Utrecht University

Text and Media Analytics

Assignment 2

Analyzing Emotional Dynamics in YouTube Comments: Biden vs. Trump

Instructor: Nguyen Dennis

Vasilopoulos Konstantinos (1415638)

Word Count: 2072

Table of Contents

Introduction

In today's digital world, social media sites have become places for political debate, where everyone can quickly communicate their opinions, and reactions. YouTube is different from the other platforms since it is not just a place to store multimedia content but has also a dynamic community where people can remark and share their thoughts. By looking at these comments made by users, you can get a unique look at how people feel about political leaders and events.

The aim of this research is to explore the emotional dynamics present in YouTube comments related to the two prominent political figures.

To achieve this, we used advanced Natural Language Processing (NLP) techniques, specifically transformer-based models like RoBERTa, which provide better contextual understanding (**Agrawal et al., 2021).**

We seek to detect and analyze the prevalence of various emotions—such as anger, disgust, fear, joy, sadness, surprise, and neutrality—. This analysis endeavors to uncover patterns and differences in public emotional responses, thereby contributing to a deeper understanding of the sociopolitical landscape as reflected in digital interactions.

Given that Joe Biden and Donald Trump have very different public personas, it makes sense to conduct a comparative emotion analysis of their YouTube comment sections. This study contributes to the increasing amount of research on natural language processing (NLP), discussing politics, and identifying emotions

Literature Review

The study of emotion in online political discourse has become a prominent area of research within computational social science and natural language processing (NLP).

Historically, emotion detection in textual data relied on lexicon-based approaches such as the NRC Emotion Lexicon **(Mohammad & Turney, 2013)** or WordNet Affect. These resources mapped words to predefined emotional states like joy, anger, or sadness. While foundational, these models often struggled with context sensitivity—particularly in politically nuanced or sarcastic texts—leading to misclassification **(Staiano & Guerini, 2014)**. To address these limitations, deep learning models, and more recently, transformer-based architectures, have become the gold standard for emotion detection.

BERT (Bidirectional Encoder Representations from Transformers), introduced by **Devlin et al. (2019)**, revolutionized NLP by enabling bidirectional context encoding, allowing models to interpret the meaning of a word based on its surrounding words. Building on this, **Liu et al. (2019)** proposed RoBERTa (Robustly Optimized BERT Pretraining Approach), which removed the Next Sentence Prediction task and trained on significantly more data. RoBERTa demonstrated superior performance across a variety of classification benchmarks, including sentiment and emotion classification.

These models have been applied in political domains with significant success. **Wisnubroto and Çılgın (2023)** analyzed over 260,000 YouTube comments from the 2023 Turkish general elections using emotion analysis tools. Their study found that anger and joy were particularly dominant, with emotion trends correlating closely with key political events. Their research emphasized YouTube as an effective medium for real-time sentiment tracking during political campaigns.

In the U.S. context, **Shevtsov et al. (2021)** investigated emotional framing in Twitter content during the transition from President Trump to President Biden. Using automated text analysis and emotion labeling, they discovered that Trump's digital discourse often invoked anger and fear, whereas Biden's messaging was more frequently associated with hope and reassurance. These findings align with **Jamdar et al. (2015)**, who observed that emotional patterns in political language tend to diverge significantly depending on rhetorical strategy and target audience.

Recent research has also explored hybrid model architectures. **Zhang et al. (2023)** proposed a RoBERTa-LSTM hybrid model for social media sentiment analysis, achieving higher accuracy than either architecture alone. By combining RoBERTa's contextual embeddings with LSTM's sequential learning capacity, the model captured both the content and flow of emotional expression—an approach particularly suited for threaded discussions like YouTube comment chains.

These studies all show how useful and effective transformer-based models are when applied to political social media data. They also highlight a recurring observation: emotional engagement on the internet varies not only with the subject matter but also with the person in charge of it.

