



UNIVERSITY
OF WARSAW

Sentiment and Topic Analysis of Donald Trump's 2024 Campaign Speeches

Analyzing Political Communication Strategies

Porimol Chandro

Ho Thi Hoang Nhu

Major in Data Science and Business Analytics

Faculty of Economic Sciences, University of Warsaw



Project Objectives



- Analyze sentiment in Trump's speeches.
- Track sentiment trends over time.
- Categorize speeches into key campaign topics.
- Utilize advanced sentiment analysis techniques.



Dataset Collection

- **Data Sources:**
 - Election campaign speech delivered by Donald Trump in Kaggle.
- **Selection Criteria:**
 - Relevance to the 2024 campaign.
- **Challenges:**
 - Ensuring data completeness and accuracy.





Data Preprocessing

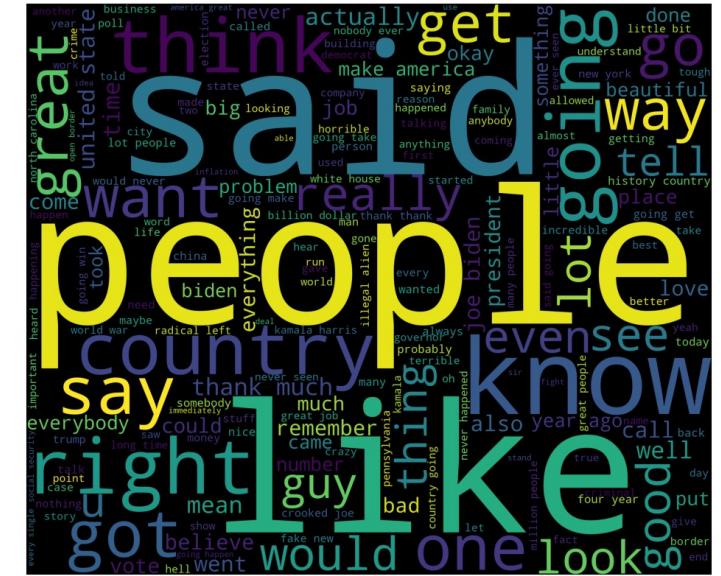
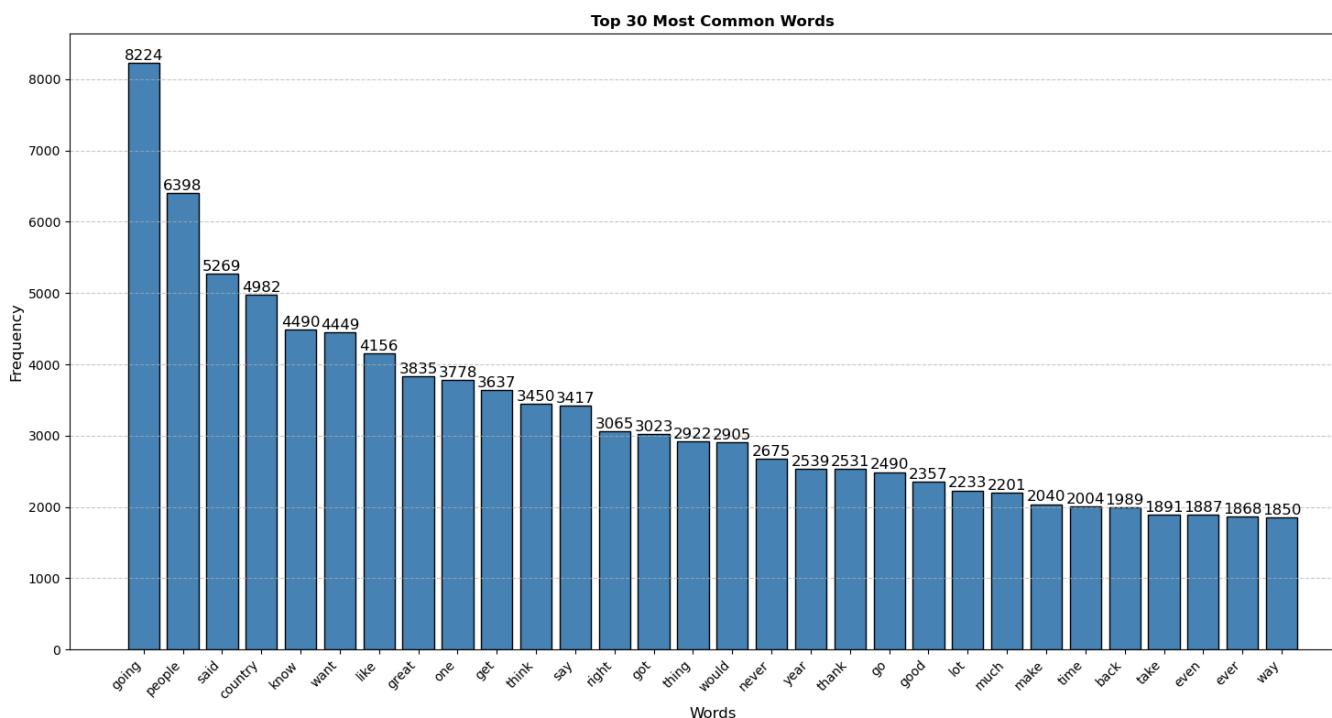
Designed to extract essential metadata and content from file names and speech transcripts., to ensure a structured and efficient approach to organizing and analyzing the dataset.

