Analyzing Public Opinion on YouTube: A Comparative Study of Trump vs. Biden and Trump vs. Harris

Sonia Borsi

September 9, 2024

Abstract

This paper employs sentiment analysis and topic modelling to examine the opinions expressed in YouTube comments from two videos: one on the 2024 Trump vs. Biden debate and another on Kamala Harris vs. Trump. Sentiment analysis using VADER revealed that comments on the Harris vs. Trump video were more positive. Topic modelling with LDA identified themes of leadership and politics. The study offers insights into public opinion in the context of the 2024 election. The full analysis is available on my GitHub repository.

1 Introduction

In the context of the digital social era, social platforms such as YouTube have assumed a pivotal role as a primary site for public discourse, particularly in relation to salient political events. The objective of this study is to analyse the sentiments and the main topics present in the comments of two YouTube videos published by the Wall Street Journal. The first video concerns the debate between Biden and Trump in the context of the 2024 presidential campaign. The second video features experts engaged in a comparative analysis of Trump and the emerging candidate, Kamala Harris. This analysis has two objectives: firstly, to identify the sentiments expressed in the comments below both videos and the principal topics that emerge from public discussion. Furthermore, an investigation is conducted to ascertain whether there have been any alterations in these sentiments over time. This study employs natural language processing (NLP) techniques, including the VADER tool for sentiment scoring and the Latent Dirichlet Allocation (LDA) tool for topic modelling, in order to examine how public perception and dialogue may have shifted as the 2024 election landscape evolved. The objective of this analysis is to elucidate the evolving nature of public opinion surrounding the two videos in question, thereby providing a more nuanced understanding of how the candidacies of Trump, Biden, and Harris are perceived by the online community. The study employs two extensive datasets of comments extracted from the two videos using the YouTube API, which are subsequently processed and analysed.

2 Related work

Sentiment analysis (SA) has gained prominence in natural language processing as a tool for extracting public opinion from text, especially on platforms such as YouTube. SA is used to determine the polarity of comments—positive, negative, or neutral—using machine learning (ML) algorithms or lexicon-based approaches. Studies, such as those by Mehmood et al. (2018) [6], have shown that sentiment analysis applied to social media data can reveal trends in public opinion, particularly in response to real-world events. In a political context, Feng et al. (2023) [3] demonstrated that sentiment analysis of data from social platforms can be used to predict election outcomes, highlighting that public sentiment on these platforms can indicate broader voting trends. In addition to sentiment classification, topic modeling techniques such as Latent Dirichlet Allocation (LDA) have been used to identify hidden topics in large textual datasets. Blei, Ng, and Jordan (2003) [12] introduced LDA as a generative probabilistic model that represents documents as mixtures over latent topics, making it highly effective for text classification and collaborative filtering tasks. Liu (2012) [10] also applied this combination of SA and topic modeling to various domains, including elections and consumer reviews, offering deeper insights into the relationship between sentiment and key topics. For example, LDA helps group words into topics, while sentiment analysis classifies the sentiment expressed in these topics. The combined use of SA and LDA creates a more holistic approach to understanding public opinion by revealing not only how people feel but also what they are discussing. Moreover, Badjatiya et al. (2017) [11] emphasized that this hybrid approach of integrating SA with topic modeling is particularly useful in analyzing large datasets and extracting meaningful trends that reflect public sentiment and key thematic concerns. By utilizing LDA to structure textual data and sentiment analysis to gauge emotional reactions, this method provides deeper insights, making it invaluable for applications in politics, marketing, and media analysis.

3 Methodology

3.1 Data Collection

The first video selected for analysis is the presidential debate between Donald Trump and Joe Biden (*"Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"*). The video can be accessed via this link. The video, published by the Wall Street Journal, has been viewed 22 million times and has attracted almost 130,000 comments. The second video to be considered in this analysis is one in which the Wall Street Journal poses a series of questions to a Democratic strategist, a political scientist and a pollster (*"Kamala Harris: A Stronger Candidate Against Trump? — WSJ"*). The questions pertain to the potential advantages and disadvantages of Kamala Harris, who is running against the former President, Donald Trump, available at the link. The video has been viewed 2.3 million times and has attracted 15,000 comments. The number of interactions is demonstrably lower than that of the previous video, but at the time of this analysis, it represents one of the most popular YouTube videos in which a comparison between Kamala and Trump is presented. In this analysis, data was collected using the YouTube Data API to obtain comments about the two videos published by the Wall Street Journal. The two videos selected for analysis were identified by unique video IDs, which were then used to automate the retrieval of comments. Given the considerable number of comments on both videos, it was necessary to select a subset for analysis. To ensure that the analysis could be conducted with the available computational resources, a batch of 5,000 comments from each video was selected for examination, and the comments have been exported to a CSV file (one for each video).

3.2 Data Preprocessing

In order to conduct a comparative analysis of the sentiments expressed by users in the comments of the two videos mentioned above, an extensive pre-processing of the data has been undertaken in order to prepare it for analysis. The text data was subjected to cleaning and normalisation procedures to ensure its suitability for subsequent analysis.

Text Cleaning The initial phase of data pre-processing in this study entailed the cleansing of textual data, specifically the raw comments extracted from the two videos. Text cleaning is a pivotal step in the domain of NLP, aiming to transform the data into a uniform and coherent format, thereby reducing inherent variability within the data set. This process entailed a series of pivotal transformations, including:

- Unicode Normalization: this process aims to standardize the data by converting text to a standard ASCII format, which is particularly beneficial when dealing with non-standard characters or text encoded in disparate formats. In fact characters are converted to their nearest ASCII equivalent. It is also possible to remove special characters and symbols from the text.
- Lowercase Conversion: This stage serves to guarantee uniformity and to reduce the dimensionality of the data by treating words with different cases as the same token.
- Removal of Special Symbols: this process involved removing URLs and HTML tags from the comments, as they can introduce noise into the data.
- Expansion of Contractions: the feature utilises the Contractions library to transform abbreviated forms into their complete counterparts.

Furthermore, additional transformations have been implemented, including the removal of punctuation and digits (including periods, commas, exclamation marks and so forth), the replacement of newlines with spaces, and the collapsing of multiple spaces into a single space. In order to enhance the analytical value of the text, certain words that add no meaningful value have been removed.

Text Processing

- The tokenisation of text, whereby the text is split into individual units called tokens, represents a crucial step in the preprocessing phase. The purpose of this step is to convert raw text into a structured format that can be easily processed by algorithms, such as those used in the present analysis.
- Lemmatization is the process of reducing words to their base or root form, known as the "lemma." This method considers the context and meaning of the word in order to ensure that it is transformed into a meaningful base form. In order to conduct this analysis, the Natural Language Toolkit (NLTK) has been employed. This is a tool that utilises the WordNet lexical database in order to identify the appropriate lemma for a given word, taking into account its part of speech.

3.3 Sentiment Analysis

Sentiment Analysis is the computational study of people's opinions, attitudes and emotions toward an entity (Medhat, Hassan, & Korashy, 2014) [5]. The sentiment analysis of the comments on the two YouTube videos comprised two main phases: the first was a sentiment scoring phase (negative, positive, neutral), during which all the comments were processed through Vader. The second was a feature engineering phase, during which additional features were extracted to complement the scores. Finally, the comments were categorised based on the compound scores in the sentiment categorisation phase.

