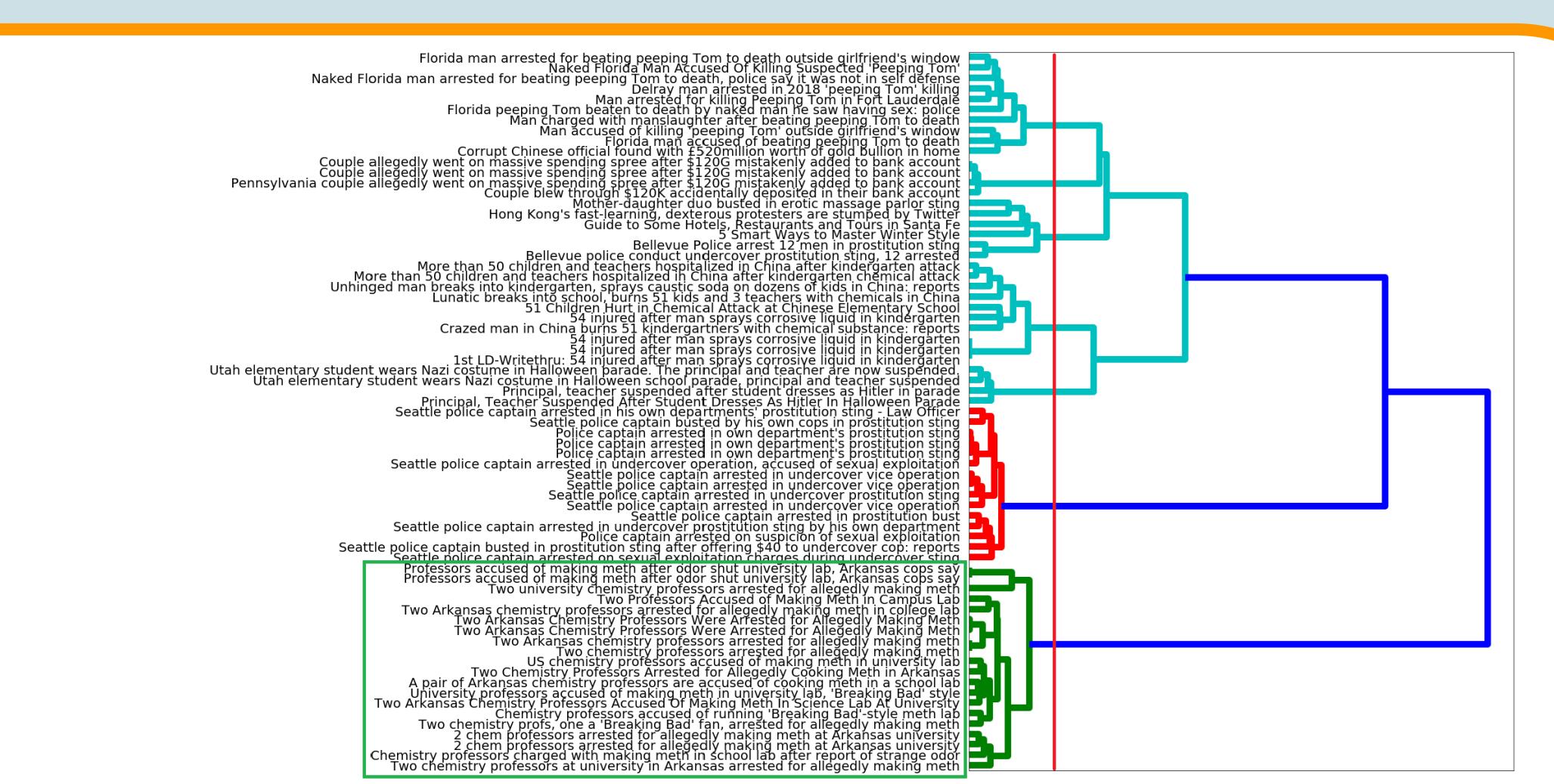
Stage 2 – Applying Machine Learning Techniques