- **Text Normalization:** Convert all text to lowercase, remove punctuation.
- **Tokenization:** Split text into words or phrases.
- **Lemmatization:** Convert words to their base or root form.
- **Stop Words Removal:** Eliminate common words that do not add significant meaning (e.g., "and", "of").

	date	location	speech	file_path	cleaned_speech
0	2022-11-07	Pa For Dr Oz	\nYou're going to elect Dr. Mehmet Oz to the U...	data/text/trump-rally-in-pa-for-dr-oz-11-05-22...	going elect mehmet oz united state senate goin...
1	2022-11-16	Unknown	\nWell, thank you very much, and on behalf of ...	data/text/former-president-trump-announces-202...	well thank much behalf melania entire family w...
2	2023-02-01	New York Attorney Generals Fraud Investigation	\nI do.\n\nGood morning.\n\nSure.\n\nYes...	data/text/excerpts-of-donald-trumps-deposition...	good morning sure yes yes sure thanks think kn...
3	2023-03-06	Cpac	\nWell, thank you very much and I'm thrilled t...	data/text/trump-speaks-at-cpac-2023-transcript...	well thank much thrilled back cpac thousand gr...
4	2023-03-14	Davenport Iowa	\nOh, thank you very much. Thank you very much...	data/text/donald-trump-speech-in-davenport-iow...	oh thank much thank much thank great governor ...

