Emotional Framing in Congressional Tweets During the Capitol Riot

Niccolo' Cibei

Keywords: Emotion classification, Transformer models, Sentiment analysis

1 Introduction

Political discourse often relies on emotional framing to shape public perception, mobilize support, and frame events within broader ideological narratives. Social media platforms, and Twitter in particular, provide a direct channel for political actors to express such framing in real time.

This study focuses on U.S. congressional tweets surrounding the January 6, 2021 Capitol Riot, a moment of acute political crisis. Using a transformer-based emotion classification model fine-tuned on the GoEmotions dataset, combined with explainability techniques such as attention heatmaps and SHAP token attribution, we examine how emotional tone shifted before, during, and after the event.

By analyzing both macro-level (8-day) and micro-level (58-hour) time windows, and conducting case studies of prominent Democratic and Republican leaders, the work aims to identify not only which emotions were expressed, but also the rhetorical functions they served in crisis communication.

2 Methods

We detected emotional tone using the `SamLowe/roberta-base-go_emotions` transformer, a RoBERTa-base model fine-tuned on Google's GoEmotions dataset (28 emotion labels). To improve interpretability, we mapped these labels into broader functional categories (e.g., fear, anger, sadness, joy, pride, surprise, moral/empathic, disgust, neutral), enabling clearer comparison across parties and time periods.

Each tweet was classified to obtain the top predicted emotion, probability, and grouped category. Classification was run in batches with GPU acceleration, producing outputs for all five temporal datasets.

For rhetorical interpretation, we applied token-level explainability methods: attention heatmaps from the model's final layer and SHAP (SHapley Additive exPlanations)

values to highlight words driving each prediction. These visualizations reveal which lexical choices carried emotional weight in political discourse.

3 Dataset

We used tweets from official U.S. congressional member accounts, retrieved from the public congresstweets repository[1], filtering out non-member and committee accounts. The data was cleaned to remove retweets, mentions, URLs, and artifacts, and matched with metadata (screen_name, party, chamber) from the companion user file. For analysis, we produced five datasets: two *macro-level* (4 days before and after January 6, 2021) and three *micro-level* (24h before, 10h during, and 24h after the riot). In total, the final corpus contains over 11,500 tweets, each with timestamp, author, party affiliation, and processed text ready for emotion classification and explainability analysis.

4 Results

4.1 Macro-Level Analysis (4-Day Windows)

We compared the distribution of emotions in congressional tweets from the four days preceding the Capitol Riot (January 2–5, 2021) and the four days following it (January 7–10, 2021), as shown in Figures 1–3.

Overall, the emotional landscape of congressional communication changed noticeably after January 6. While *neutral* and *joyful* tones dominated in both periods, the balance between them shifted substantially. The proportion of *neutral* tweets increased by roughly five percentage points, suggesting a move toward more restrained or factual language. Conversely, expressions of *joy* dropped sharply, by around fifteen percentage points, reflecting a collective withdrawal from positive emotional framing. At the same time, *sadness* more than tripled, rising from about three percent to over eleven percent of tweets, while smaller but notable increases appeared in categories such as *moral* and *fear*.

When broken down by party, these shifts reveal distinct rhetorical strategies. Democratic members displayed a marked increase in *neutral* framing (+8.9%) and a rise in *sadness* (+7.7%), suggesting an effort to adopt a more institutional and procedural tone in response to the crisis. Republican members, on the other hand, exhibited a stronger surge in *sadness* (+12.8%) combined with a greater use of *moral* appeals (+3.4%) and a decrease in *neutral* tone (–4.1%). This pattern indicates a more emotionally charged and value-oriented rhetoric in their post-riot communication.

Taken together, these results highlight how the two parties navigated the emotional aftermath of the event in contrasting ways: Democrats gravitated toward composure and institutional language, while Republicans leaned into moral and affective framing. The divergence underscores the strategic use of emotional tone as a political tool during moments of national crisis.