Data Methodology

The datasets used in this study are the sum of 36231 YouTube comments extracted from official videos of Joe Biden and Donald Trump in year 2023. Each entry in the dataset includes the following columns:

- **video_id**: The unique identifier of the YouTube video to which the comment belongs.
- **comment_id**: The unique identifier of the comment.
- **comment_text**: The textual content of the comment.
- **author**: The username of the individual who posted the comment.
- **comment_date**: The date and time when the comment was posted.
- **comment_like_count**: The number of likes the comment received.
- **in_reply_to**: If the comment is a reply, this field contains the ID of the parent comment.

It is worth noting that there are ethical and representational considerations relevant to the dataset. While all the data was publicly available, we ensure that comments are analyzed at an aggregate level to preserve user anonymity. Additionally, YouTube comment sections can exhibit varying degrees of moderation and audience bias depending on the nature of the content and the channel's demographics. These factors may introduce noise or skew the emotional tone of responses.

Before conducting any analysis, we examined the datasets for missing values using the **.info()** method in pandas. The Biden dataset contained 15,731 entries and the Trump dataset 20,500 entries, each with 13 columns. Both datasets showed excellent data completeness, with almost all columns fully populated. Only two fields in the Biden dataset—Comment Text Display and Comment Text Original—contained a single null value each (fig. 1). These minor gaps are statistically insignificant and do not pose a risk to downstream processing. This high level of completeness ensured a reliable foundation for subsequent preprocessing and emotion detection.

```
[6]: df_b.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15731 entries, 0 to 15730
Data columns (total 13 columns):
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   Comment ID                 15731 non-null  object
 1   Video ID                   15731 non-null  object
 2   Author Channel ID          15731 non-null  object
 3   Author Display Name        15731 non-null  object
 4   Author Profile Image URL   15731 non-null  object
 5   Author Channel URL         15731 non-null  object
 6   Comment Text Display       15730 non-null  object
 7   Comment Text Original      15730 non-null  object
 8   Comment Can Rate           15731 non-null  bool
 9   Comment Viewer Rating      15731 non-null  object
 10  Comment Like Count         15731 non-null  int64
 11  Comment Published At       15731 non-null  object
 12  Comment Updated At         15731 non-null  object
dtypes: bool(1), int64(1), object(11)
memory usage: 1.5+ MB
```

```
[8]: df_t.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20500 entries, 0 to 20499
Data columns (total 13 columns):
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   Comment ID                 20500 non-null  object
 1   Video ID                   20500 non-null  object
 2   Author Channel ID          20500 non-null  object
 3   Author Display Name        20500 non-null  object
 4   Author Profile Image URL   20500 non-null  object
 5   Author Channel URL         20500 non-null  object
 6   Comment Text Display       20500 non-null  object
 7   Comment Text Original      20500 non-null  object
 8   Comment Can Rate           20500 non-null  bool
 9   Comment Viewer Rating      20500 non-null  object
 10  Comment Like Count         20500 non-null  int64
 11  Comment Published At       20500 non-null  object
 12  Comment Updated At         20500 non-null  object
dtypes: bool(1), int64(1), object(11)
memory usage: 1.9+ MB
```

Fig. 1

To ensure the validity of our emotion analysis, we first verified the dominant language in the dataset. Using the **langdetect** library, we identified the primary language of each cleaned comment. Results showed that over 70% of comments in datasets were in English (fig. 2), confirming that the dataset is suitable for English-based NLP models like RoBERTa.

Prior to this step, we applied standard text preprocessing. Comments were lowercased, emojis were stripped, URLs removed, and unnecessary punctuation and whitespace were cleaned. This process helps ensure consistency and improves the model's ability to capture meaningful emotional cues in the text.

```
🔵 Biden Dataset Language Distribution 🔵 (%):
language
en       70.777446
no        3.458140
fr        3.089441
error     2.555464
af        2.313902
ro        2.154981
nl        1.538364
da        1.214163
it        1.099739
fi        1.080669
Name: proportion, dtype: float64
```

```
🔴 Trump Dataset Language Distribution 🔴 (%):
language
en       79.219512
error     6.126829
fr        1.921951
no        1.536585
ca        1.526829
es        1.131707
pt        0.839024
af        0.785366
cy        0.682927
ja        0.531707
Name: proportion, dtype: float64
```

Fig. 2

In this study, emotion classification was performed using a transformer-based model. We defined seven emotion categories—**anger**, **disgust**, **fear**, **joy**, **neutral**, **sadness**, and **surprise**—and applied a pre-trained RoBERTa model to predict these labels for each comment. The cleaned comments were tokenized and passed through the model in batches, where probabilities were computed and the most likely emotion was selected using softmax and argmax.