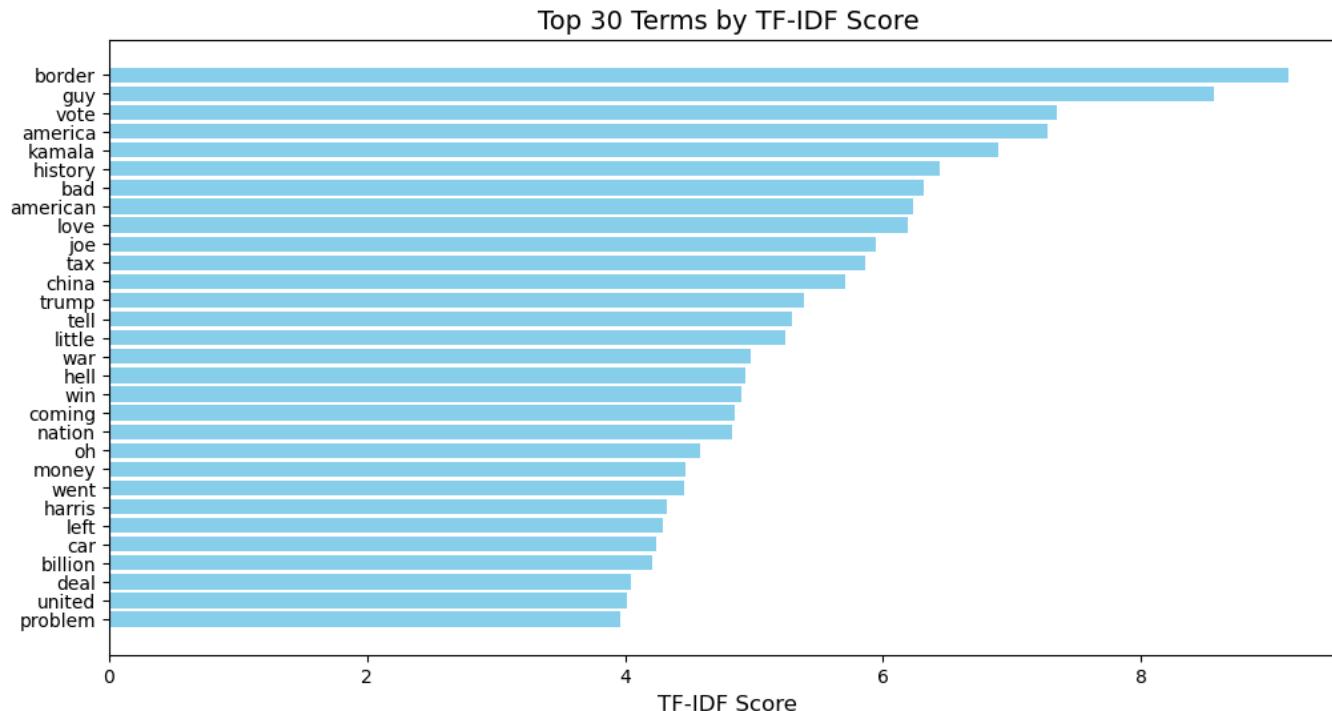
Exploratory Data Analysis (EDA)

- "people," "know," "country," and "great" show a focus on connecting with the public and addressing national issues.
- Positive terms ("thank," "good") and negative ones ("problem," "bad") reflect a mix of hope and critique.



"going," "people," and "country" stand out prominently in both visuals, reinforcing their significance in the campaign's messaging.

TF-IDF Vectorization



- **Immigration & Economy:** Terms like "border", "tax", and "money" reflect a focus on immigration and economic issues.
- **Patriotism:** Words like "America", "nation", and "united" emphasize national pride.
- **Opposition:** Names like "Joe", "Kamala", and "Harris" signal a focus on political rivals.
- **Emotional Tone:** Terms like "bad", "war", and "hell" suggest charged rhetoric to energize supporters.

Latent Dirichlet Allocation (LDA)



- The Dirichlet Allocation (LDA) model helps analyze latent topics in a text corpus by grouping related keywords into topics.
- Each topic represents a thematic analysis of the data and can help us better understand Trump's concerns during his campaign.