Scores and Feature extraction In order to analyse the sentiment expressed in the two YouTube videos under consideration in the present analysis, the Valence Aware Dictionary and Sentiment Reasoner (VADER) was employed. This is a widely used instrument in the field of NLP especially for analysing social media data, as it is highly sensitive to both the positive or negative valence of emotions expressed in textual data, as well as the intensity of those emotions. Indeed, VADER has been developed to address the challenges posed by informal language, including slang and abbreviations, which are prevalent on social media. This makes it an optimal choice for this study. The VADER tool provides a compound score, which represents the overall sentiment of a text, as well as separate scores for positive, neutral, and negative sentiment components. The score include:

- **Compound Score**: A normalized score representing the overall sentiment, ranging from -1 (most negative) to +1 (most positive).
- **Positive, Neutral, and Negative Scores**: These scores reflect the proportion of the text that expresses positive, neutral, and negative sentiment, respectively.

Furthermore, additional text-based metrics have been calculated, including the popolarity, the length of the comments, the word count, and the sentiment shift, which is calculated as the difference between the positive and negative scores. This provides insight into the overall sentiment balance within a comment. Additionally, readability has been assessed using the Flesch-Kincaid grade level, which estimates the complexity of the text. Comments that are more readily comprehensible may be more likely to convey a discernible sentiment.

3.4 Topic Modeling

Topic modelling is a machine learning technique that is primarily employed to discern the existence of topics, which can be defined as clusters of words that frequently co-occur. More specifically, a topic model identifies themes by identifying repeated patterns of words and grouping these patterns of words into topics that reflect the content of the documents (Churchill & Singh, 2022) [4]. In this study, Latent Dirichlet Allocation, a popular topic modelling algorithm, has been employed.

Corpus Preparation and Vectorization Prior to the application of the LDA model, the comments associated with the two YouTube videos were prepared using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorisation technique. This is an extremely effective method which transforms the text data into a matrix, with each row representing a document (in this case, the comment) and each column a word. The values in the matrix indicate the relative importance of each term within the comment in comparison to the entire corpus. This approach serves to highlight words that are not only frequent but also contextually significant, thereby enhancing the effectiveness of the LDA model in identifying pertinent topics.

Latent Dirichlet Allocation (LDA) In order to ascertain the distribution of topics across the comments, the Latent Dirichlet Allocation (LDA) algorithm has been employed. LDA is a three-level hierarchical Bayesian model, in which each item of a collection is modeled as a finite mixture over an underlying set of topics (Blei, Ng, & Jordan, 2003) [12]. It identifies a set of underlying topics within a collection of documents by determining the probability distribution of words across topics and topics across comments. Following the fitting of the LDA model, the performance was evaluated using coherence scores, a measure of the semantic similarity between words within the same topic. High scores indicate the presence of topics that are both interpretable and meaningful. The coherence score is a pivotal metric in guaranteeing that the topics generated by the LDA model are not merely statistically significant, but also meaningful within the context of the data. In this study, the coherence score was calculated for a range of topic numbers in order to identify the optimal number of topics that best represented the data.

N-gram Analysis In order to enhance the efficacy of the LDA model in topic modelling, N-gram analysis has been integrated. This analysis employed TF-IDF vectorisation, enabling the extraction of the most frequent unigrams (comments) from the corpus and facilitating a more detailed understanding of the topics that commenters discussed.

4 Experimental results

The analysis of the two YouTube videos yielded several noteworthy experimental results that could prove useful for further investigation and consideration into the evolving sentiments of the American public in relation to political events.

Sentiment Analysis Results The sentiment analysis, has yielded the scores associated with the three primary sentiments (neutral, positive, and negative) for both videos. These scores are presented in the table 1. The comparative analysis of the Trump vs. Biden and Kamala vs. Trump debates revealed the following key observations: in comparing the Trump and Biden debates, 37.31% of the comments were positive, with an average popularity score of 0.0526. In contrast, the Kamala vs. Trump debate elicited a greater proportion of positive comments, at 41.75% and a higher average popularity score, at 0.0985. The proportion of negative comments was markedly lower in the Kamala vs. Trump debate, accounting for 24.20% of the total. The results demonstrate that the Kamala vs. Trump video was met with a more favourable response than the Trump vs. Biden debate video, characterised by a reduction in negative commentary and an elevated average popularity score.

Table 1: *Comparison of Sentiment Analysis Between Presidential Debate 2024 and Kamala vs Trump comparison video*

Event	Trump vs Biden	Kamala vs Trump
Number of Positive Comments	1816	1551
Number of Neutral Comments	1619	1265
Number of Negative Comments	1432	899
% of Positive Comments	37.31%	41.75%
% of Neutral Comments	33.26%	34.05%
% of Negative Comments	29.42%	24.20%
Average Popolarity	0.0526	0.0985

Additionally, the appendix contains Figures 9/10, which present unigrams of comments categorised by sentiment (positive, neutral, and negative) for both videos under analysis, illustrating the frequency of the most common words within each sentiment category, thereby offering insights into the specific language used by commenters depending on their sentiment. Moreover, in order to facilitate a more nuanced understanding of the differences in topics and between neutral, negative, and positive sentiments for both videos, word clouds were generated. In the Trump vs. Biden debate, positive comments frequently include words such as "please," "confidence," and "Trump," which suggest support for Trump. The neutral comments frequently mention both "Biden" and "Trump," indicating a balanced discussion devoid of strong emotional tones.

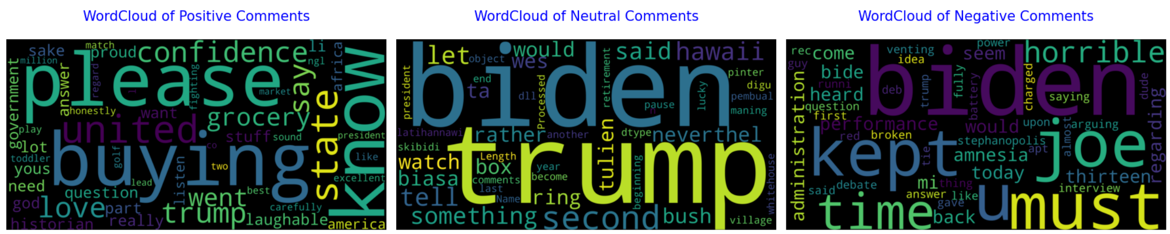

Figure 1: Wordclouds for positive, negative and neutral comments, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"

The negative comments are characterised by the frequent mention of "Biden" and "horrible," reflecting a criticism that is primarily directed at Biden. In the Kamala vs. Trump debate, positive comments emphasise the words "Harris" and "win," indicating support for her. Neutral comments mention both "Trump" and "Kamala," reflecting a more impartial perspective. The negative comments frequently mention both Trump and Biden, indicating that criticism was directed at both figures and at broader political dissatisfaction.