The model was applied separately to both Biden and Trump datasets. The resulting emotion labels were stored and later used for comparative analysis. Frequency distributions of each emotion were normalized and visualized to identify patterns in emotional tone across the two groups. This approach enables a nuanced comparison of public sentiment toward each political figure.

Results

The emotion distribution in YouTube comments reveals notable differences between Biden and Trump audiences. For **Biden**, the most dominant emotion is **neutral** (44.6%), followed by **joy** (17.3%) and **surprise** (14.9%). **Anger** appears in 9.1% of the comments, while **sadness** (6.8%), **disgust** (4.2%), and **fear** (3.1%) are present in smaller proportions.

In **Trump** comments, **joy** is the most prevalent emotion at **42.3%**, surpassing **neutral** at **32.6%**. This indicates a higher frequency of overtly positive expressions. **Surprise** (7.7%) and **anger** (7.1%) follow, with **sadness** (6.9%), **fear** (2.4%), and **disgust** (1.1%) comprising the smallest shares. (fig 3)

These results suggest that while Biden's comment section leans toward neutrality and moderate positive tone, Trump's audience exhibits a more emotionally charged and positively skewed response, especially in expressions of joy.

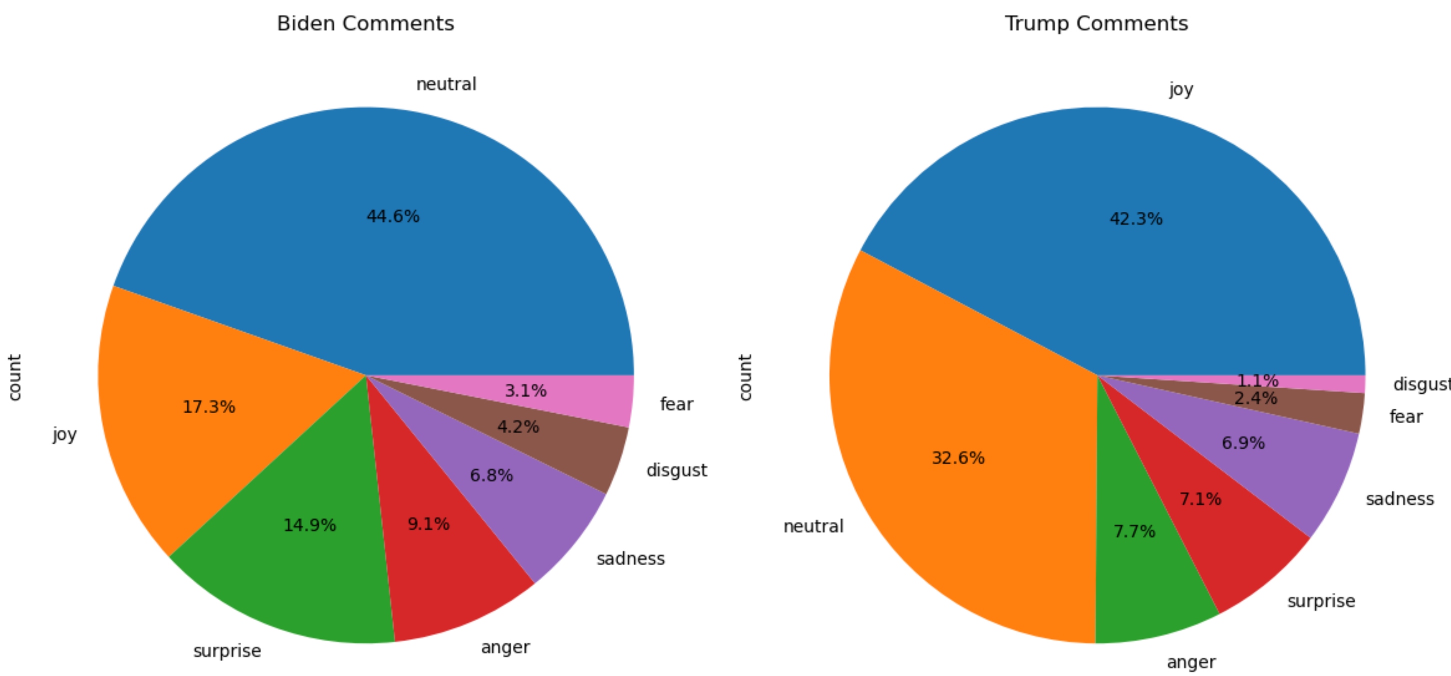

Fig 3.

These observations are further supported by the bar chart (fig 4), which offers a direct visual comparison of emotional proportions between the two candidate groups. The stark contrast in *joy* and *neutral* categories is particularly evident, reinforcing the conclusion that Trump-related discourse evokes more overtly positive sentiment, while Biden's comment sections reflect a more subdued and balanced emotional profile. Notably, the relative presence of *surprise* and *anger* in

Biden comments, compared to their lower frequencies in Trump’s, may indicate divergent audience engagement patterns or expectations.

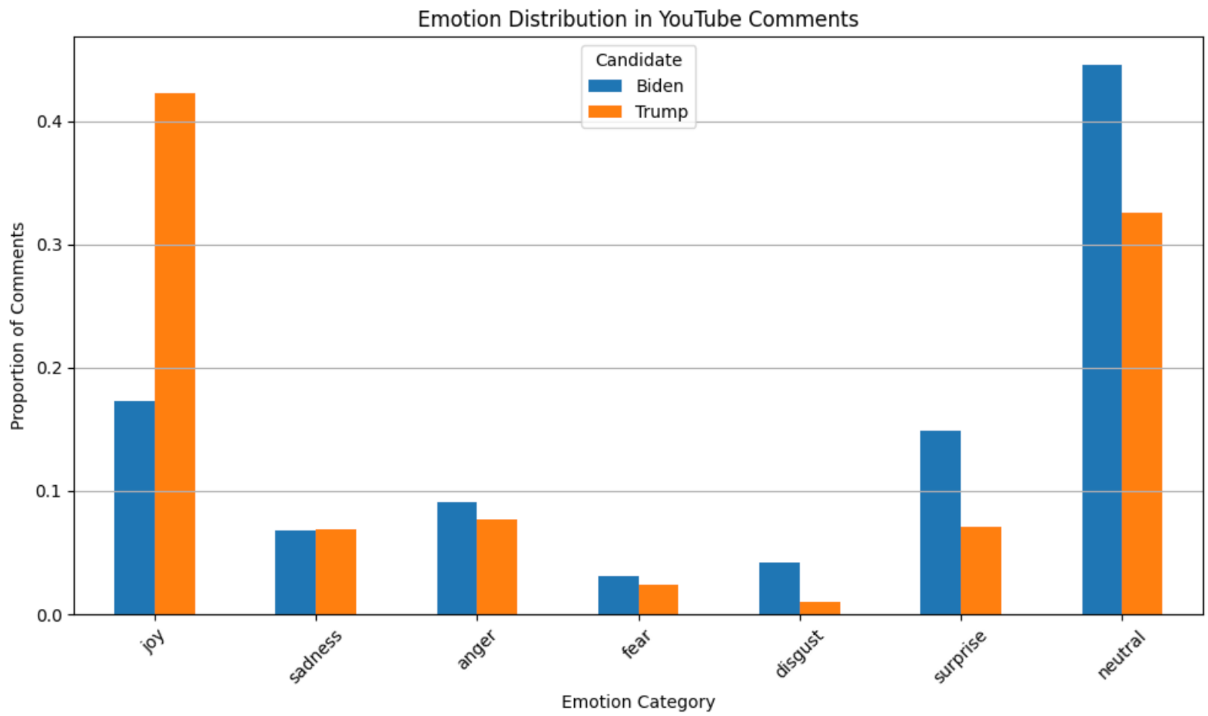

Fig. 4

The word clouds further reflect these emotional tones. Biden's comments featured expressions of affection and support like *"te amo"*, *"love"*, and *"president Biden"*. Trump’s comments included words such as *"God bless"*, *"vote"*, and *"country"*, indicating strong nationalist and religious rhetoric. (fig. 5)

