

Emotional Rhetoric: A Sentiment Analysis in Modern Argentine Presidencies

Franco Reinaldo Bonifacini (41540A)
Università degli Studi di Milano

July 10, 2025

1 Introduction

This project explores the emotional content in the speeches of Argentina’s last three presidents: **Mauricio Macri (center-right wing)**, **Alberto Fernández (center-left wing)**, and **Javier Milei (extreme right wing)**. Each of these leaders represents a very different political ideology, and their way of speaking to the public reflects that. By analyzing the **emotions expressed in their speeches**, this project aims to understand how their communication styles differ, and how they may use emotion (consciously or not) as a tool to influence or persuade citizens.

Emotions play a key role in political discourse. Sentiment analysis, a method that uses natural language processing to detect emotions in text, has become an important tool in understanding how political messages are shaped. In this project, we apply sentiment analysis techniques to presidential speeches, looking for **patterns in how each leader expresses emotions like joy, anger, sadness, or interest**.

2 Research Question and Methodology

2.1 Research Question and Motivation

The main motivation behind this project is to analyze the last three presidents of Argentina (Mauricio Macri, Alberto Fernández, and Javier Milei) from an emotional perspective. Argentina is a **country historically marked by political instability, corruption, and one of the most cyclical economies in the world**. Based on both general observation and personal experience, many Argentine politicians seem to shape their speeches according to the political context and public sentiment, often prioritizing electoral gain over long-term solutions.

This project seeks to understand whether emotional language is present in the speeches of these presidents, and if so, whether it reflects an effort to align with public opinion or to trully represent their believes for imprpvoing the country. In this context, the central research question is: **Do the last three Argentine presidents express emotions in their public speeches in a way that supports the idea that they adjust their message to the political context and popular opinion?**

2.2 Methodology

For the implementation of the project, different models were used to analyze the presidential speeches. One of the first technical steps in this project was the translation of the original spanish-language speeches into english. For this task, it was used the **Helsinki-NLP (opus-mt-es-en)** model. It was chosen for its ability to preserve contextual meaning, which is crucial for later stages such as emotion and topic analysis. Its performance has been validated in various multilingual NLP applications, making it a dependable option for this type of pre-processing.

For the detection of emotions in text, I relied on **SamLowe’s RoBERTa-based** model, a model trained on Google’s GoEmotions dataset. This dataset features **over 28 emotion labels** derived from Reddit comments, offering a rich and diverse emotional taxonomy. The model stands out for

its widespread adoption in the NLP community, supported by its frequent usage and transparent performance reporting. While it performs well in detecting common emotions such as amusement, anger, or sadness, it shows some limitations with subtler emotions like realization or relief. To simplify analysis and improve interpretability, the emotions were grouped into broader categories, rather than working with all original labels.

Finally, for topic classification, two models were tested:

- **BERTopic_Wikipedia model:** a pre-trained instance of BERTopic trained on a large set of Wikipedia articles. Although this model was reliable, the topics outputted were not too clear and useful for further analysis.
- **roberta-base_topic_classification_nyt_news model:** trained on over 250,000 news headlines from The New York Times. Its focus on current and varied topics made it a suitable tool for identifying major themes in the presidential speeches. Compared to other available models, this one demonstrated better balance in accuracy and topic diversity, likely due to the quality and consistency of the training data.

Since the local machine available did not have access to an NVIDIA GPU, all models were executed using **Google Colab Pro**, which provided access to cloud-based GPU acceleration (**Tesla T4**). This significantly reduced processing time compared to CPU-only execution and allowed for the efficient handling of large language models and batch inference.

3 Methodological Framework

3.1 Data Collection and Preparation

The initial dataset was based on a **scraping script** available on GitHub [2], which included **presidential speeches from the Casa Rosada website up to October 2024**. This code was then used to scrape additional speeches to complete the dataset up to 2025.

Some speeches could not be retrieved automatically. In those cases, the texts were manually copied from the official Casa Rosada website [1].

To process the data, the **title_content_split** function from VC Studio was used to separate the titles from the speech content. Then, the **cleaned_content function** was applied to remove statements made by other participants, keeping only the parts spoken by the President of Argentina.

Due to the varied structure of the speeches, unlike more predictable formats like interviews or monologues, a manual review was necessary. This helped ensure that the extracted content **accurately reflected only presidential speech**, despite inconsistencies in format.