[1] https://github.com/alexlitel/congresstweets

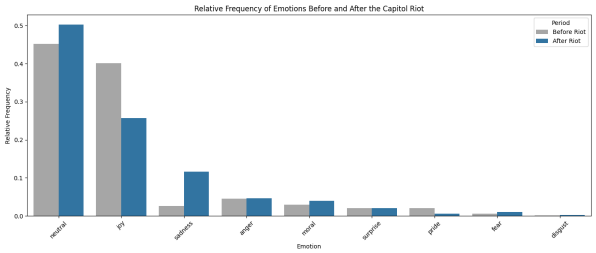


Fig. 1: Relative frequency of emotions before and after the Capitol Riot (all members).

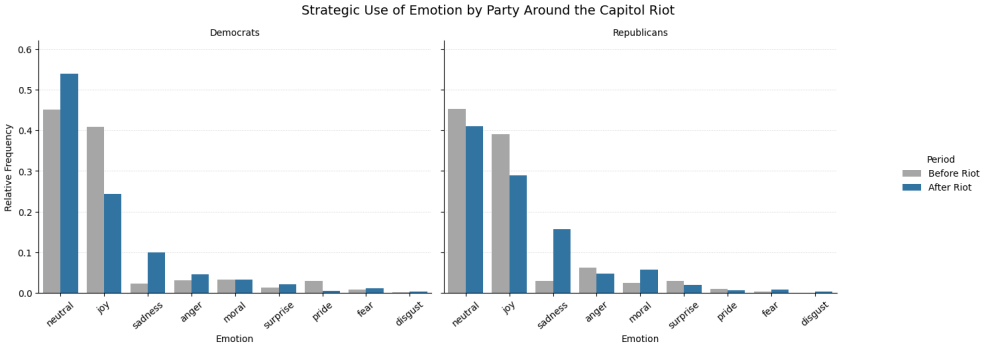


Fig. 2: Relative frequency of emotions before and after the Capitol Riot by party.

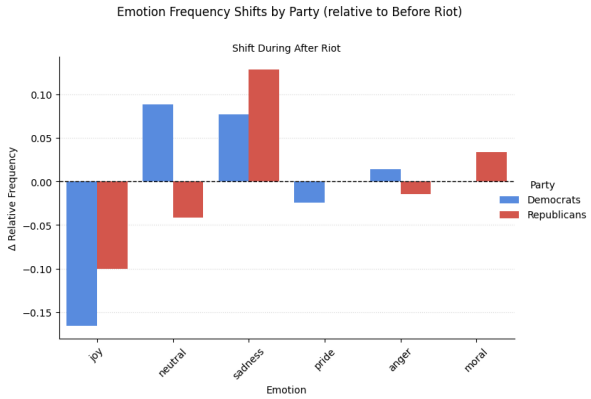


Fig. 3: Change in relative frequency (After – Before) by party.

4.2 Micro-Level Analysis (24-Hour Windows)

To capture rapid emotional dynamics surrounding the Capitol Riot, we examined three consecutive 24-hour periods: the day leading up to the riot (January 5, 19:00–January 6, 19:00 UTC), the riot itself (January 6, 19:00–January 7, 05:00 UTC), and the

immediate aftermath (January 7, 05:00–January 8, 05:00 UTC). Figures 4–6 illustrate how emotional expressions shifted across these critical hours.

Overall, the riot corresponded to a sharp disruption in the emotional tone of congressional communication. During the event, the proportion of *neutral* tweets dropped precipitously—from about 57

Party-specific trajectories reveal deeper rhetorical contrasts. Republican members exhibited the most pronounced shift during the riot, with a dramatic decline in *neutral* tone (–0.38) and substantial increases in both *anger* (+0.23) and *moral* appeals (+0.12). Democratic members also moved away from neutrality (–0.13), though their rise in emotional expression was more moderate: *moral* language increased by +0.08, *anger* and *sadness* each by +0.04. After the riot, Democrats' tone rebounded more quickly toward neutrality, whereas Republicans maintained elevated levels of *sadness* and *moral* framing, indicating a sustained emotional orientation.