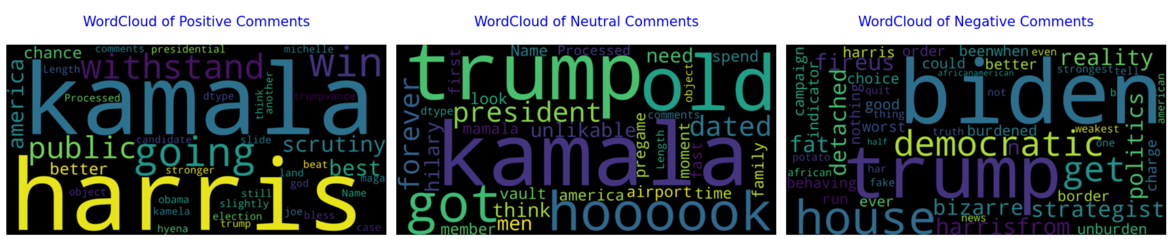

Figure 2: Wordclouds for positive, negative and neutral comments, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

Topic Modeling Results The results of the topic modelling, using Latent Dirichlet Allocation, led to the identification of two principal topics for the comments in each video. Each topic resulting from this analysis is constituted by a set of words, with each word being associated with a weight that can be considered as a measure of the importance of the word in the context of the overall topic. The most salient words and their respective weights are presented in Table 2 and Table 3 for both videos.

The results of the topic modelling performed on the video of the debate indicate the emergence of two distinct topics. Topic 1 is primarily concerned with the overarching themes of the debate, with a notable focus on the candidates "Trump" and "Biden," who have the highest word weights. Topic 2 is concerned with the concept of national identity and the role of leadership in this context. It is noteworthy that the terms "country," "American," and "president" are particularly prevalent in this regard. While the terms "Biden" and "Trump" are also of significance in this topic, the discussion here places greater emphasis on their roles in relation to national issues.

Table 2: *Words and Weights for Topic 1 and Topic 2 for the video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"*

Topic 1		**Topic 2**	
Word	**Weight**	**Word**	**Weight**
people	177.4996	country	136.5976
world	178.2672	american	137.2251
debate	188.4894	president	149.5883
time	188.5250	america	156.3601
country	190.2030	joe	241.8029
one	196.1886	debate	241.9783
know	209.9884	people	244.6660
president	454.3570	trump	249.1930
biden	1225.6742	like	357.3391
trump	1362.0146	biden	579.9475

In regard to the video comparing Kamala and Trump, Topic 1 is concerned with the 2020 U.S. election, with particular attention paid to figures such as Trump and Kamala Harris. The focus is on the aspects of voting, winning, and Harris's identity as a woman of colour. Topic 2 is concerned with the presidency, and includes discussions of Trump, Biden, and Kamala Harris, as well as references to media coverage and the transition of power.

Table 3: *Words and Weights for Topic 1 and Topic 2 of the video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"*

Topic 1 K		**Topic 2 K**	
Word	**Weight**	**Word**	**Weight**
trump	414.950083	trump	493.146866
harris	401.161043	kamala	386.207179
kamala	331.219477	harris	361.387902
vote	288.370978	president	244.064324
woman	231.581279	biden	227.962730
win	200.902257	kamala harris	136.005728
american	165.131060	going	87.651718
biden	157.288311	wsj	87.511164
would	155.701600	joe	84.369650
black	152.056879	would	83.469436

Furthermore, a LDA visualisation has been conducted utilising the pyLDAvis tool, which offers an interactive visual representation of the identified topics and their interconnections. The visualisations are included in the appendix for further investigation of the topic distribution and word importance across the debates. Furthermore, Principal Component Analysis (PCA) has been employed with the objective of visualising the intertopic distances between the two topics for each video, which is beneficial in identifying the spatial relationship between the topics (FIGURES 1 and 2 in the appendix).

Further details regarding the analysis can be found in the Appendix of the present paper.

5 Conclusion

This study presents a comparative analysis of the comments associated with two key YouTube videos in relation to the American Presidential campaign of 2024. The first is the 2024 Trump vs. Biden presidential debate, and the second is a video discussing Kamala Harris as a candidate in the context of her comparison to Trump, following Biden's renouncement. The two significant political videos have been subjected to sentiment analysis using VADER and topic modelling using LDA, resulting in the extraction of valuable insights into public opinion surrounding these events. Indeed, the sentiment analysis indicated that Kamala Harris was perceived more favourably than in the Trump vs. Biden debate, which was characterised by a high level of polarisation. The topic modelling results served to further highlight the existence of distinct themes in each video, with discussions around leadership, national identity, and policy emerging as key points in both cases. This analysis demonstrates the dynamic nature of the sentiments expressed by individuals on social media platforms such as YouTube in relation to political events, illustrating how public opinion can evolve in response to shifts in political contexts. Further research could build on this by analysing a more extensive range of videos and comments, or by incorporating additional natural language processing techniques in order to gain a deeper understanding of the ways in which digital discourse influences and reflects political narratives.

References

[1] Kumar, V., & Subba, B. (2024). A TfidfVectorizer and SVM based sentiment analysis framework for text data corpus. *Proceedings of IEEE International Conference on Big Data*, 2098-2103.

[2] Singh, R., & Tiwari, A. (2021). YouTube Comments Sentiment Analysis. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 5(5), 1-11.

[3] Feng, G., Chen, K., Cai, H., & Li, Z. (2023). A Hybrid Method of Sentiment Analysis and Machine Learning Algorithm for the U.S. Presidential Election Forecasting.

[4] Churchill, R., & Singh, L. (2022). The Evolution of Topic Modeling. *ACM Computing Surveys*, 54(10s), 1-35.

[5] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.

[6] Mehmood, K., Essam, D., & Shafi, K. (2018). Sentiment Analysis System for Roman Urdu. In *Science and Information Conference* (pp. 29-42). Springer, Cham.

[7] Athar, A. (2014). Sentiment Analysis of Scientific Citations (No. UCAM-CL-TR-856). *University of Cambridge, Computer Laboratory.*

[8] Poria, S., Cambria, E., Gelbukh, A., Bisio, F., & Hussain, A. (2015). Sentiment Data Flow Analysis by Means of Dynamic Linguistic Patterns. *International Journal of Computational Linguistics and Applications.*

[9] Turney, P. D., & Mohammad, S. M. (2014). Experiments with Three Approaches to Recognizing Lexical Entailment. *Journal of Artificial Intelligence Research*, 51, 275-310.

[10] Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Morgan & Claypool Publishers.*

[11] Badjatiya, P., Gupta, S., Gupta, M., & Varma, V. (2017). Deep Learning for Hate Speech Detection in Tweets. *Proceedings of the 26th International World Wide Web Conference (WWW)*, 759-760.

[12] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.

[13] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP).* arXiv preprint arXiv:cs/0205070.

6 Appendix

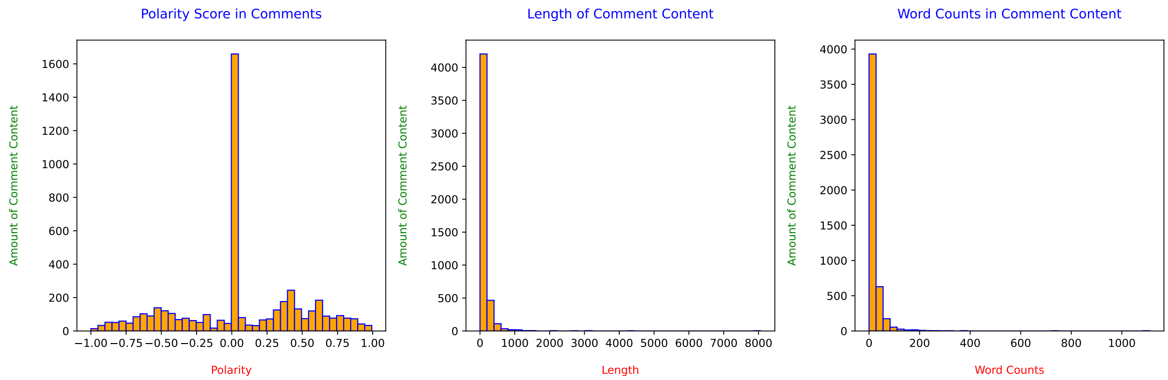

Figure 3: Analysis of YouTube Comments by Polarity, Length, and Word Count, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"; This figure presents three histograms analyzing various aspects of YouTube comments. The first histogram (left) shows the distribution of comment polarity scores, indicating whether the comments are negative, neutral, or positive, with a sharp spike at neutral polarity (0). The second histogram (center) illustrates the distribution of comment lengths, with the majority of comments being short in length. The third histogram (right) represents the word count distribution of comments, showing that most comments have fewer than 100 words. These graphs provide an overview of the sentiment, length, and word frequency characteristics of the dataset.

