

Politics of emotions or propaganda?

Timur Rezepov 34177A

Master's Degree Course in Data Science for Economics.

Natural Language Processing course, A.Y. 2024-2025.

Prof. Alfio Ferrara.

Dott. Sergio Picascia, Dott.ssa Elisabetta Rocchetti.

1 Introduction

This project explores how emotional language is used in political texts to influence perception and manipulate audience response. The data analyzed in the project is the transcription of US 2020 presidential debates available on Kaggle¹.

Emotion classification models based on pre-trained BERT architectures have been shown to outperform other approaches (Hsu and Ku, 2018²). To obtain emotion annotations, the DistilBERT Classifier³ was fine-tuned using some techniques and methods discussed in ‘GoEmotions: A Dataset of Fine-Grained Emotions’⁴.

AI Usage Disclaimer

Parts of this projects have been developed with the assistance of Generative AI tools. Specifically it was used to:

- Perplexity AI as an exploration tool for fine-tuning practices
- ChatGPT to refine some parts of the report text

2 Research question and methodology

Emotional language is a powerful tool in political speeches, used to persuade, mobilize, and connect with audiences on a deeper level. For example, words evoking fear or anger create a sense of urgency, while disapproval and guilt can create negative impressions of opponent's views.

¹US 2020 presidential debates:

<https://www.kaggle.com/datasets/headsortails/us-election-2020-presidential-debates>

²<https://doi.org/10.18653/v1/W18-3505>

³https://keras.io/keras_hub/api/models/distil_bert/distil_bert_text_classifier/

⁴<https://arxiv.org/abs/2005.00547>

This project aims to analyze the emotional language used by Donald Trump and Joe Biden in their speeches during the 2020 U.S. Presidential Election. Precisely, each candidate uses distinct communication strategies tailored to their target audiences, seeking to evoke specific emotional responses and strengthen their position.

The approach used comprises of building emotion profiles: the quantitative characteristic of how often different emotions(sentiments) used in the speech. Specifically the analysis focuses on:

- emotion flow throughout a single event;
- emotion profile during a specific debate event;
- comparison of emotional profiles across different debate events;
- emotion distribution by topic.

As part of the project, an emotion classification model was developed. The model is based on the pre-trained DistilBERT architecture, fine-tuned using the GoEmotions dataset.

3 Experimental results

3.1 Model training

The GoEmotions dataset consists of 58,000 Reddit comments, each labeled with one or more of 27 emotions or marked as Neutral, making it a multi-label classification task. Exploratory data analysis showed that the dataset is imbalanced. To address this, severely underrepresented classes were excluded, and only those emotions relevant to political discourse were retained—resulting in a set of 19 emotion classes.

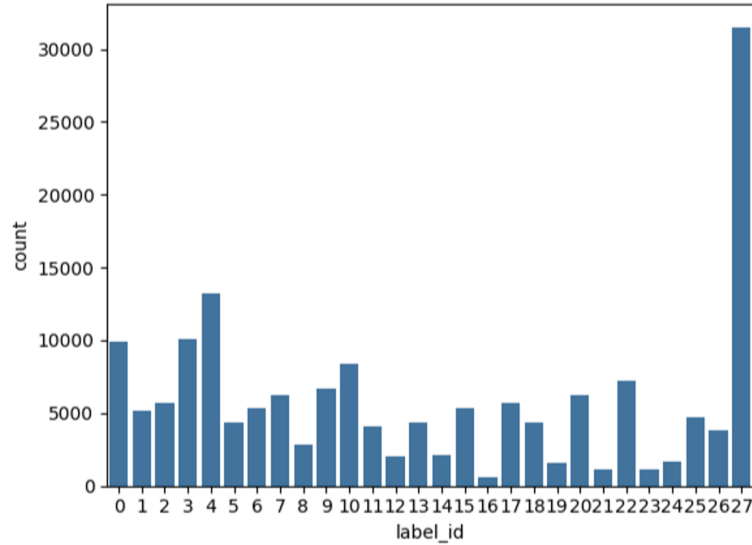


Fig. 1 Original class distribution in GoEmotions dataset

The model training parameters and performance metrics are summarised in Table 1. The resulting accuracy can be considered reasonable given the complexity of the problem:

- multi-classification problem with 19 classes;
- imbalance of the dataset;
- emotional tone is subjective and each text can be classified differently by different experts.