CSV of articles and metadata

Dataset Aggregation & Pre-Processing
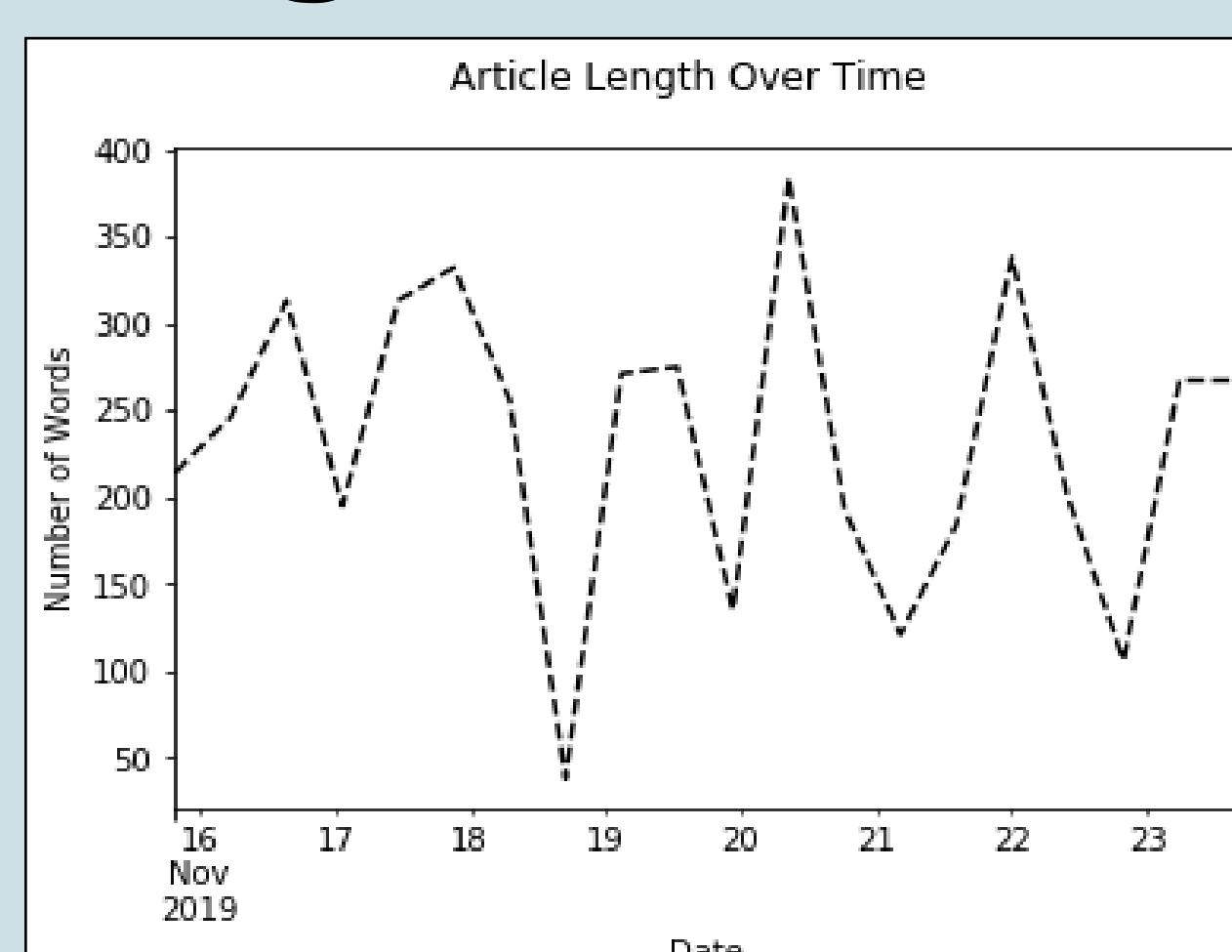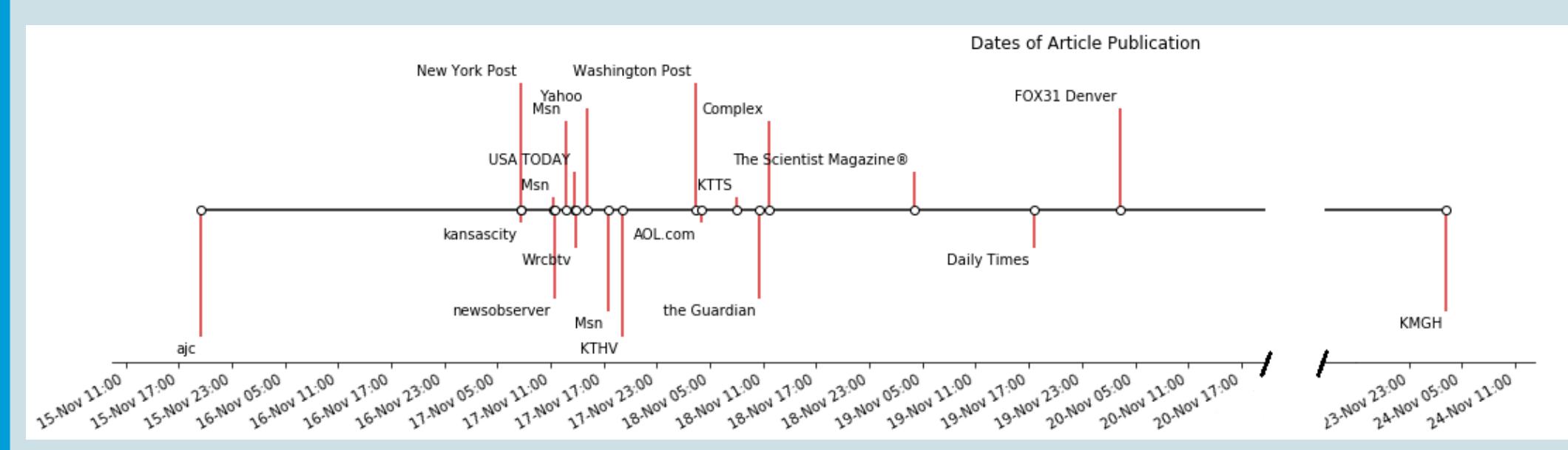
Similarity Score for various news articles

Clustering

Stage 3 – Data Analysis

Dates of Article Publication

Article Length Over Time

Title Length Over Time

Total Facebook Interactions Over Time

Analyzing Propagation of Articles