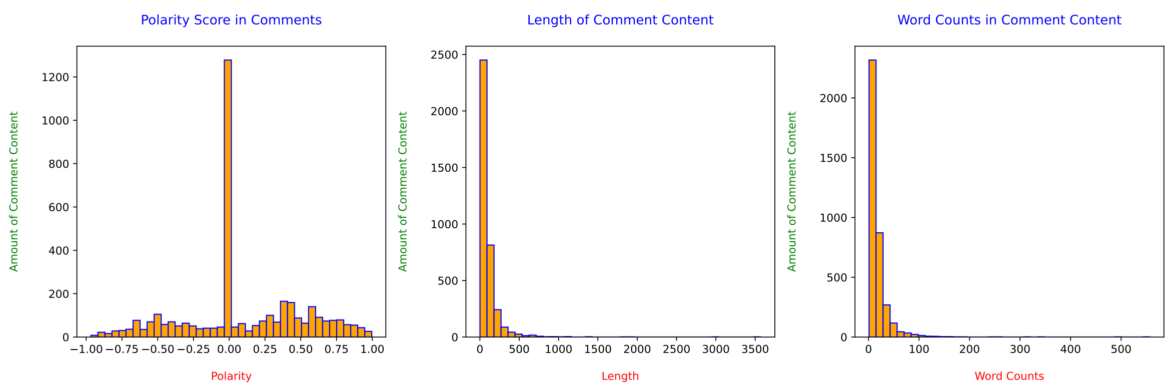

Figure 4: Analysis of YouTube Comments by Polarity, Length, and Word Count, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

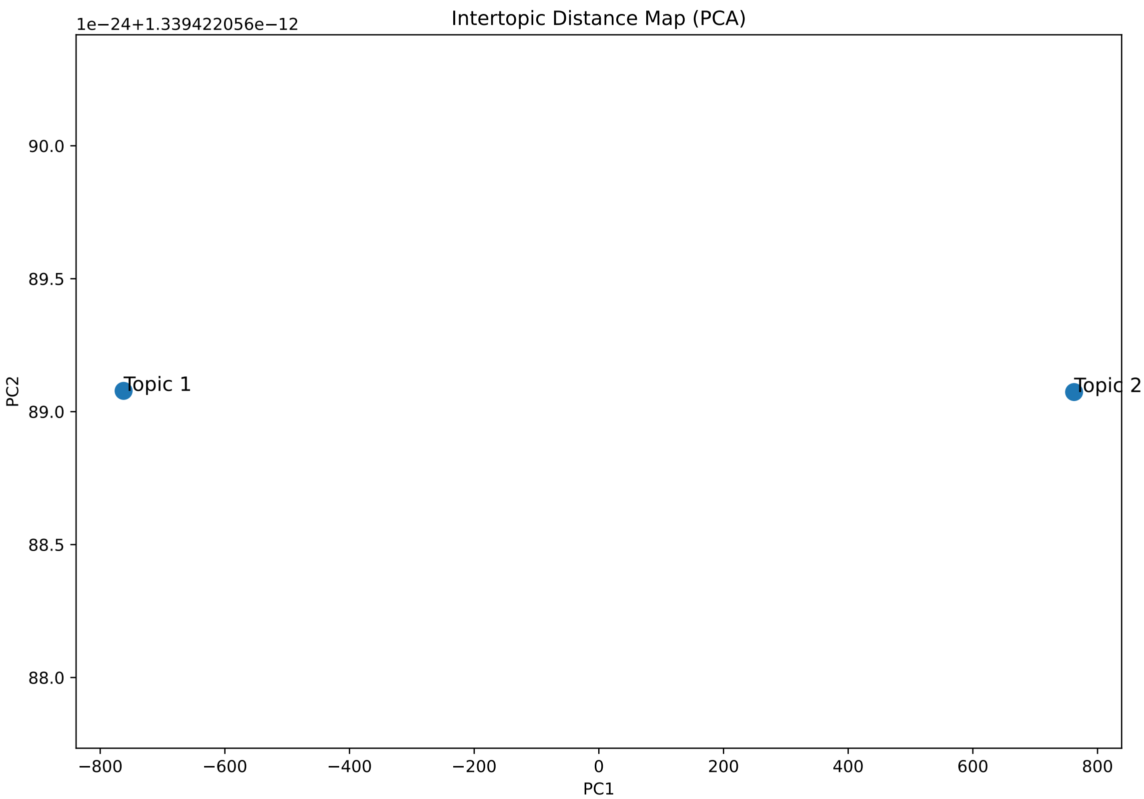

Figure 5: Intertopic distance map, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ". The map represents two distinct topics, each visualized as a point in the 2D plane. The distance between the topics indicates their dissimilarity, with Topic 1 and Topic 2 being shown as relatively distant from each other, reflecting the divergence in content or themes associated with each topic discussed during the debate.