Last but not least, very short speeches, defined as those **with three or fewer sentences**, were excluded from the dataset because they were considered insufficient for the goals of the project. Additionally, **outliers were removed based on the rule of values exceeding the Upper bound ($Q3 + 1.5 * IQR$) in terms of text length**. An exception was made for speeches by President Javier Milei, since the limited number of available cases (less than two full years in office) would have significantly reduced the dataset. The final number of speeches per president was as follows:

| President | Count |
|-------------------|-------|
| Mauricio Macri | 619 |
| Alberto Fernández | 592 |
| Javier Milei | 13 |

3.2 Speech Segmentation

Continuing with the preparation of the speeches, each text was segmented into **chunks of no more than 80 words**. When **punctuation marks** were present, the text was split accordingly; otherwise, it was cut after reaching the 80-word limit. The threshold of 80 words was selected to **preserve the contextual meaning of longer speech passages**, which are common in presidential speeches. Shorter segments could risk losing semantic value relevant to the analysis.

The function **split_speeches** was used for this task, combining sentence tokenization with word-based splitting. Additionally, very short segments (fewer than 5 words) were merged with the previous chunk to maintain coherence.

The segmented data was then transformed into a flat structure using the **explode()** function in pandas, resulting in a DataFrame where each row represents a single speech segment, along with its corresponding president and date.

3.3 Translation (Spanish to English)

To prepare the data for emotion classification using English-language models, all speech segments were translated from Spanish to English. This was done using the **MarianTokenizer** and **MarianMTModel** classes from the transformers library, specifically the **"Helsinki-NLP/opus-mt-es-en"** model.

The **translate** function was implemented to process a list of Spanish text inputs. The function tokenizes the input and generates the corresponding English translations using a transformer-based approach.

Initial tests using example sentences from President Javier Milei showed that the translations were sufficiently accurate for the purpose of this project. For instance:

| Original | Translation |
|-----------------------------|----------------------------------|
| "Hola a todos." | "Hello, everyone." |
| "Yo soy el león." | "I'm the lion." |
| "Yo también los amo." | "I love you too." |
| "Viva la libertad, carajo." | "Long live the fucking freedom." |

To ensure smooth performance and avoid crashes in Google Colab, the dataset was divided into **blocks of 50 sentences**. Each block was translated separately, which allowed for reliable execution without exceeding memory limits.

3.4 Emotion Classification

To detect emotions in each speech segment, the model **"SamLowe/roberta-base-go_emotions"** was used through the transformers library's pipeline interface. The classification was performed using a text classification pipeline, which returns the full list of predicted emotions along with their confidence scores.

To efficiently process the dataset, a the **batch_emotion** function was implemented to classify emotions in batches of 50 text segments at a time. This helped optimize runtime and reduce memory usage when working in Google Colab.

Once predictions were obtained, **top_emotion** function was used to extract the top emotion per segment (the emotion label with the highest confidence score). This allowed for a simpler interpretation and analysis of results.

The original model outputs **27 distinct emotion labels**. To make the results more interpretable and suitable for analysis, these emotions were grouped into six broader categories:

| Group Emotion | Emotions |
|---------------|---|
| Joy | gratitude, optimism, approval, joy, admiration, amusement, excitement, pride, relief, love, caring. |
| Sadness | sadness, remorse, embarrassment, confusion, disappointment. |
| Anger | anger, annoyance, disgust, disapproval. |
| Fear | fear, nervousness, surprise |
| Interest | curiosity, realization, desire. |
| Neutral | neutral. |

This grouping helped simplify the emotional landscape while preserving the richness of the original classification.

3.5 Topic Classification

To identify the main topics in each speech segment, several approaches were tested.

The first approach used was **BERTopic**, a well-known method for topic modeling based on transformer embeddings and clustering. However, in this case, the generated topics were not suitable for analysis. Many topic labels were unclear or overly technical, such as:

- 941_bourgeoisie_capitalist_bourgeois_capitalism.
- 414_motivation_motivations_motivational_motivate.

This made interpretation and grouping difficult.

After discarding BERTopic, several pretrained topic classification models from **Hugging Face** were explored. One promising option was **valurank/distilroberta-topic-classification**, but it had very few downloads, which introduced uncertainty about its reliability [3].

