RISE RESEARCH INNOVATION SCHOLARSHIP ENTREPRENEURSHIP

Northeastern University

Analyzing News Articles From DIfferent Political Sources: Climate Change

Natalie Kwong, Abigail Valladolid, Sheryl Cheng; John Rachlin

Background

In today's polarized media, news articles from left, right, and center-leaning sources often present distinct narratives on the same topics, shaping public perception and societal discourse. This project aims to analyze these differences by examining popular common words, sentiment polarity, and vocabulary richness which could uncover the linguistic and emotional strategies employed by diverse news companies.

Methods

Our analysis utilizes a framework implemented by a class designed to process analyze, and visualize textual data from text files. It also has the ability to support custom parsers to handle specific preprocessing needs. The framework builds a meta-dictionary to store metrics for each text file of choice, including word counts, sentiment polarity, and vocabulary richness. Users can load text files or provide URLs which converts to text files using a utility we created. Although additional methods for visualizations can be included, the framework currently includes a sankey diagram, a bubble chart, and a sunburst chart.

Conclusion

The visualizations reveal distinct linguistic trends across left, right, and center-leaning articles surrounding climate change. More left leaning articles display a broader vocabulary of environmental themes whereas right leaning articles have more use of repetitive language of government policies. These findings reflect differences in media portrayal, which are further complicated by personal biases and how we identify our political standing. Though humans might interpret these patterns through the lens of polarizing perspectives, an NLP model offers a more objective analysis of analyzing linguistic patterns in written contexts. This contrast highlights the value of using NLP to uncover trends that might be skewed by human interpretation.

Findings

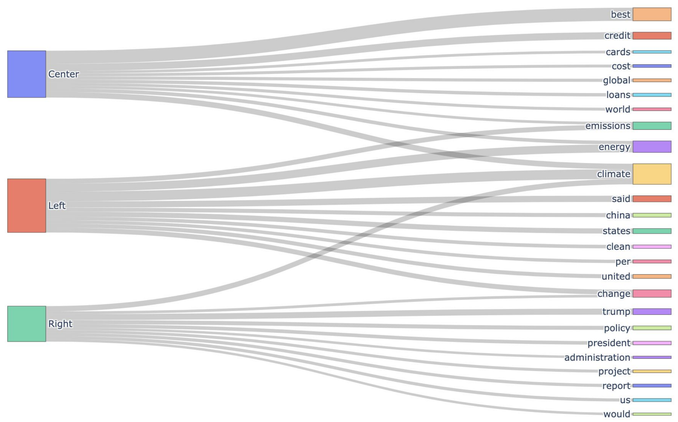

Figure 1: Word Count Sankey

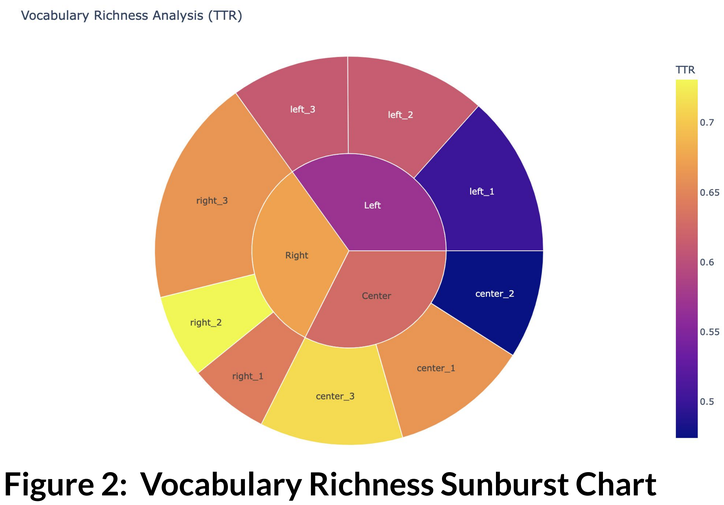

Figure 2: Vocabulary Richness Sunburst Chart

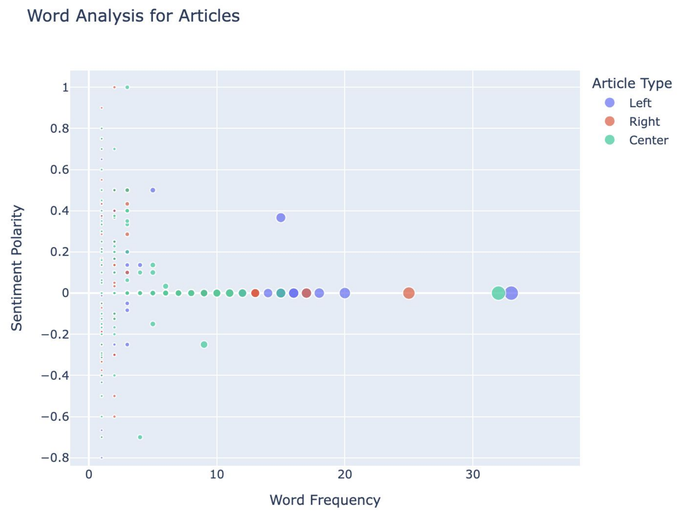

Figure 3: Sentiment Polarity Bubble Chart

Next Steps

For future analysis, we would analyze these common words within the context of their political leanings. For example, the same word can be used in a multitude of articles, but within different contexts to portray a distinct story or bias. One way to model this is by utilizing a word or token network with n-grams.

Descriptions

Figure 1: The Sankey maps the political leanings to the most common words and displays popular shared words between them.
Figure 2: The bubble chart displays the sentiment polarity of the most common words. The size of the bubble represents the frequency of that word and the color represents its political leaning.
Figure 3: The sunburst chart displays the vocabulary richness of each article within each political leaning.

Sources

https://www.nytimes.com/interactive/2024/11/19/climate/china-emissions-fossil-fuels-climate.html
https://www.cnn.com/2024/07/15/climate/polar-ice-melting-day-length-time/index.html
https://abcnews.go.com/US/chris-wright-wrong-about-clean-energy-transition-experts/story?id=115971534
https://www.bbc.com/news/articles/cwy9dwg7rpeo
https://www.forbes.com/sites/nilsrokke/2024/10/08/climate-change-action-cant-wait-for-perfect-solutions/
https://thehill.com/opinion/energy-environment/4993135-trump-climate-change-impact/
https://www.foxnews.com/politics/biden-admin-says-climate-change-poses-national-security-risks-sends-delegation-address-them
https://thedispatch.com/article/what-trump-opponents-dont-get-about-project-2025/
https://cbn.com/news/us/climate-emergency-these-horrific-weather-events-long-ago-tell-true-story