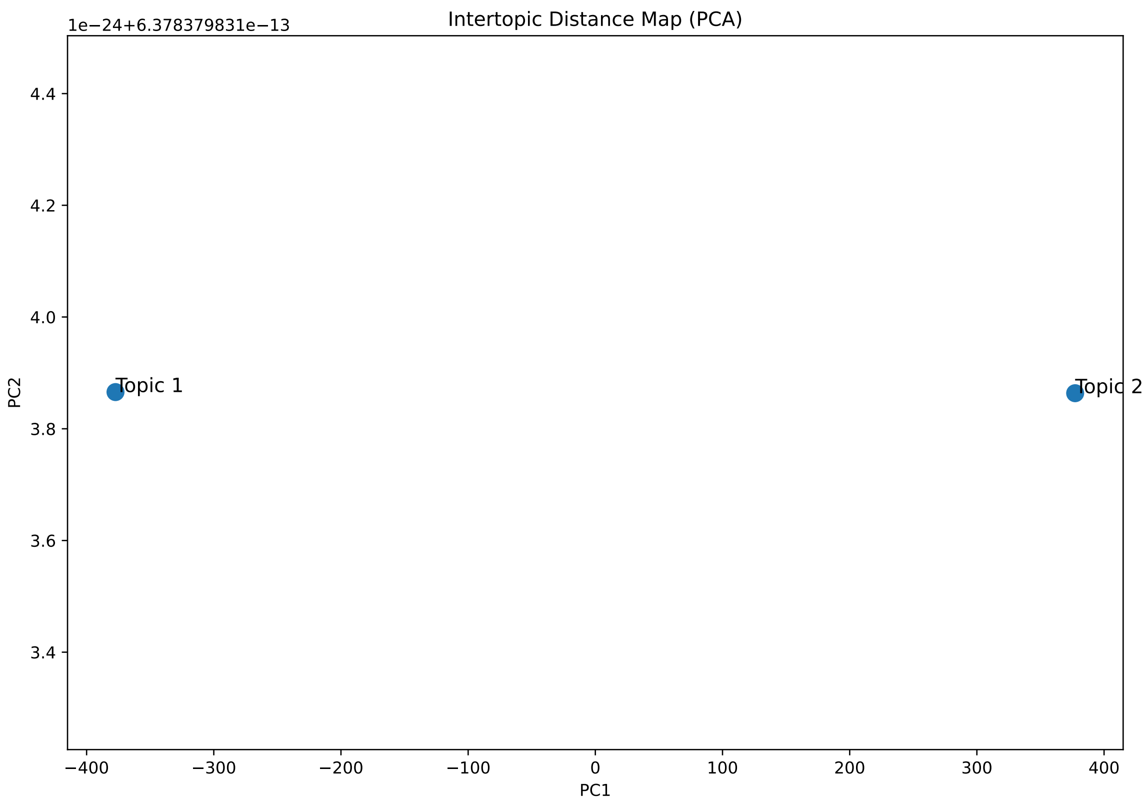

Figure 6: Intertopic distance map, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

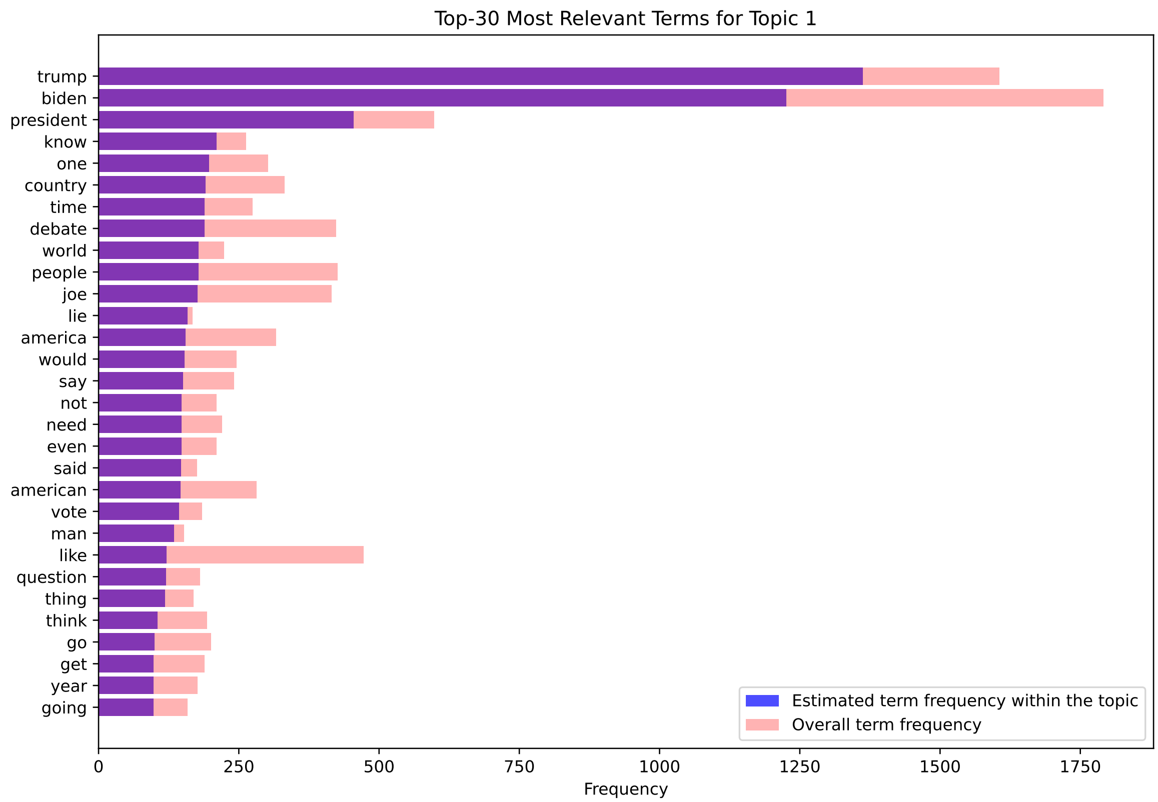

Figure 7: Top 30 Most Relevant Terms for Topic 1, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"; The bar chart displays the top 30 most relevant terms for Topic 1, ranked by frequency. The purple bars represent the estimated term frequency within the topic, while the pink bars show the overall term frequency across the entire dataset. Terms like "Trump," "Biden," and "president" dominate the topic, reflecting the prominence of political discussion. Other frequent words include "country," "world," and "America," indicating broader thematic concerns within this topic. The graph helps visualize which terms are most strongly associated with Topic 1 and how their frequency within the topic compares to their overall occurrence.

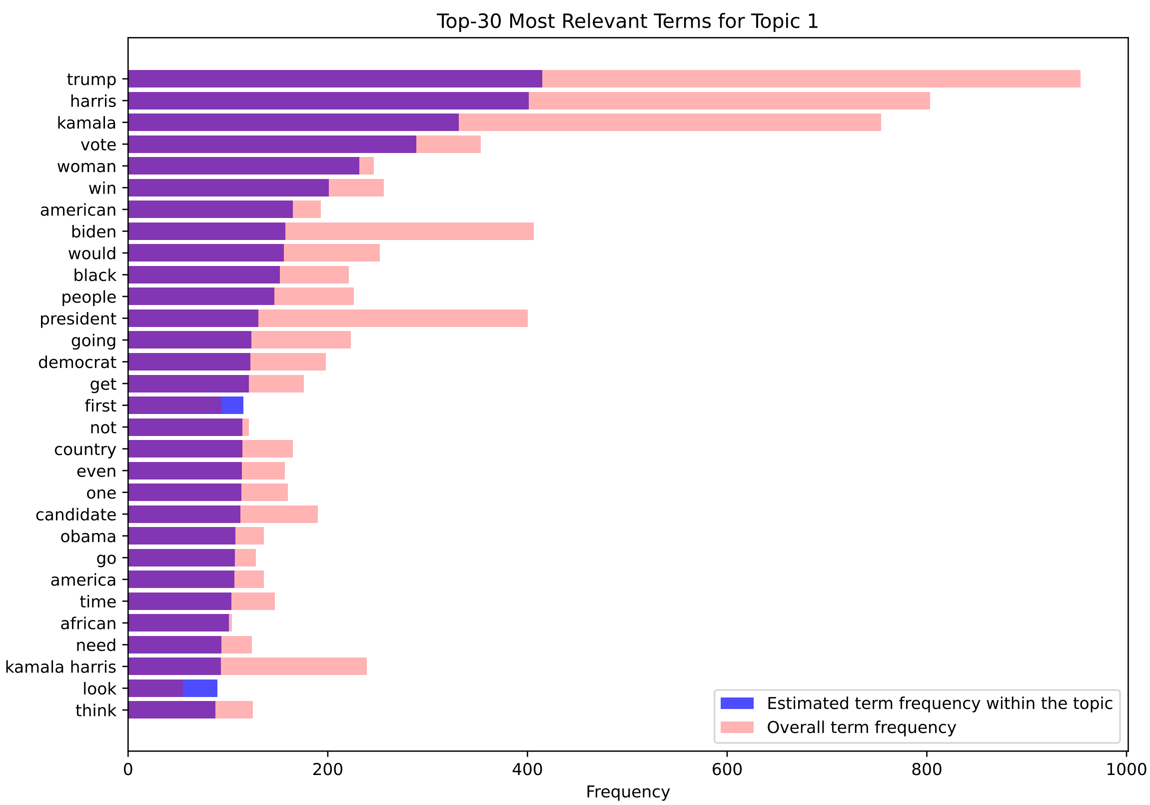

Figure 8: Top 30 Most Relevant Terms for Topic 2, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"