These findings suggest that while both comment sections are largely neutral, Trump’s audience shows higher positive affect, whereas Biden’s comment space is more emotionally varied.

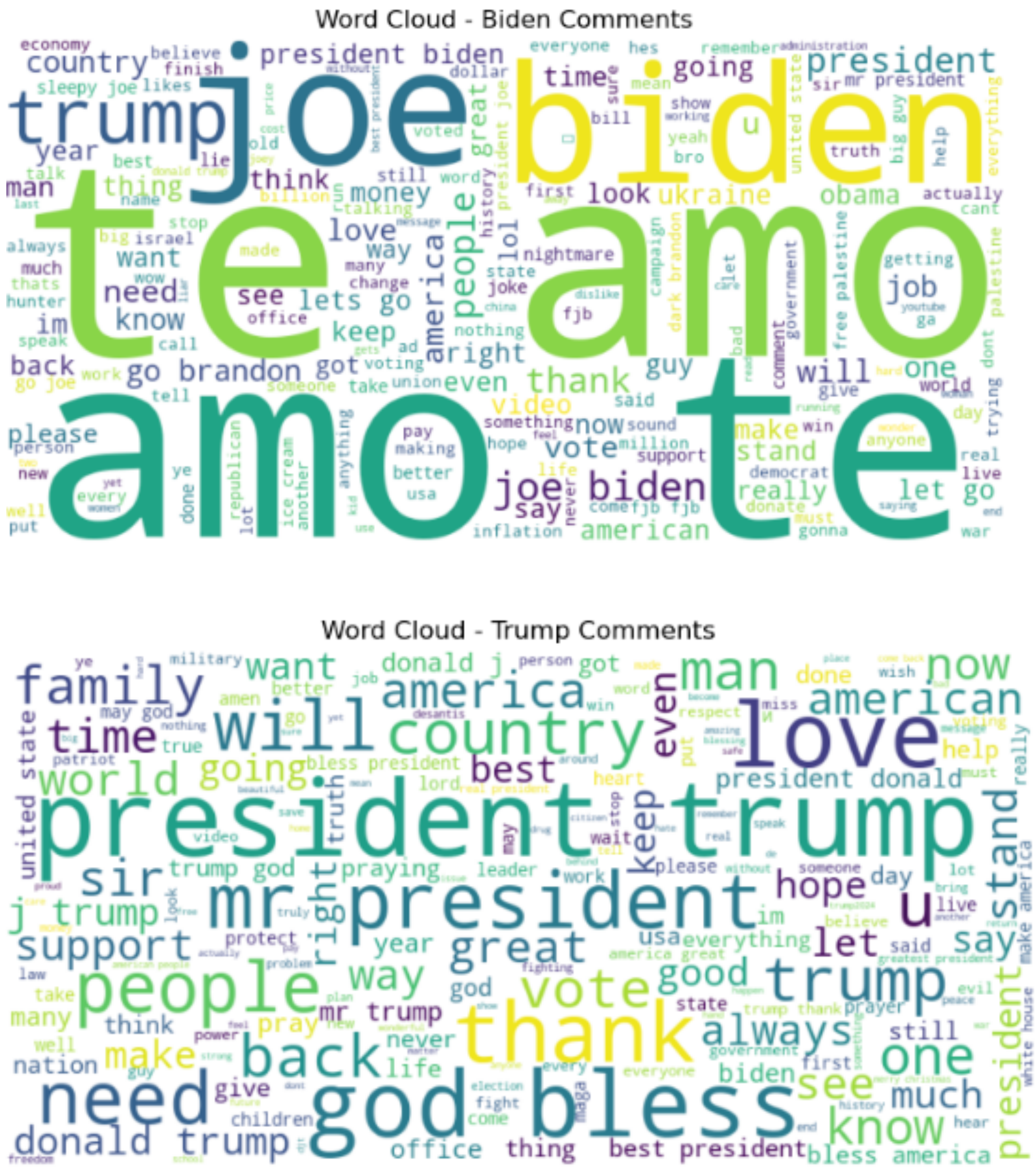
Word Cloud - Biden Comments
te amo
amo te
joe
biden
trump
Word Cloud - Trump Comments
president trump
mr president
god bless
thank
love
people
country
will