These patterns suggest that the riot provoked a temporary breakdown of neutral, procedural rhetoric across both parties, replaced by emotionally charged and morally framed discourse. *Anger* and *moral* language surged as members sought to condemn the violence and position themselves within competing moral narratives. For Republicans, the stronger and more persistent emotional shift may reflect the rhetorical need to assert legitimacy or distance themselves from the perpetrators. Democrats, meanwhile, appeared to balance expressions of outrage with a faster return to institutional language, consistent with a strategy aimed at projecting stability and continuity in governance. Taken together, these findings show how, even within hours of a crisis, partisan actors diverge in how they deploy emotion and institutional tone to shape the early public framing of events.

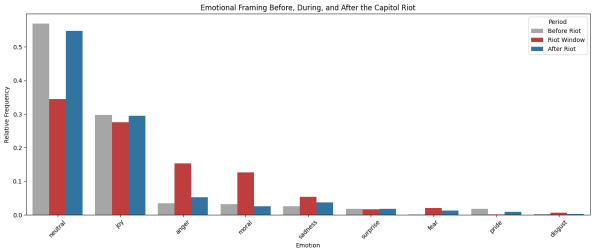


Fig. 4: Emotion distributions in the 24h before, during, and after the Capitol Riot (all members).

4.3 Pelosi vs. McCarthy

A comparative analysis of rhetorical and emotional framing was conducted in the tweets of two key congressional leaders:Nancy Pelosi (Democrat) and Kevin McCarthy (Republican). We focused on two distinct moments of January 6, 2021: the period during the riot, characterized by urgent, real-time reactions, and the hours immediately after, when discourse shifted toward reflection, accountability, and interpretation.

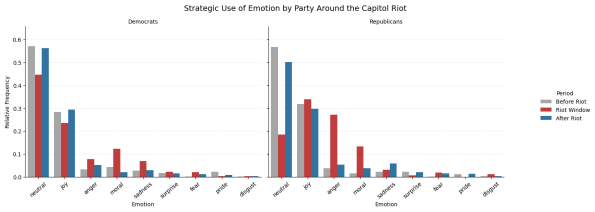


Fig. 5: Emotion distributions in the 24h windows, by party.

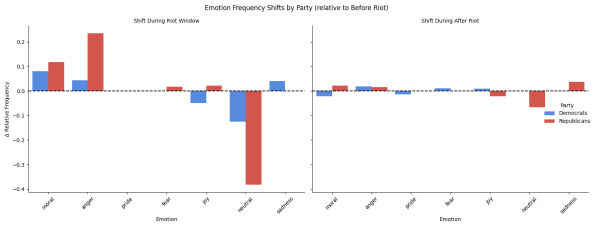


Fig. 6: Change vs. 24h-before baseline by party. Left: Riot Window; Right: After Riot.

Each tweet was analyzed using a RoBERTa-based model fine-tuned on the GoEmotions dataset to identify its primary emotional tone. To complement this emotion classification, we applied two explainability techniques designed to uncover how these emotions were linguistically expressed. First, we generated attention heatmaps from the model's final transformer layer to visualize which tokens most strongly influenced the emotion prediction. Second, we used SHAP (SHapley Additive exPlanations) values to quantify each word's contribution to the predicted emotional label, thereby providing a fine-grained view of rhetorical emphasis.

Together, these methods allow for a nuanced understanding of how emotional intensity and rhetorical focus varied between the two leaders. By combining quantitative emotion scores with interpretability visualizations, we can observe not only what emotions were conveyed, but also how specific lexical and rhetorical choices contributed to each emotional framing. The following subsections qualitatively interpret these results, highlighting key differences in tone, emotional salience, and communicative strategy between Pelosi and McCarthy at critical moments of the crisis.

4.3.1 Pelosi vs. McCarthy During the Riot

During the Capitol riot, Speaker Pelosi and Minority Leader McCarthy employed markedly different rhetorical tones, despite responding to the same unfolding crisis.