Eventually, the model **dstefa/roberta-base_topic_classification_nyt_news** [4] was selected. It is based on RoBERTa and trained on New York Times news data. While not specifically trained on political speech, it provided clear and generalizable topic labels, which were appropriate for this analysis.

The **assign_roberta_topic_labels** function was implemented to classify the speech segments in batches of 50, returning both the predicted topic label and its associated confidence score. These values were added directly to the dataset for further analysis.

3.6 Explainability (SHAP)

To better understand how the emotion classification model makes decisions, **SHAP (SHapley Additive exPlanations)** was used to analyze a subset of predictions. SHAP is a model-agnostic framework that explains machine learning outputs by assigning an importance value to each input feature. These values, called Shapley values, come from cooperative game theory and measure the individual contribution of each feature to a prediction [7].

Initially, **attention-based analysis** was considered to interpret model behavior. However, this approach **presents limitations**. Although attention weights highlight tokens the model “focuses” on, they do not necessarily indicate true importance or causality [5]. In practice, attention scores may emphasize irrelevant words or ignore meaningful ones, making interpretation less reliable.

SHAP was selected instead, as it provides **more robust, interpretable explanations**. It allowed to see which specific words or phrases contributed most to each predicted emotion. This method was found to be **easier to implement and more insightful** for this project’s goals than attention heatmaps.

The **shap_explanation** function was implemented to apply SHAP to the sentences in the dataset, displaying both the top five predicted emotions and a visual explanation of the most influential tokens. This method helps validate the model’s reasoning and ensures that emotionally charged or meaningful words are indeed driving the predictions.

4 Experimental Results

4.1 Context

To better understand the tone, content, and emotional patterns in each speech, it is important to consider the political and historical context in which each president governed. Argentina has gone through significant changes in political leadership and ideology over the last decade. Each administration faced different national and global challenges, which shaped the way they communicated with the public.

Mauricio Macri (2015–2019), became president after 12 years of Kirchnerist governments (center-left). Representing the coalition Cambiemos, which united parties across center-right and centrist ideologies, his election marked a shift in Argentine politics toward a more market-friendly and business-oriented approach. Despite initial optimism, his administration struggled to control **inflation**, which remained a central economic issue. The failure to stabilize the economy and meet expectations eventually affected his popularity and became one of the defining criticisms of his presidency.

Alberto Fernández (2019–2023), won the election with the support of former president Cristina Fernández de Kirchner, who became his vice president. This alliance marked the return of Kirchnerism to power. His presidency began during the **COVID-19 pandemic**, and Argentina imposed one of the **longest lockdowns** in the world. While the early response was cautious and got most of the population’s approval, over time it deeply impacted businesses, workers, and the economy. Public trust declined due to several controversies, such as holding private gatherings during strict lockdowns, offering vaccines to political allies before priority groups, and a general refusal to acknowledge policy failures. In public speeches, Fernández often maintained an optimistic tone, even when economic and social indicators were worsening.

Javier Milei (2023–present), is a disruptive political figure who broke away from traditional party structures. He won the presidency with La Libertad Avanza, an extreme-right libertarian movement, after gaining massive popularity through media appearances and a direct, confrontational communication style. His speeches often include provocative remarks and are characterized by strong anti-establishment rhetoric. He frequently ends public addresses with the phrase “**¡Viva la libertad, carajo!**” (“**Long live the fucking freedom.**”). His administration introduced deep economic reforms aimed at reducing state spending and inflation. While some macroeconomic indicators have improved, the population has yet to feel relief in everyday life. His confrontations with political opponents, public figures, and even foreign leaders have placed him at the center of several controversies. Nonetheless, he maintains strong support from segments of the population that view him as a necessary break from the past.