Fig. 5

Discussions and Conclusion

The aim of this study was to determine whether Large Language Models (LLMs) could reliably identify emotional patterns in YouTube comments about two well-known political figures, Donald Trump and Joe Biden. We separated the over 36,000 comments into seven groups according to our emotions using the "j-hartmann/emotion-english-roberta-large" model. We then examined our results of each candidate. The results suggest that although LLMs can spot general emotional patterns, not all applicants show the same emotional reaction. In practice, these responses might differ significantly both within and between candidate groups.

Most of the time, the comments on both candidates' pages were either neutral or positive. But 42.3% of the comments on Trump's website were good, while only 17.3% of the comments on Biden's page were “good”. Biden's views were more diverse and fair. This means that things other than the information itself, like culture, context, and ideology, can change how emotionally involved people are online.

On the other hand, this strategy has several serious issues. The method doesn't take into consideration cultural background, sarcasm, political jargon, or the wider picture of a conversation, all of which could make emotional cues less evident or change them. It doesn't do that; instead, it looks at each comment on its own. It's also tricky to establish generalizations from the sample data because the rules for moderation, the video, and the individuals who watch it all affect YouTube comments.

These results demonstrate that it's harder to employ LLMs in political discussions because they can only pick up on basic emotional patterns. They don't know enough about politics and society to really get the nuances of political action online. Reading literature alone can make it hard to interpret metaphor, irony, and ideological signaling.

Future research could look at more than one piece of information, such the tone of voice, the mood of photographs, or even how comment threads change over time, to learn more about how individuals show their thoughts in different settings. We could also learn more about how feelings change over time by seeing how real-life political events like scandals, debates, and policy announcements change how people act online.

In short, LLMs like RoBERTa can help us uncover emotional patterns in political content online, but they don't give us the whole picture because they don't contain enough context and detail

References

- Agrawal, Y., Shanker, R. G. R., & Alluri, V. (2021). Transformer-based approach towards music emotion recognition from lyrics. In *European Conference on Information Retrieval* (pp. 167–175). Springer.

- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. NAACL.

- Hartmann, J. (2022). *Emotion English RoBERTa Large*. Hugging Face. https://huggingface.co/j-hartmann/emotion-english-roberta-large

- Jiménez-Preciado, P. et al. (2025). *The Power of Words from the 2024 U.S. Presidential Debates*. Computational Political Science Review.

- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach*. arXiv preprint arXiv:1907.11692.

- Mohammad, S. M., & Turney, P. D. (2013). *Crowdsourcing a word–emotion association lexicon*. Computational Intelligence, 29(3), 436-465.

- Shevtsov, A., Oikonomidou, M., Antonakaki, D., & Ioannidis, S. (2021). *Emotion rhetoric and political polarization in U.S. elections*. arXiv preprint arXiv:2106.09447.

- Staiano, J., & Guerini, M. (2014). *DepecheMood: A Lexicon for Emotion Analysis from Crowd-Annotated News*. ACL.

- Wisnubroto, I., & Çılgın, B. (2023). *Emotion Analysis on YouTube Comments for Turkish Elections*. International Journal of Computational Social Science.

- Zhang, L., Li, C., & Zhao, Y. (2023). *Hybrid RoBERTa-LSTM model for social media sentiment analysis*. Journal of Intelligent & Fuzzy Systems, 45(2), 1373–1382.