The classification confusion matrix is shown in Figure 2.

Table 1 Model training parameters and performance metrics

Optimizer	Adam
Learning Rate	1e-5
Number of epochs	2
Accuracy	0.29

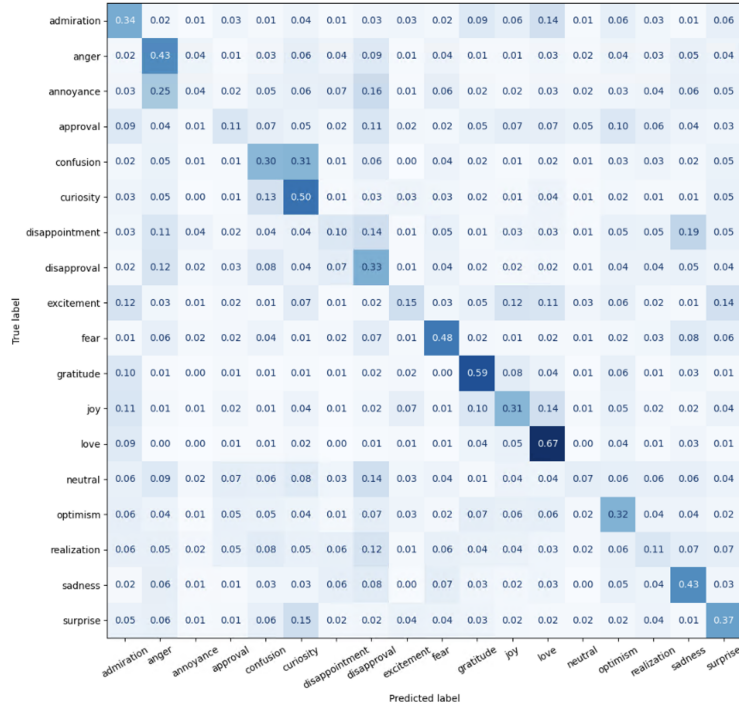


Fig. 2 Confusion matrix of the classifier

3.2 Analysis of emotional language

This article presents the key insights, while the more detailed analysis is available in the accompanying Jupyter Notebook within the repository.