4.2 Results: Emotions and Explainability

The emotion classification results provide a general overview of how the three most recent Argentine presidents express emotions in their public speeches:

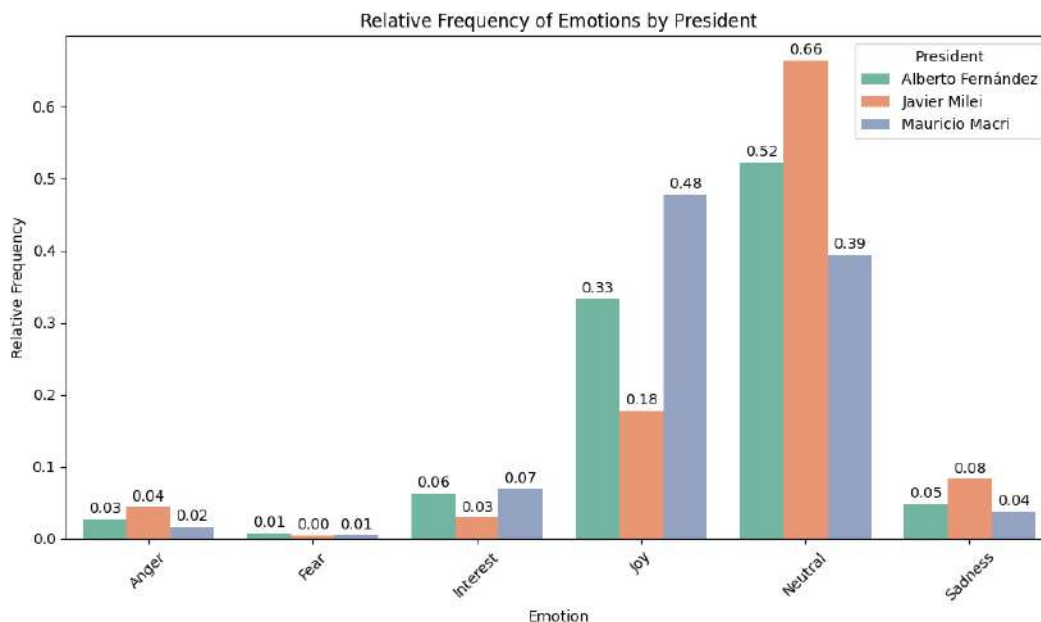


Figure 1: Frequency of Emotion by President

One of the first observations is the high frequency of the **neutral emotion** across all presidencies. This suggests a tendency toward institutional or politically correct communication, possibly to keep a formal tone or avoid polarization.

When looking at the emotional distribution per president:

- **Mauricio Macri’s** speeches show the highest frequency of **joy** (0.48), followed by **neutral** (0.39). Less frequent but still present are **interest** (0.07), **sadness** (0.04), and **anger** (0.02).

- **Alberto Fernández** presents a similar emotional pattern, with **neutral** (0.52) and **joy** (0.33) leading, followed by **interest** (0.06), **sadness** (0.05), and **anger** (0.03).
- **Javier Milei** displays a different profile. While **neutral** (0.66) remains dominant followed by **joy** (0.18), there is a noticeably higher presence of **sadness** (0.08) and **anger** (0.04), with interest appearing less frequently.

The emotion fear appeared with extremely low frequency across all speeches and was therefore excluded from further analysis.

These emotional patterns align with the political context and rhetorical style of each leader. **Macri and Fernández**, representing more traditional political approaches (center-right and center-left, respectively), tend to emphasize **joy and interest**, possibly in an attempt transform negative realities (such as inflation, economic crises, or the COVID-19 pandemic) in a more hopeful or optimistic light. Their speeches suggest a focus on persuasion and perception management, consistent with the Argentine tradition of political optimism in discourse. In contrast, **Milei’s** discourse, rooted in an anti-establishment and confrontational style, shows a higher frequency of **anger and sadness**, particularly during his campaign year. These emotions support his strategy of highlighting the country’s decline and blaming past governments, particularly the political class, for Argentina’s long-standing crises.

Interestingly, word clouds reinforce these patterns, which are shown in Figure 11 for Macri, Figure 12 for Fernández, and Figure 13 for Milei, in the appendix. In speeches expressing joy, **Macri and Milei** frequently use terms associated with **openness and global integration**, such as “world” and “freedom,” reflecting their liberal economic views. **Fernández**, instead, uses language centered on **national production and labor**, consistent with his protectionist politics. For **anger**, Macri and Fernández use more conventional, generalized language, while **Milei’s rhetoric is openly confrontational**, with words like “politicians,” “wrong,” or “fucking freedom.” This highlights his break from traditional political communication. In the case of **sadness**, each leader’s vocabulary reflects their main obstacles or crises—whether inherited or caused during their own administration: **inflation (Macri), the pandemic (Fernández), and economic collapse (Milei)**. Regarding interest, key words like “change” (Macri), “realized” (Fernández), and “economy/market” (Milei) reflect their core political messages.