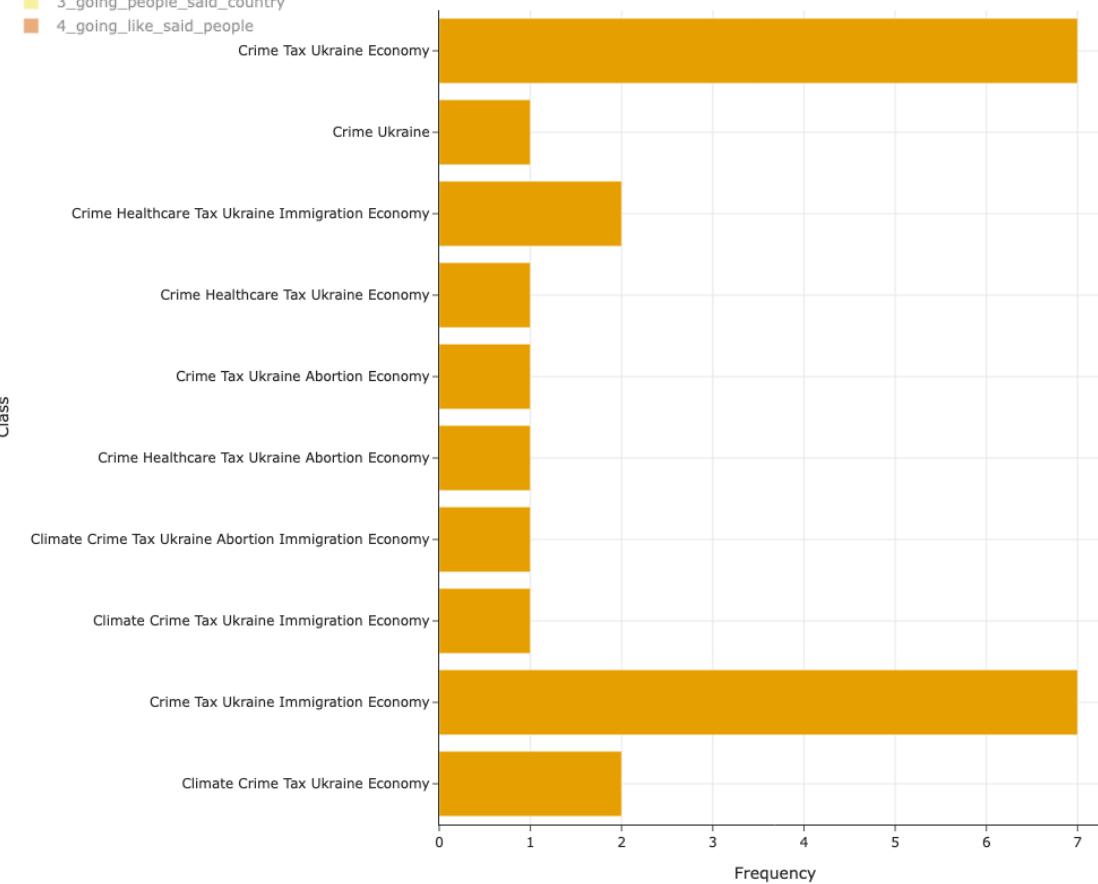
BERT for topic analysis



Global Topic Representation

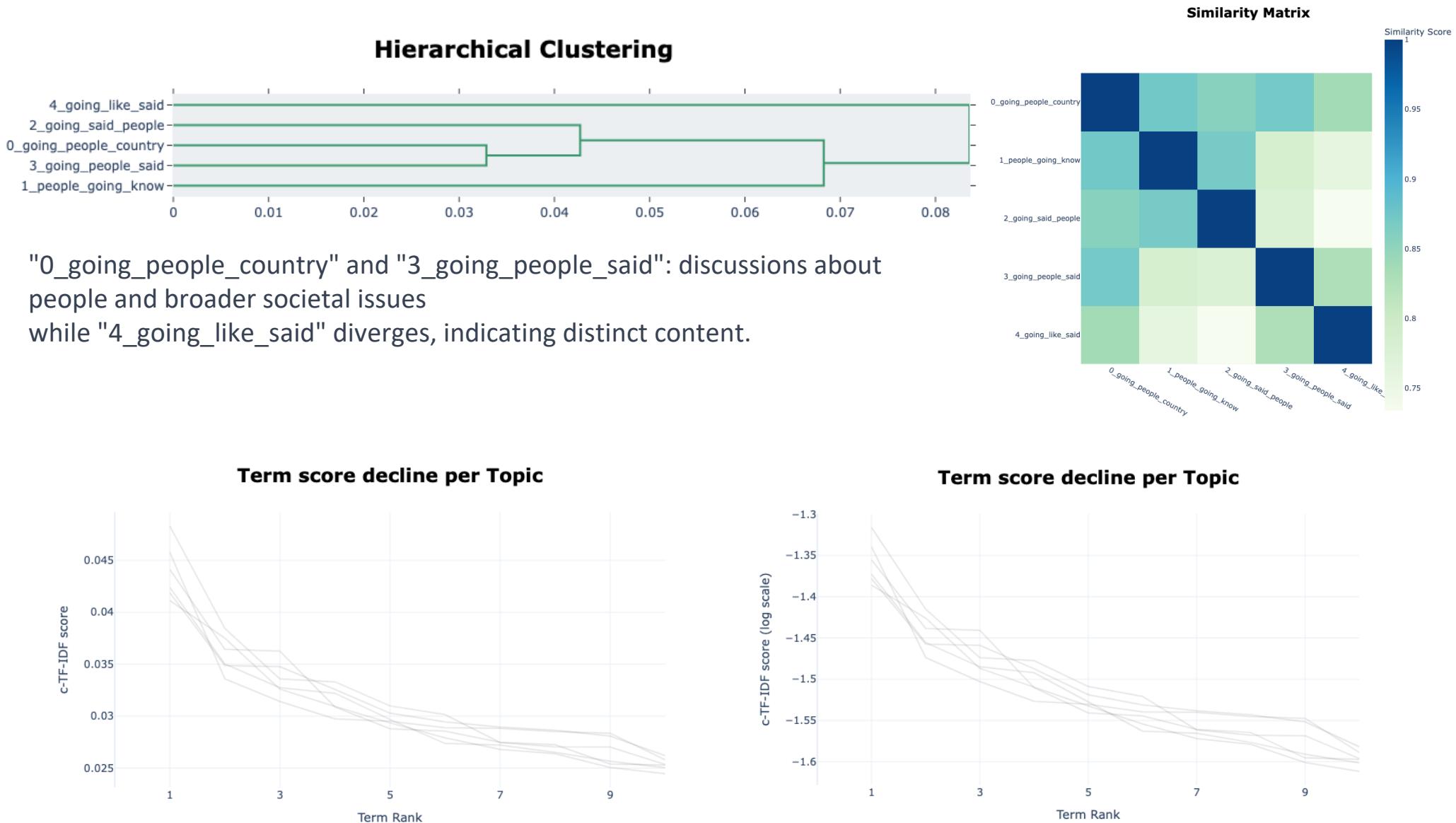
0_going_people_country_know
1_people_going_know_said
2_going_said_people_country
3_going_people_said_country
4_going_like_said_people

Topics per Class



- focus on **crime, taxes, Ukraine, and economic challenges** as recurring and interconnected themes.