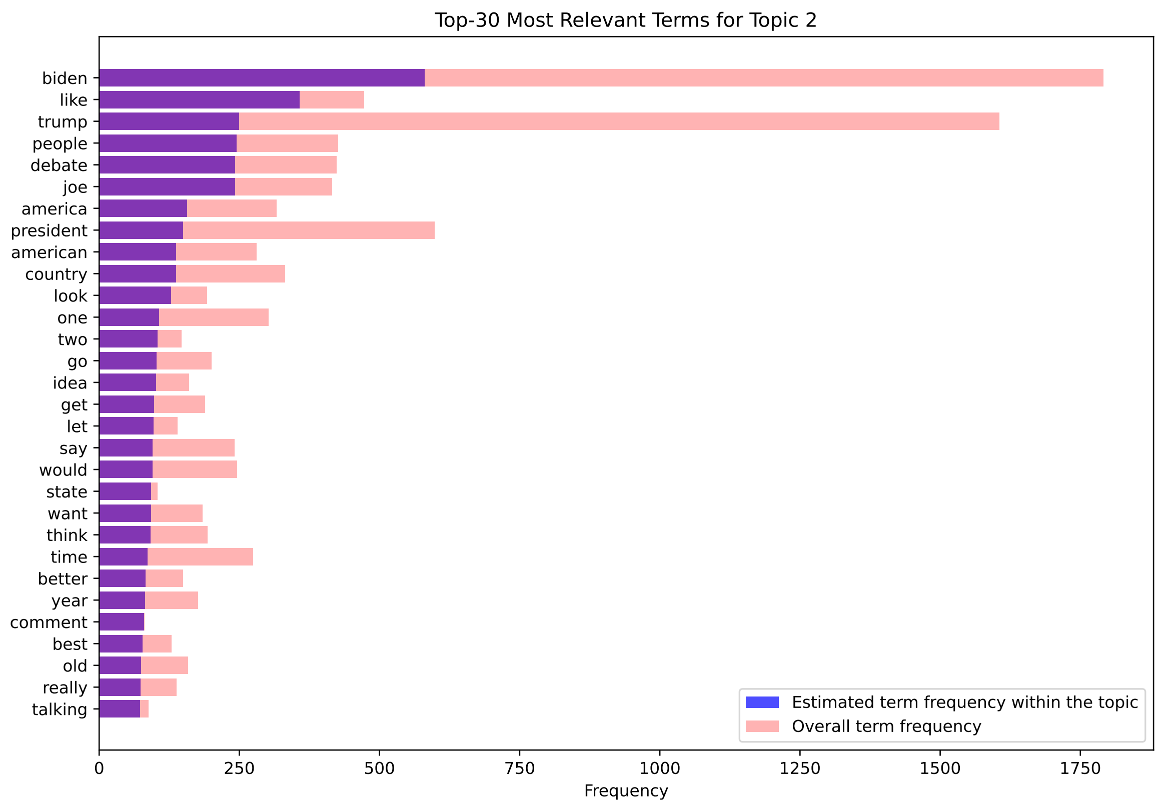

Figure 9: Top 30 Most Relevant Terms for Topic 1, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

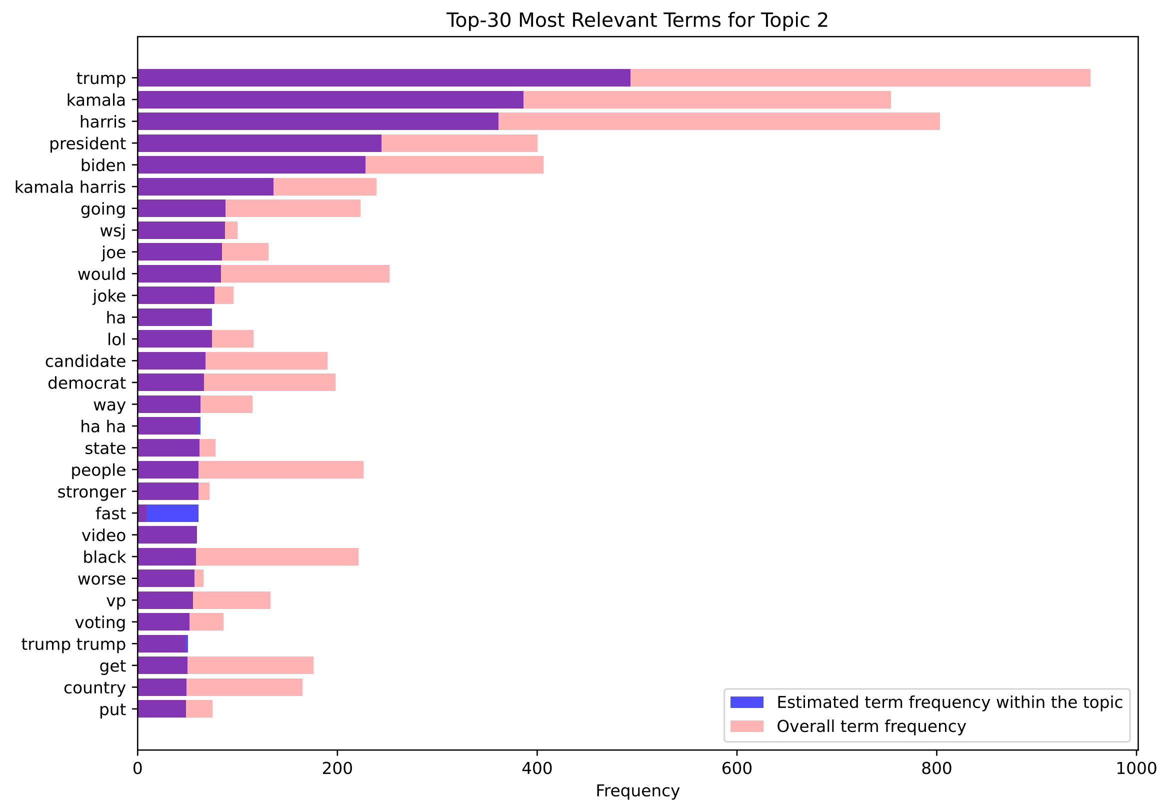

Figure 10: Top 30 Most Relevant Terms for Topic 2, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