A clear example of **Fernández using joy strategically** appears when he expresses **admiration** (grouped under joy) toward Sergio Massa, then serving as his economy minister. Despite Massa leading to an annual inflation of **211% in 2023**, with monthly peaks of 25%, SHAP highlights the word “best” as key in describing him as the optimal choice in the following elections, revealing a contrast between rhetoric and economic reality.

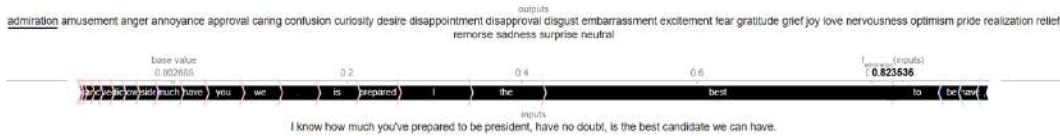


Figure 2: SHAP results for Fernández (Joy)

In addition, for **sadness** in **Fernández’s** discourse, the SHAP analysis highlights emotions of **disappointment and grief**, particularly in a sentence where he blames Macri for the condition in which he received the country, suggesting that this deterioration facilitated the impact of the **COVID-19 pandemic**. SHAP emphasizes terms like “inherit,” “very,” and “damaged,” illustrating an attempt to shift responsibility onto the previous administration rather than fully assuming accountability for managing the crisis.



Figure 3: SHAP results for Fernández (Sadness)

Emotion also evolved across time. **Macri and Fernández** both show a decline in **joy** throughout their mandates, reflecting the growing public dissatisfaction as economic and social challenges accumulated. Toward the end of their terms, **interest and anger** increase, suggesting attempts to re-engage public opinion or redirect blame.

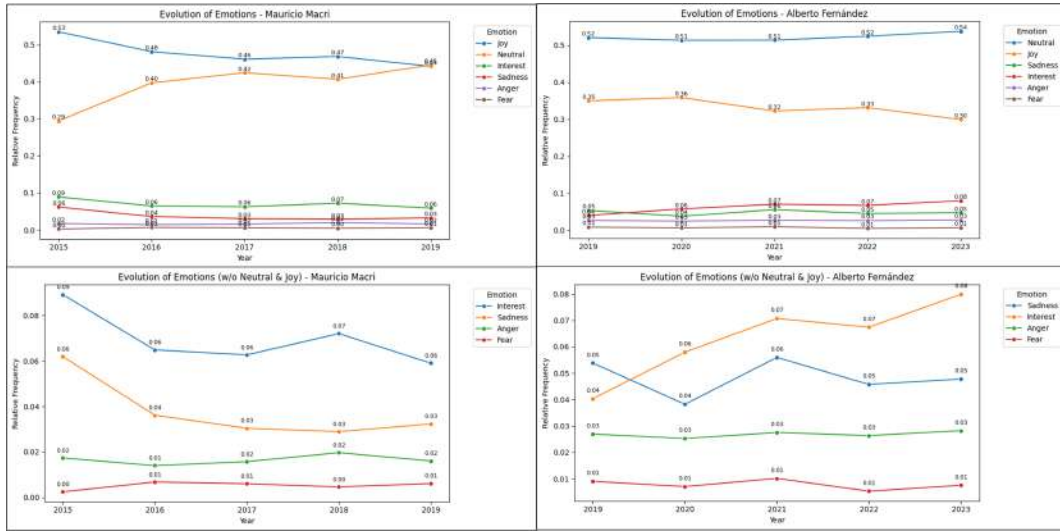


Figure 4: Evolution of Macri and Fernández

A **SHAP-based analysis** further illustrates **Macri's emotional discourse**, particularly around **anger and interest**. In a case of **anger** (disapproval), he criticized the previous government for hiding economic data, with SHAP highlighting negative terms like "think," "don't," and "But" as key drivers. For **interest** (realization), a representative example is his emphasis on Argentina's reintegration into the global economy during his presidency, supported by SHAP values on terms like "realized," reinforcing his message of openness and progress.