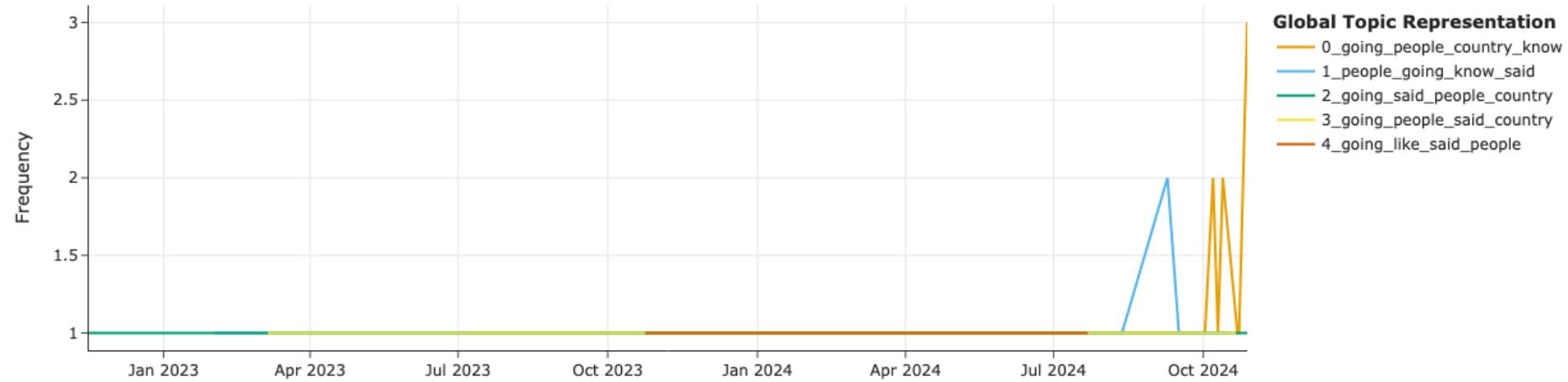
BERT for topic analysis



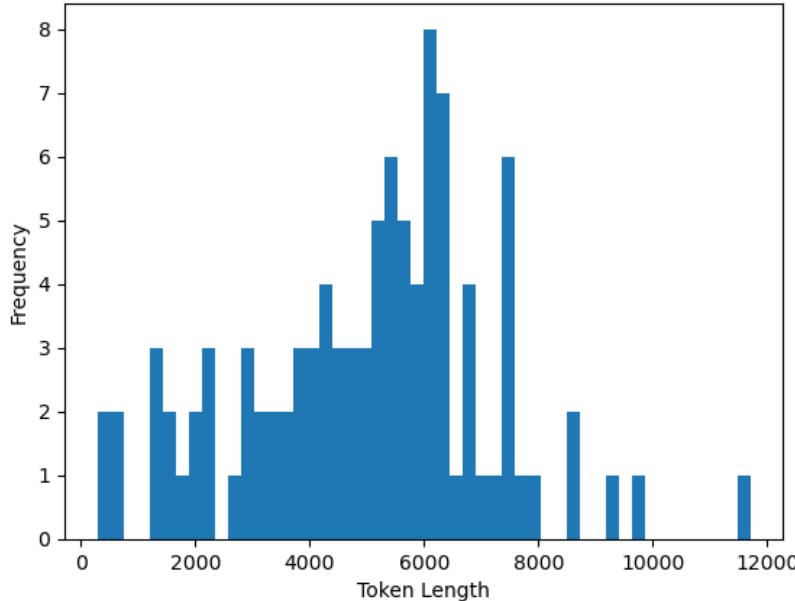


BERT for topic analysis

Topics over Time



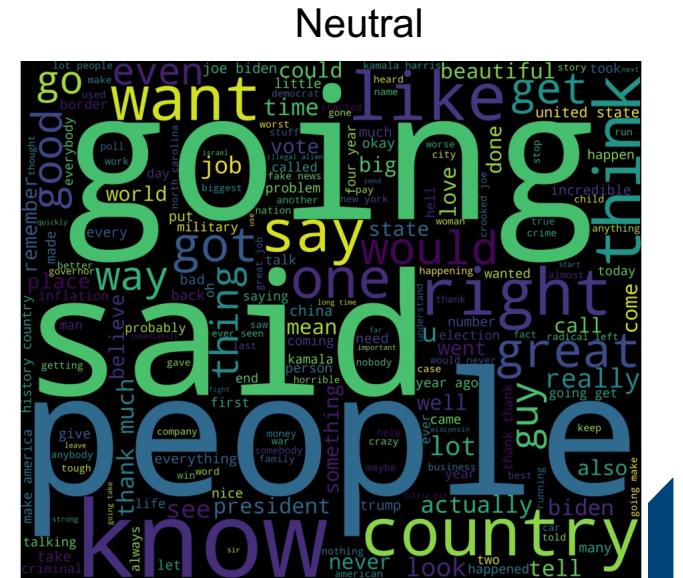
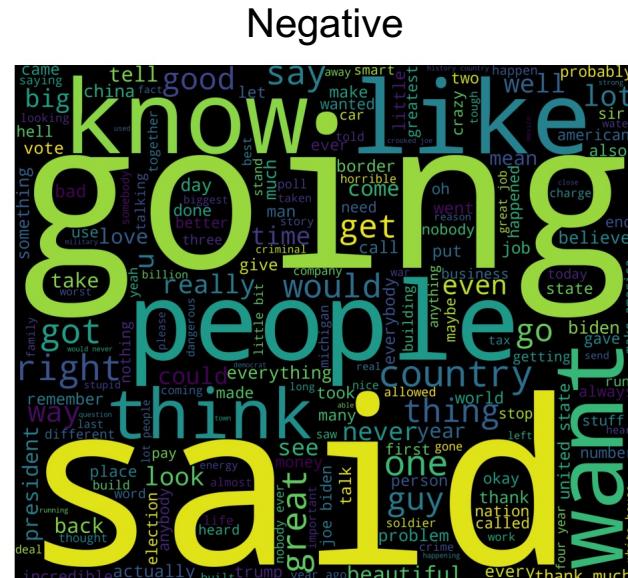
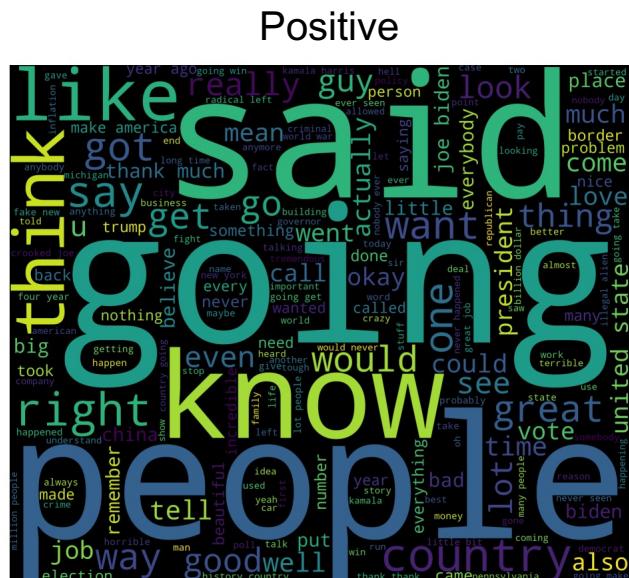
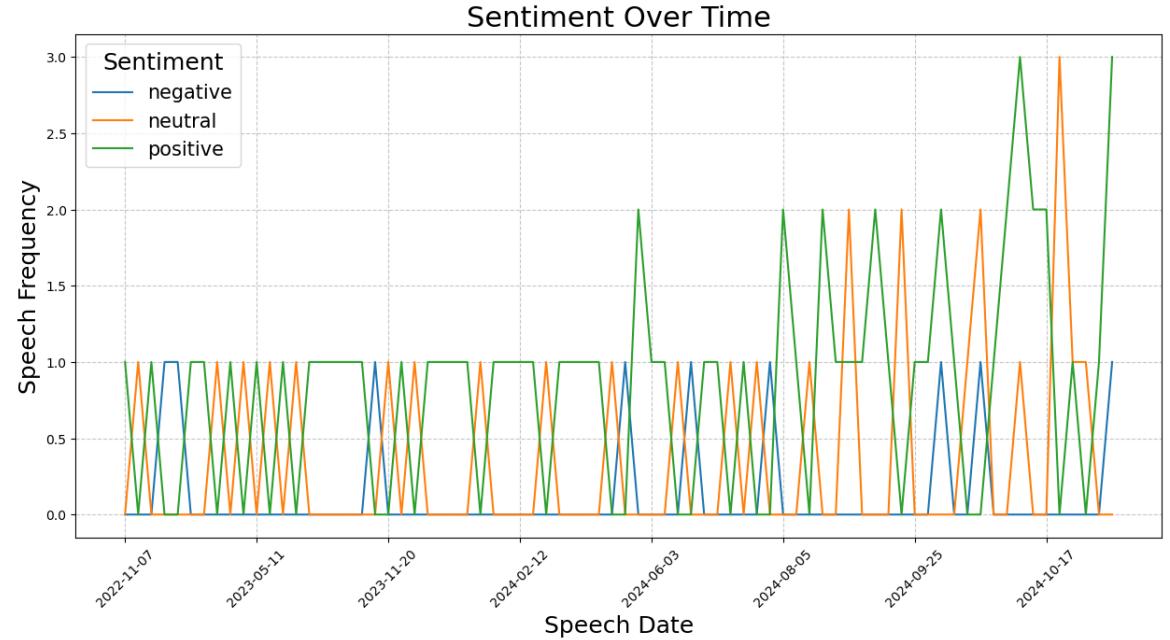
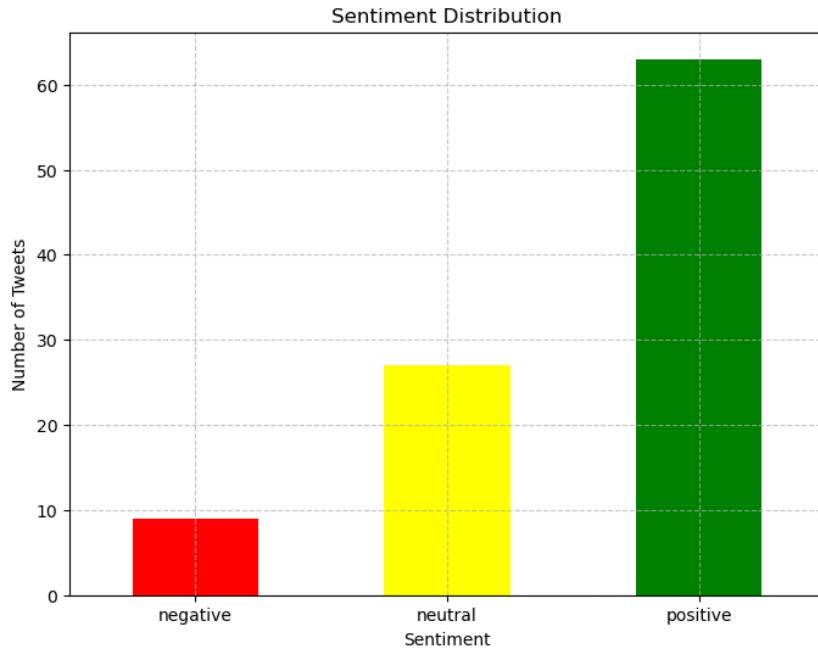
Token Length Distribution



At the end of October, Donald Trump used words in topic 0 in his meetings to connect with voters, emphasize understanding of the country, and encourage political action.



Sentiment Analysis Based on BERT



Conclusion



Dominant Themes: Focus on immigration, economy, and rivals to address voter concerns.

Emotional Engagement: Use of charged language to connect emotionally and foster urgency.

Temporal Dynamics: Fluctuating sentiment as the campaign nears key electoral periods.

Overall: Sentiment and topic analysis are key to understanding political communication strategies and voter perception. Future campaigns could broaden topics to engage more voters and address diverse interests.

Q&A



Appendix

- **Source code:**
 - **Github repo:** <https://github.com/porimol/topic-modelling>
 - **Notebook:**
<https://github.com/porimol/topic-modelling/blob/main/election-campaign-sentiment-analysis.ipynb>





Thanks for listening!