3.2.1 Emotion flow (single event)

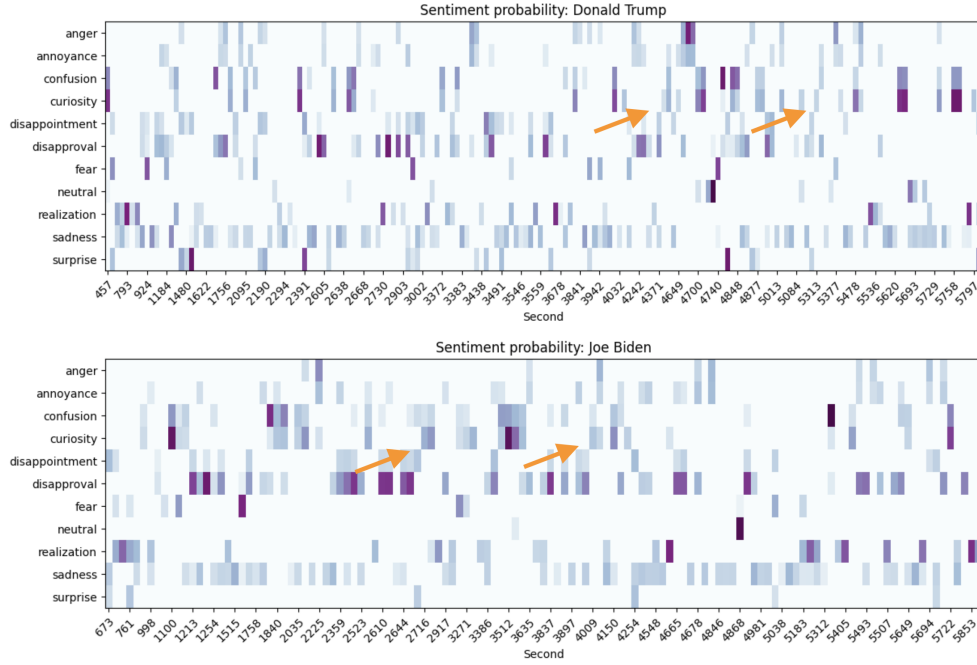


Fig. 3 Emotion (sentiment) flow

Figure 3 demonstrates probabilities assigned by the classifier to speech during the 2nd presidential debates. The emotional language of Donald Trump is more disperse and variable. Many emotions are short-lived and alternate. Biden's language is more clustered: there are **series** of emotions. Also there are less **anger** spikes. This suggests the usage of more reflective, moral tone, while Trump prefers to stay unpredictable and provokable.

In the speech of both an interesting pattern can be noted: the consecutive chain of negative emotions **disapproval** - **disappointment** - **curiosity** - **confusion** - **annoyance** highlighted with orange arrows.

3.2.2 Emotion profile (single event)

Figure 4 demonstrates emotion profile: share of each emotion expressed during the event. Both speakers display a mostly similar profile. For instance, they use **approval**

a lot to sound more constructive. Trump's profile is more variable, he expresses **anger**, **admiration** and **surprise** suggesting a more expressive manner. The usage of **confusion** and **curiosity** is high: Trump uses questions to provoke the oponent, who, on the other hand, defended himself with **disapproval** and **annoyance**.