Figure 5: SHAP results for Macri

For **Milei**, early trends suggest a shift from **emotionally charged campaign discourse** to a

more neutral and optimistic tone in office. While **sadness** dominated his campaign narrative, mainly used to emphasize Argentina’s collapse, his presidential tone incorporates more **joy and neutrality**, likely reflecting efforts to project confidence in his economic policies and early signs of stabilization.

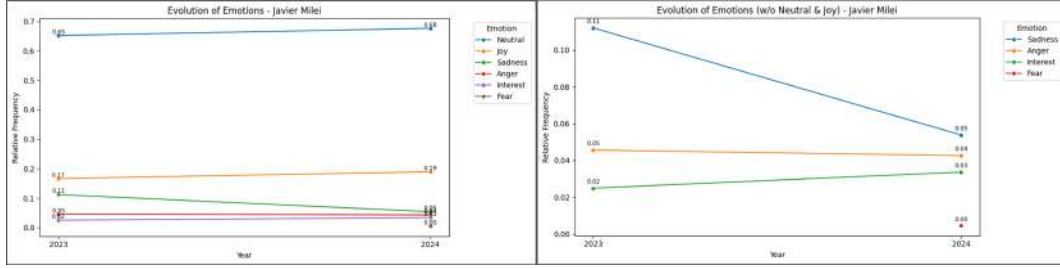


Figure 6: Evolution of Milei

When focusing on the explainability of some examples from **Milei**, the SHAP results proved to be accurate and aligned with his ideological stance. Regarding joy, a clear example is Milei expressing **admiration** for businessmen, reflecting his alignment with his party’s ideology. The SHAP values emphasized words such as “businessman” and “hero”, highlighting the president’s admiration toward them. On the other hand, in the case of **sadness**, a fitting example is a speech where he describes the state in which he received the country, primarily criticizing the model implemented by the previous ruling parties. The key word reinforcing this emotion was “heartbreaking”, which highlights his sentiments about the national situatio, feelings aligned with the perception shared by the majority of the population.

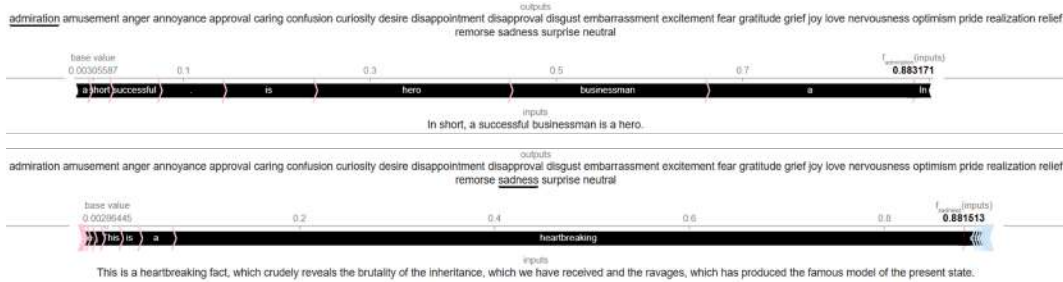


Figure 7: SHAP results for Milei

4.3 Results: Emotions by Topic

Finally, an analysis of emotions by topic reveals how each president’s discourse reflects both their ideology and the context of their mandate. While common themes like **Business and Finance**, **Health and Wellness**, **Crime**, and **Sports** appear across all presidencies, the emotional tone and focus differ significantly.

Under **Mauricio Macri** (2015–2019), **Business and Finance** dominated, aligned with his market-oriented agenda. Initially framed with **joy and interest**, the tone shifted toward **anger and sadness** as economic instability and inflation worsened. This culminated in moments like his angry speech after the 2019 PASO elections, where he blamed voters. In **Sports**, peaks of **interest and anger** followed Argentina’s Copa América loss in 2015, and **sadness** rose after the 2018 World Cup elimination.

