

Sentiment Analysis & Topic Modeling on Offshore Wind Energy in New Jersey

Isabele Bittencourt, Dr. Aparna Varde (Mentor)

Clean Energy & Sustainability Analytics Center, Montclair State University

INTRO

The plans for offshore wind energy by the shore of New Jersey still brings controversy between residents due to the wind farms' impact on wildlife, coastline, and the people's view from the beaches. In this context, we investigate public opinion on social media and discover main keywords and major topics being spread by news media on offshore wind energy in NJ.

METHODS

Apify and BeautifulSoup were used for data collection and creation of the datasets. NLTK, a python NLP library, was used for preprocessing & cleaning the data.

Sentiment analysis is a series of methods, techniques and tools used to detect and extract subjective information, such as opinion and attitudes, from language. Three models were used to perform sentiment analysis on social media data: TextBlob, VADER & SentiWordNet.

Topic modeling is a statistical method for discovering themes and main keywords in a set of documents. GENSIM is a python library for topic modeling. GENSIM's LDA model was used for topic modeling on news media data.

RESULTS

Although the overall opinion of the population in the sentiment analysis approach is somewhat positive, the neutral and negative counts are not far behind due to concerns over the impacts of offshore wind energy on wildlife and beaches, budget money etc, based on the most frequent key words used by users in the data.

The topics extracted from the LDA model show similar insights when compared to the sentiment analysis performed on social media. Furthermore, the keywords help us understand how the media has been portraying offshore wind energy in New Jersey to the masses, such as whale, deaths, Orsted, climate, opposition etc. It can be seen on the most frequent words by news media that Fox News and CBS have a much larger count of mentions about whale deaths than CNN, The Guardian and NBC, in which the most frequent word is power.