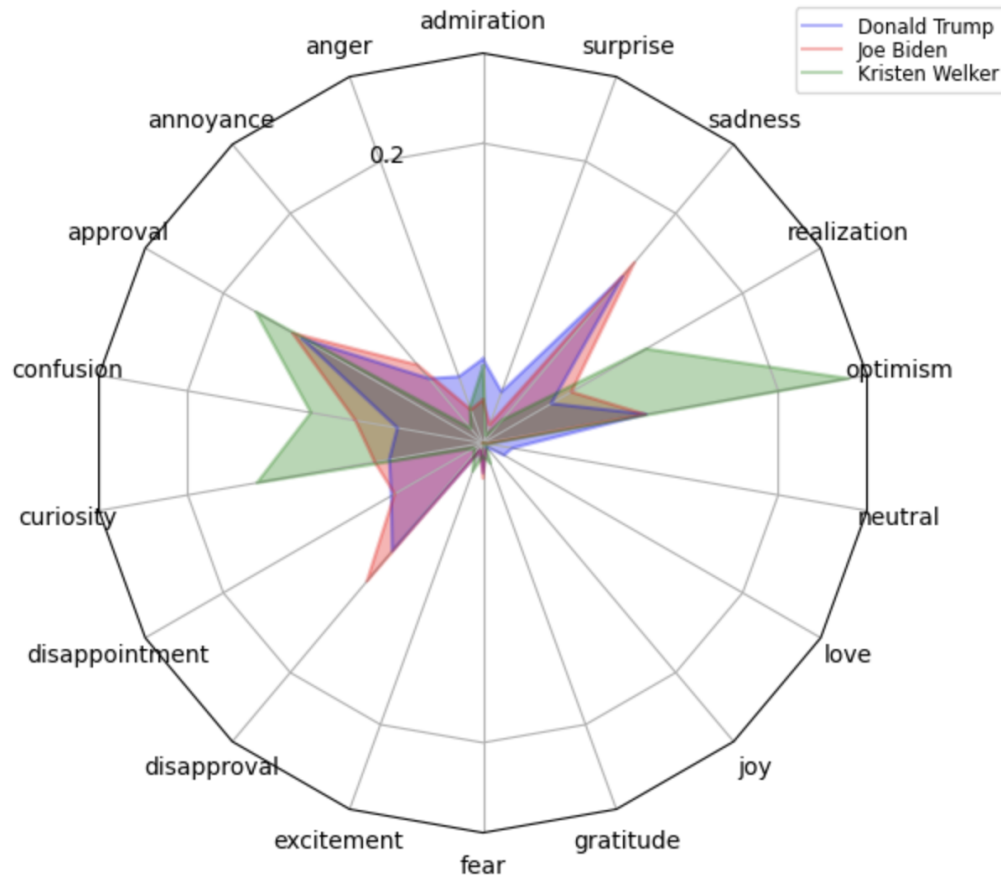


Fig. 4 Emotion profile, 2nd presidential debate

3.2.3 Emotion profiles across debates

Comparison of emotion profiles across two presidential debates is shown in Figure 5. Both speakers demonstrate the same emotional shift: they express **optimism** and **sadness** more. **sadness** - due to the fact that they discussed COVID and it's impact, while **optimism** was used to balance it and form a more hopeful, forward-looking message. Also they are less neutral, the second debate is less emotionally flat.

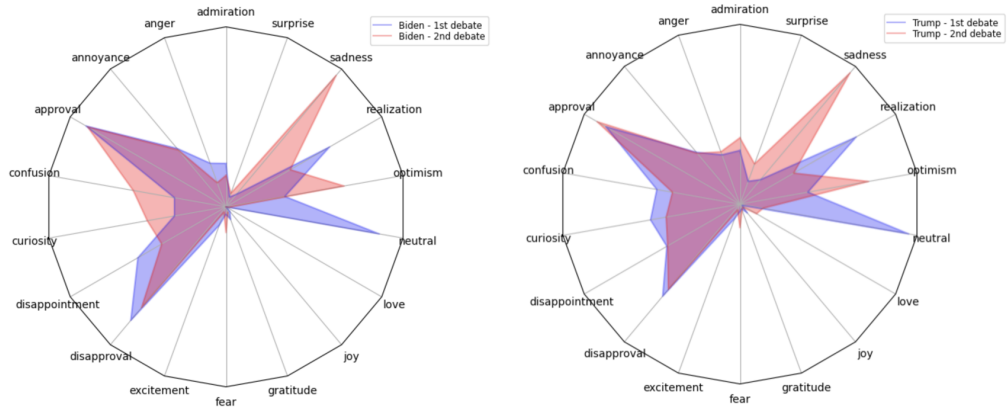


Fig. 5 Debate comparison

3.2.4 Topic modeling

Topic modeling was used to analyze how emotional language varied across different topics. The topic markup was performed using BERTopic⁵.

A clear example of emotional framing can be observed in the speakers' statements on the topic of Fracture and the Environment (Figure 6).