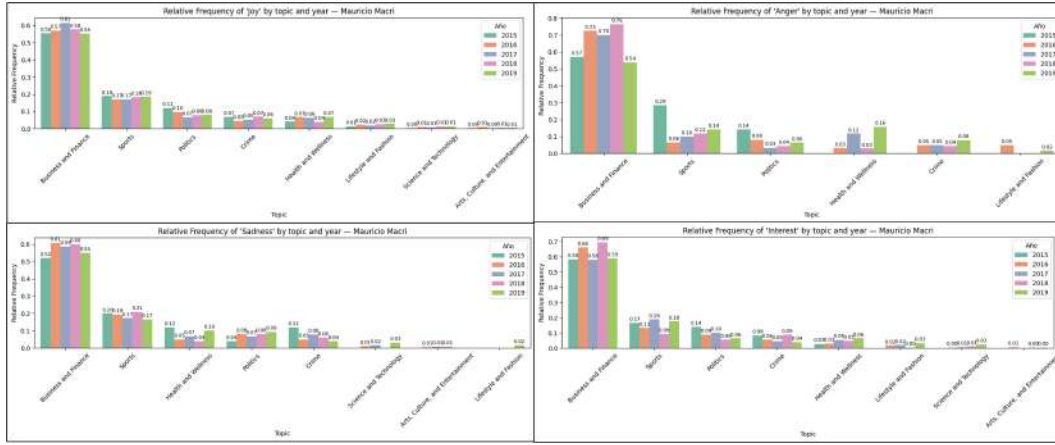


Figure 8: Emotions by Topic - Macri

Alberto Fernández (2019–2023) governed during the **COVID-19 pandemic**, which made **Health and Wellness** a central topic, with increasing **anger** and **sadness** as public support diminished due to extended quarantines and political scandals. A slight rebound of **joy** and **interest** near the end of his term may reflect efforts to project optimism. In **Business and Finance**, high levels of **sadness** in 2019, even before he took office, suggest rhetorical framing to blame Macri and justify future shortcomings. Unlike his predecessors, Fernández remained largely disengaged from **Sports**, even during the 2022 World Cup, reflecting a deliberate distancing between the government and the national team.

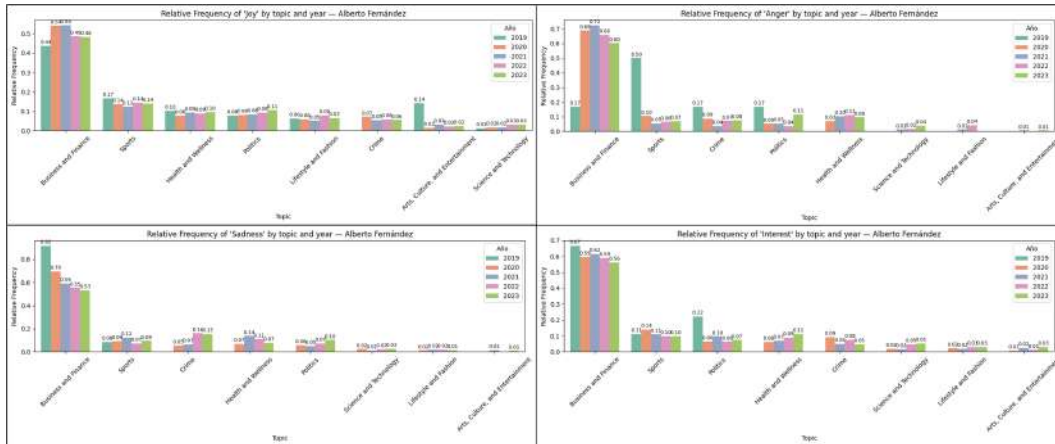


Figure 9: Emotions by Topic - Fernández

Javier Milei (2023–present) introduced a radically different tone. His discourse, represented by extreme-right and libertarian views, is emotionally polarizing. In **Business and Finance**, high **anger** and **sadness** are already present during the 2023 campaign, as he denounced prior administrations and described the country as devastated. In contrast, **Crime** is linked to **joy**, especially following Patricia Bullrich’s appointment as Security Minister and early anti-crime successes. Milei’s language often includes aggressive terms, like “fucking freedom”, reflecting his anti-establishment strategy and resonating with voters seeking drastic change.