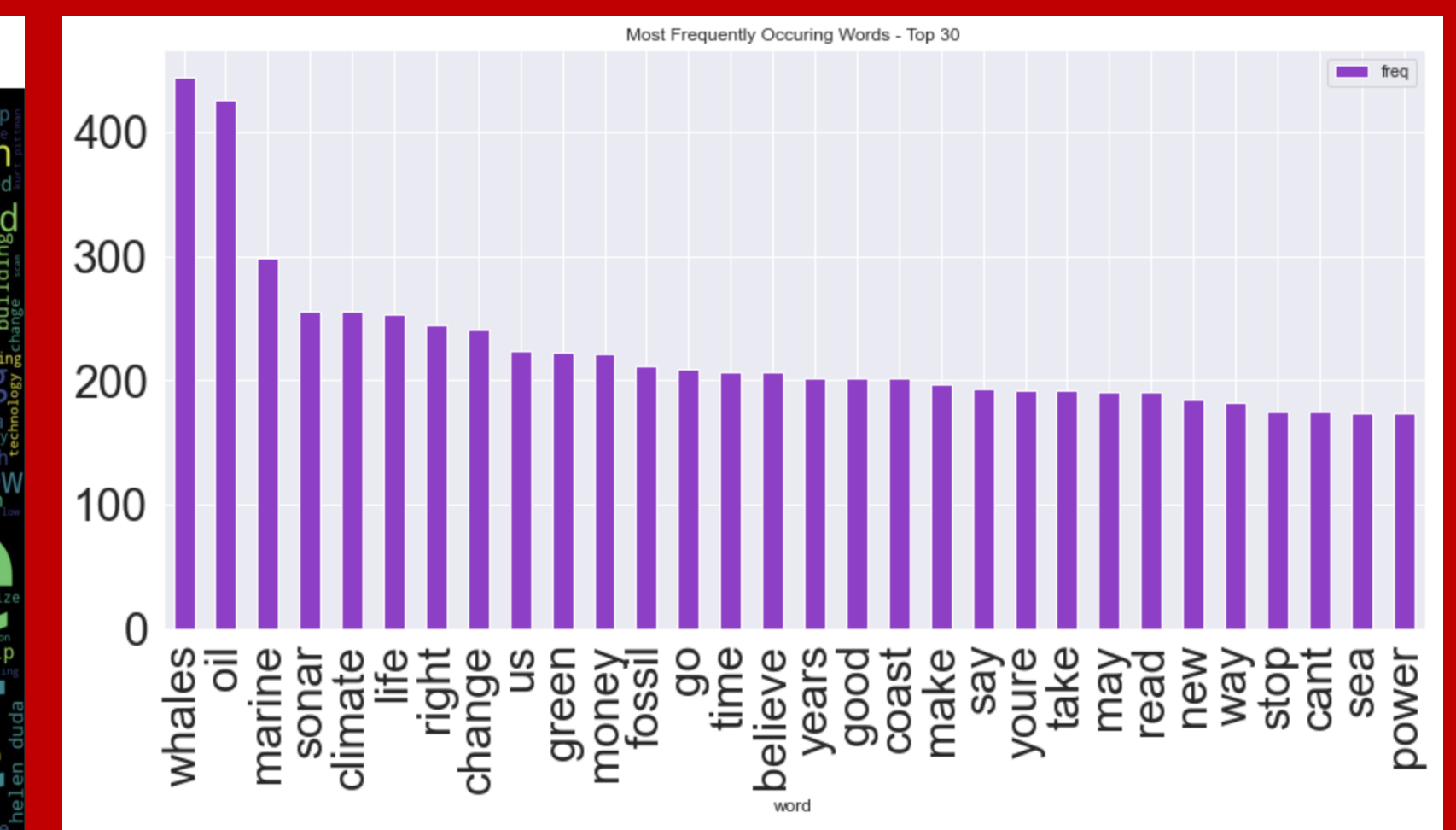
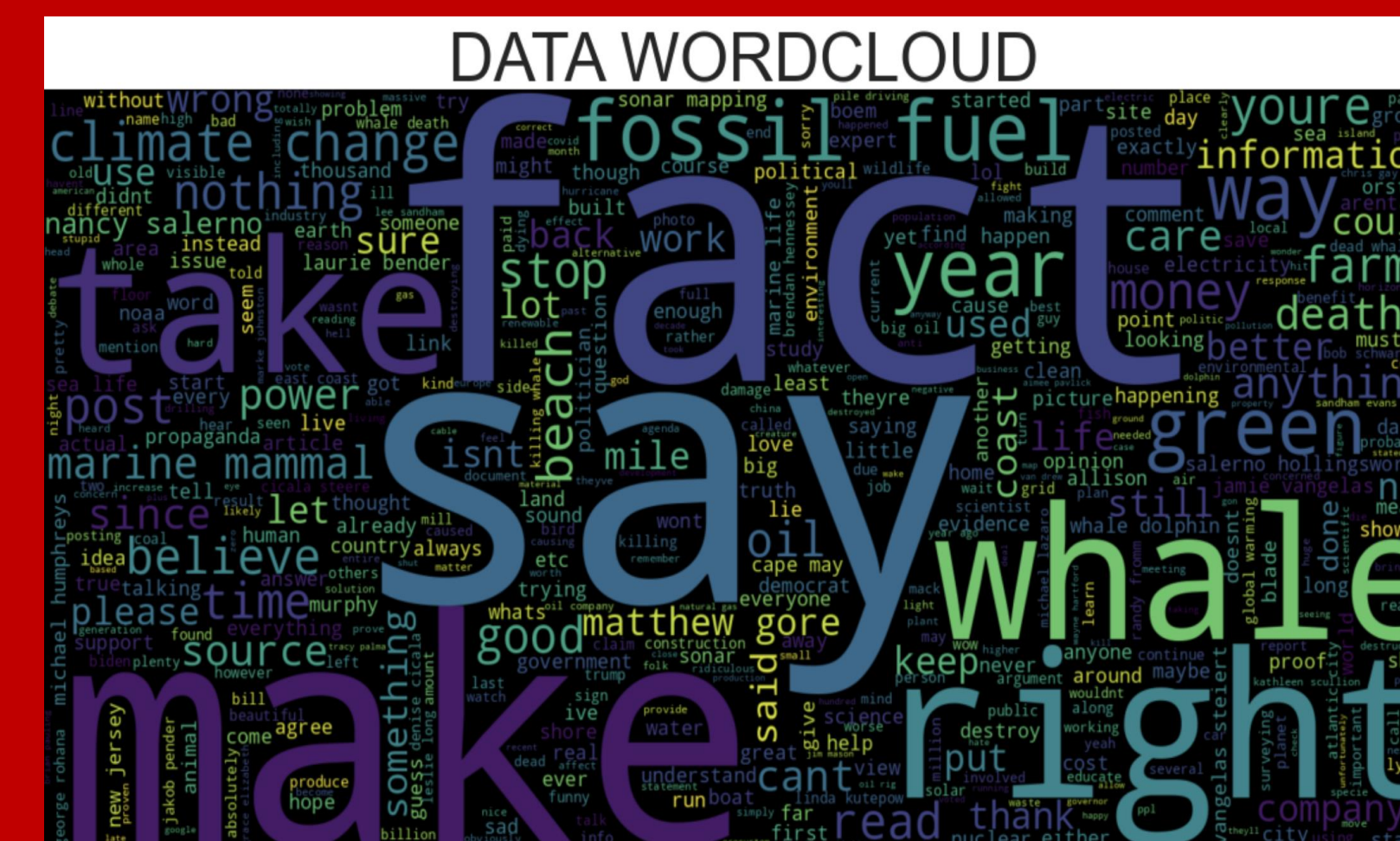
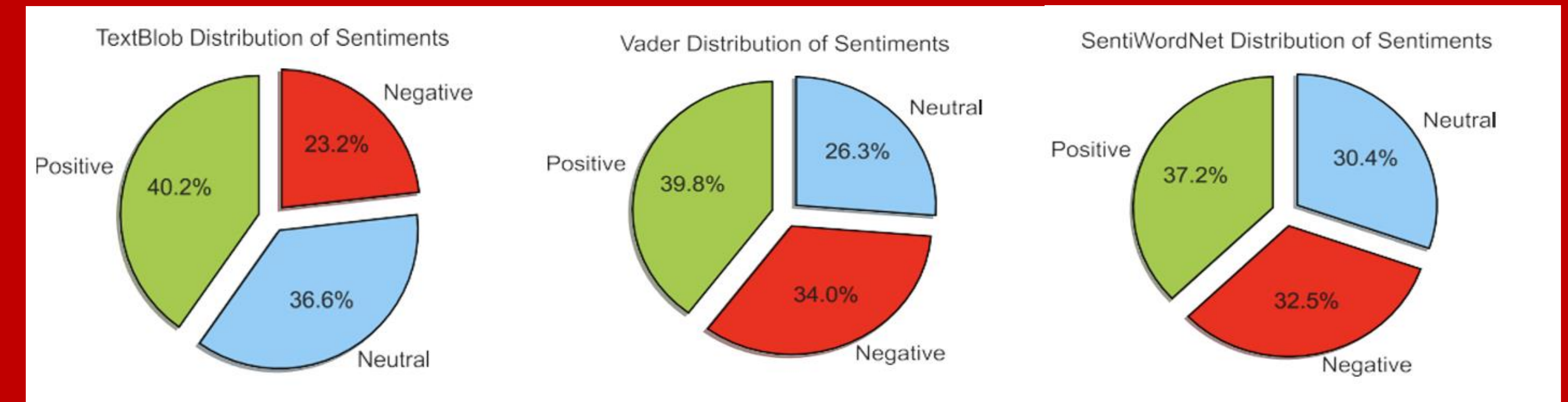
DISCUSSION

The results presented here can be used for future decisions by government leaders and interested companies on the topic, i.e. wind farms' location, projects budget, environmental impacts etc.

Tailoring policies to public opinion and identifying key concerns are an important part of the process of investing in the build of wind farms by the coast of New Jersey.

Performing similar analyses on social media posts in other regions to gain deeper insights on the reception of the policies by the common masses.

Sentiment Analysis on Facebook Data Visualization



Topic Modeling on News Media Data Visualization