⁵<https://maartengr.github.io/BERTopic/index.html>

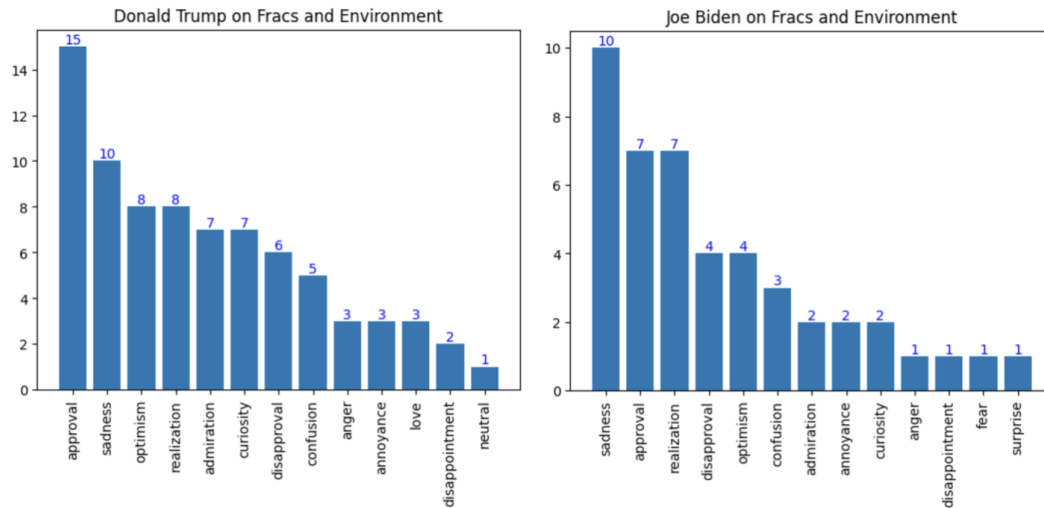


Fig. 6 Emotion language used during discussions of Fracture and Environment

Biden expresses emotionally charged narrative: his speech has high levels of realization (personal awakening), sadness (for communities suffering cancer) and optimism (regulatory actions proposals). Biden uses personal memory (story about his childhood) to form emotional appeal and mobilize support.

Figure 7 presents the SHAP⁶ explainer's output for one of Biden's statements, which is rich in words and phrases evoking the emotion of **realization**.

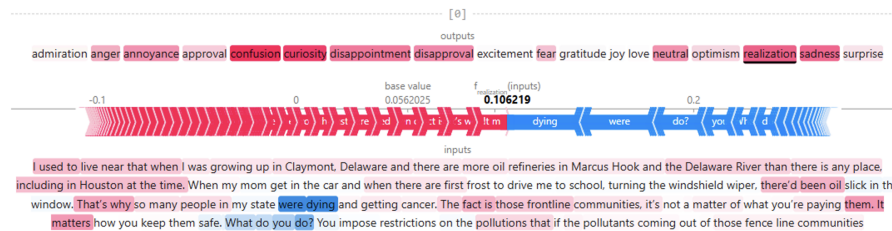


Fig. 7 Emotion language used during discussions of Fracture and Environment

Trump's leading emotion is approval. He wants to justify his policies and evoke pride. His use of admiration and optimism suggests a strategic narrative of success and progress, often appealing to economic wellbeing.

Figure 8 shows the example of Trump using **admiration** in his speech.

⁶<https://shap.readthedocs.io>

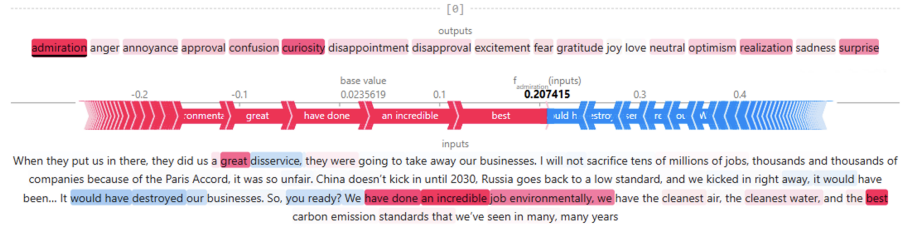


Fig. 8 Emotion language used during discussions of Fracture and Environment

3.3 Conclusion

The used model and methodology offer a foundation for emotion and sentiment analysis tasks. The analysis demonstrated that emotional language is used broadly in political speeches. In the events studied, both Biden and Trump exhibited distinct emotional profiles: Biden's language tended to be more empathetic and thoughtful, while Trump's style appeared more provocative, expressive, and self-assured.

Future work could explore the following directions:

- Improving classifier performance
- Analyzing emotion distribution over longer timeframes (e.g., across multiple years)
- Constructing emotion network graphs or conducting turn-by-turn analyses to examine how one speaker's statements influence the emotional tone of the other
- Investigating co-occurrence patterns among emotions
- Assessing the impact of emotional language on audience responses, such as through comment analysis or post-event surveys