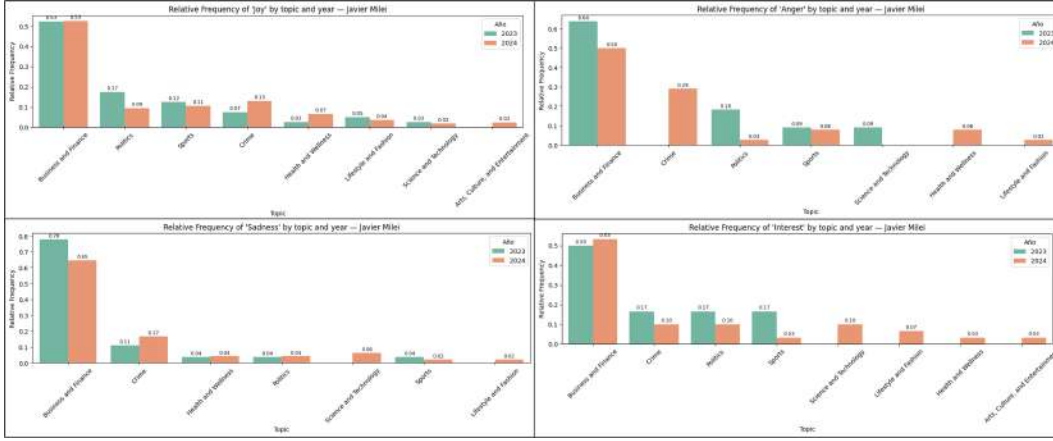


Figure 10: Emotions by Topic - Milei

5 Conclusion

This project demonstrated the value of emotion analysis and explainability for understanding presidential speeches in Argentina. By examining the emotional tone of presidential speeches, both in general and by topic, it was possible to identify distinct rhetorical patterns aligned with each president’s ideology, context, and communication strategy.

However, the study had some limitations. A significant challenge was the inconsistency of the speech formats. Many did not follow a standardized structure and required substantial manual preprocessing. Furthermore, some sentences included routine or polite phrases such as “Good Morning” or “Thank you” that, while irrelevant to political content, may have influenced the emotion classifier’s output.

Despite these obstacles, the analysis outputted meaningful insights. Both **Macri** and **Fernández** exhibited a traditional Argentine political behavior: using emotional rhetoric—especially **joy** and **interest**, to describe their own agendas positively while attributing crises to previous administrations. These presidents showed a higher frequency of emotions in general, often emphasizing optimism or assigning blame rather than acknowledging failures. In contrast, **Milei** adopted a more confrontational yet strategically neutral tone. While he also blamed former leaders, his discourse rapidly shifted, within the same year of assuming, from emotionally charged campaign rhetoric to a more composed and goal-oriented communication style.

In conclusion, combining emotion classification, topic modeling, and explainability techniques provided a valuable framework to decode the underlying strategies and shifts in presidential speeches, revealing not only the political ideology of each leader but also the emotional tools used to influence public perception.

AI Usage Disclaimer

Parts of this project were developed with the assistance of OpenAI's ChatGPT (GPT-4). The AI was used to support idea generation and provide guidance in structuring content and workflows. All AI-assisted suggestions were reviewed and adapted by me. I take full responsibility for the final content, ensuring its accuracy, relevance, and academic integrity.

References

- [1] Presidencia de la Nación Argentina. *Sitio oficial de la Casa Rosada*, (n.d.), <https://www.casariosada.gob.ar/informacion/discursos>
- [2] Agustín Benassi. "*scraper-casa-rosada*". GitHub, 25/10/2024, <https://github.com/abenassi/scraper-casa-rosada/blob/master/scraper.py>
- [3] Valurank. *distilroberta-topic-classification*, Hugging Face, (n.d.), huggingface.co/valurank
- [4] dstefa. *roberta-base_topic_classification_nyt_news*, Hugging Face, (n.d.), https://huggingface.co/dstefa/roberta-base_topic_classification_nyt_news
- [5] Hila Chefer, Shir Gur, Lior Wolf. *Transformer Interpretability Beyond Attention Visualization*, The School of Computer Science, Tel Aviv University, (n.d.).
- [6] Michelle M. H. Şeref, Onur Şeref, Alan S. Abrahams, Shawndra B. Hill, Quinn Warnick. *Rhetoric Mining: A New Text-Analytics Approach for Quantifying Persuasion*, INFORMS Journal on Data Science, (2023), <https://doi.org/10.1287/ijds.2022.0024>.
- [7] Scott M. Lundberg, Su-In Lee. *A Unified Approach to Interpreting Model Predictions*, (n.d.).

A Word Clouds



Figure 11: Mauricio Macri's Word Cloud



Figure 12: Alberto Fernández's Word Cloud