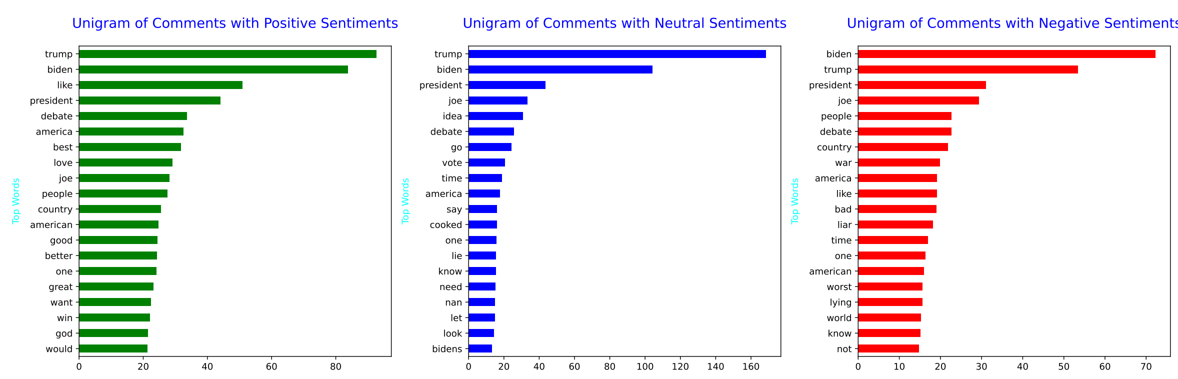

Figure 11: Unigram Distribution of Comments by Sentiment, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ". The charts are divided by sentiment: positive (left, green), neutral (middle, blue), and negative (right, red). In the positive sentiment chart, frequent terms like "Trump," "Biden," and "president" are highlighted, reflecting favorable associations in the comments. The neutral sentiment chart displays terms such as "Trump," "Biden," and "president" as well, but within a neutral context, indicating balanced discussion. On the negative side, words like "Biden," "Trump," "people," and "disaster" appear frequently, indicating critical or negative reactions from viewers. These charts collectively show the distribution of unigrams across different sentiments, offering insight into how commenters express their opinions on the debate.

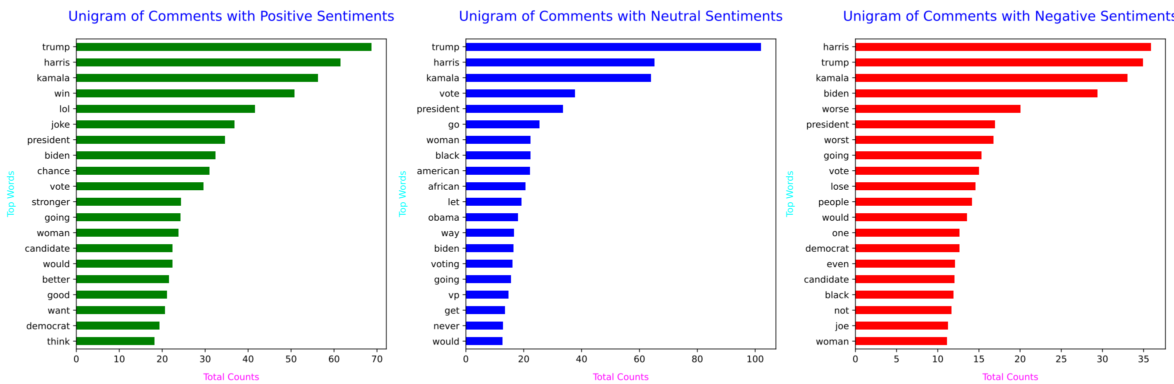

Figure 12: Unigram Distribution of Comments by Sentiment, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ". In the positive sentiment chart, the words "Trump," "Harris," "strong," and "candidate" are among the most frequently mentioned, indicating positive associations in the comments. In the neutral sentiment chart, terms such as "Trump," "Harris," and "president" are prominent, reflecting a neutral tone in the discussion. The negative sentiment chart shows frequent use of words like "Harris," "Trump," and "weak," suggesting critical or unfavorable comments, particularly directed toward Harris and Trump.

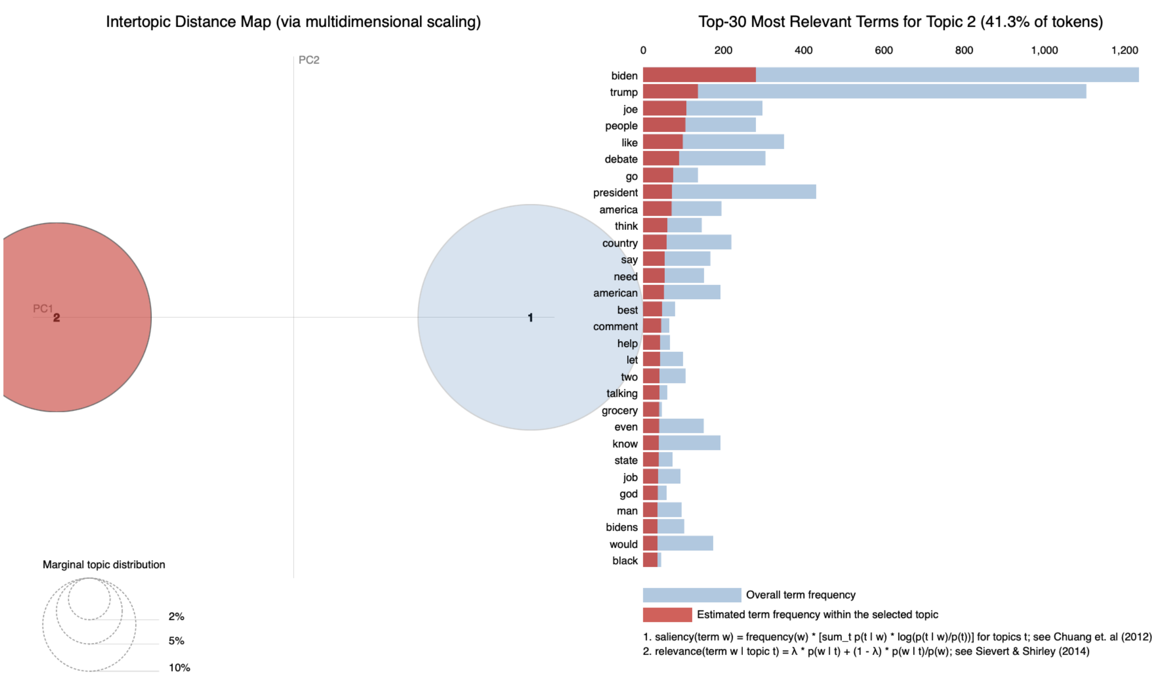

Figure 13: Intertopic Distance Map and Top-30 Most Relevant Terms for Topic 1 using pLDAvis, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ". The intertopic distance map on the left visualizes the spatial relationship between two distinct topics, with the size of the circles representing their prevalence. Topic 1, displayed in blue, is the focus of the bar chart on the right, which highlights the top 30 most relevant terms associated with this topic. The bar chart distinguishes between the estimated term frequency within Topic 1 (in red) and the overall frequency of the term across the corpus (in light blue), showing key terms.

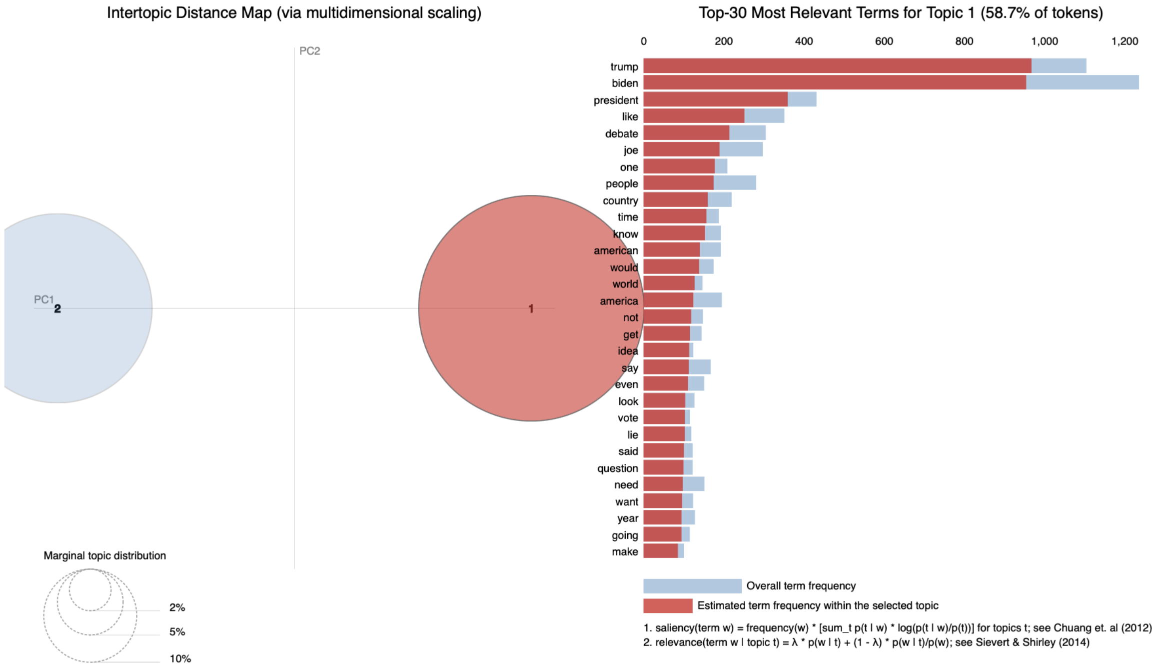

Figure 14: Intertopic Distance Map and Top-30 Most Relevant Terms for Topic 2 using pLDAvis, video: "Full Debate: Biden and Trump in the First 2024 Presidential Debate — WSJ"

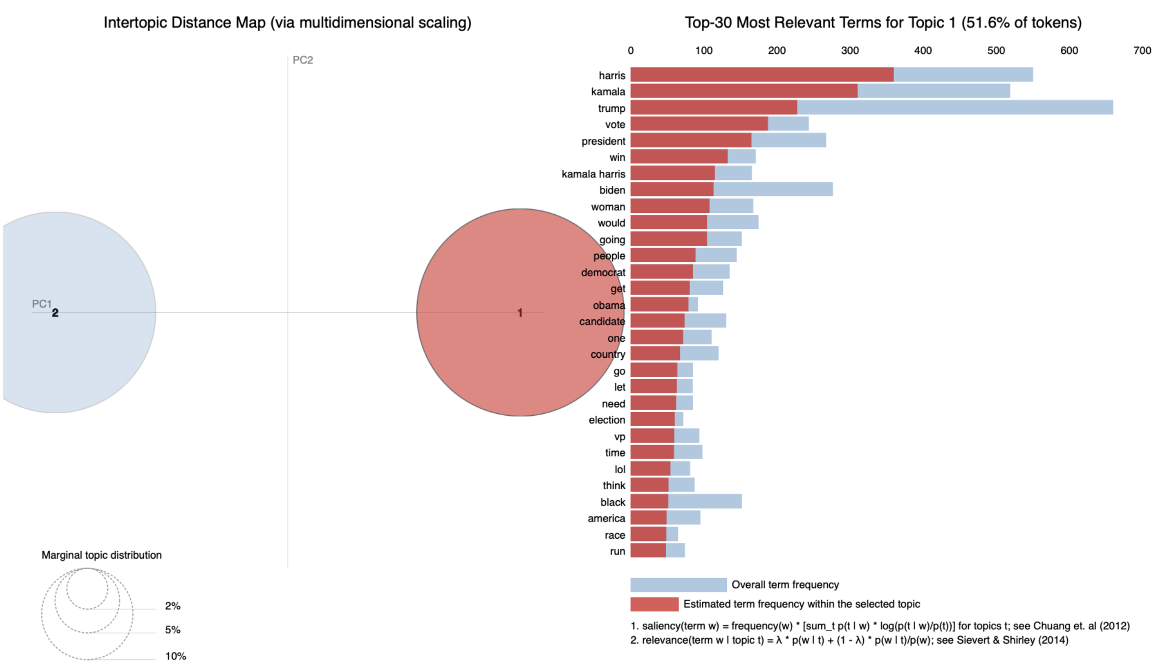

Figure 15: Intertopic Distance Map and Top-30 Most Relevant Terms for Topic 1 using pLDAvis, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"

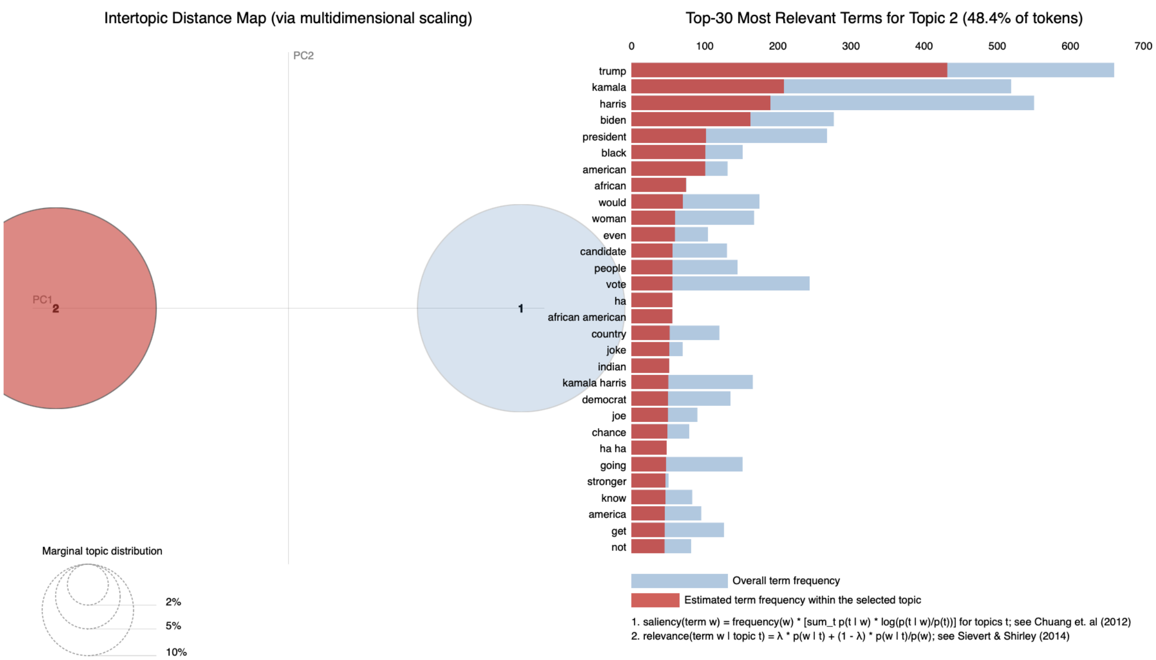

Figure 16: Intertopic Distance Map and Top-30 Most Relevant Terms for Topic 2 using pLDAvis, video: "Kamala Harris: A Stronger Candidate Against Trump? — WSJ"