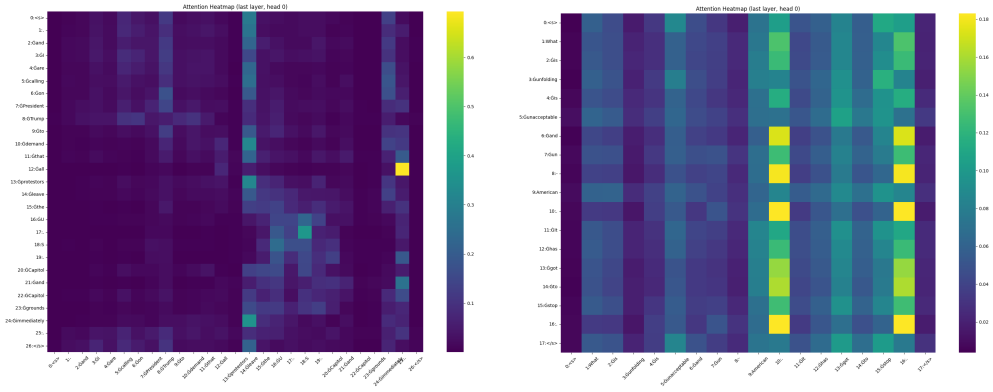
> **Pelosi:** "@SenSchumer and I are calling on President Trump to demand that all protestors leave the U.S. Capitol and Capitol grounds immediately."
>
> **McCarthy:** "What is unfolding is unacceptable and un-American. It has got to stop."

Pelosi's message was classified by the model as overwhelmingly *neutral* (0.964), with only faint traces of *approval/joy* (0.009), *anger* (0.010), and *moral* appeal (0.006). As shown in Figures 7a and 8a, the attention heatmap indicates that the model's

final-layer focus clusters around entity tokens such as "protestors", "leave", "grounds" and "immediately". This pattern suggests that the model prioritizes syntactic and semantic structure rather than affective content.

The SHAP explainability visualization further confirms that each token contributes minimally to emotional deviation. Terms like "calling" and "immediately" exert a slight influence but remain within the neutral domain. Together, these analyses demonstrate that the model interprets Pelosi's statement as a factual, procedural appeal rather than an emotional response, consistent with the dominant prediction of *neutral* (0.964).

By contrast, McCarthy's tweet reveals a distinctly emotional framing. The model classified it primarily as *anger*-related, dominated by *disapproval* (0.826) and *annoyance* (0.084), while neutral or positive emotions were marginal (*neutral* 0.034, *joy* 0.009). This rhetorical strategy situates his response within a framework of ethical judgment rather than procedural command, Figures 7b and 8b visualize this contrast. The SHAP analysis identifies "unacceptable" as the strongest positive contributor toward the *disapproval* classification, as it conveys explicit moral judgment. Meanwhile, the attention heatmap from the last transformer layer shows a clear concentration around the evaluative adjectives and the imperative "stop," indicating that the model's internal focus aligns with emotionally charged lexical items. Unlike neutral statements where attention centers on entities or syntactic dependencies, this pattern highlights how attention intensifies around affective words when the model detects emotional content. Overall, the model interprets McCarthy's utterance as an expression of moral outrage and condemnation.



(a) Pelosi attention heatmap (b) McCarthy attention heatmap

Fig. 7: Attention heatmaps during the riot for Pelosi (left) and McCarthy (right).

4.4 Pelosi vs. McCarthy After the Riot

Following the events of January 6, both Speaker Pelosi and Minority Leader McCarthy issued public statements condemning the attack on the U.S. Capitol. While both

(a) Pelosi SHAP attribution

(b) McCarthy SHAP attribution

Fig. 8: SHAP token attributions during the riot for Pelosi (left) and McCarthy (right).

messages articulated disapproval, their rhetorical strategies diverged significantly in tone, structure, and emotional framing.

> **Pelosi:** "Yesterday, American democracy came under attack — but we refuse to be bullied into abandoning our duty to work #ForThePeople."
>
> **McCarthy:** "The violent mob responsible for this attack on the U.S. Capitol deserves the full consequences of their actions under the law. The FBI is seeking to identify individuals instigating violence in Washington, D.C. We are accepting tips and digital media depicting rioting or violence in and around the U.S. Capitol on January 6. If you have information, visit."

Pelosi's post-riot statement is classified by the model primarily as *disapproval* (0.353), with secondary traces of *neutrality* and *sadness*. According to the SHAP explanation (Figure 10a), the phrase *we refuse* emerges as the dominant emotional driver, conveying moral resistance and principled rejection.

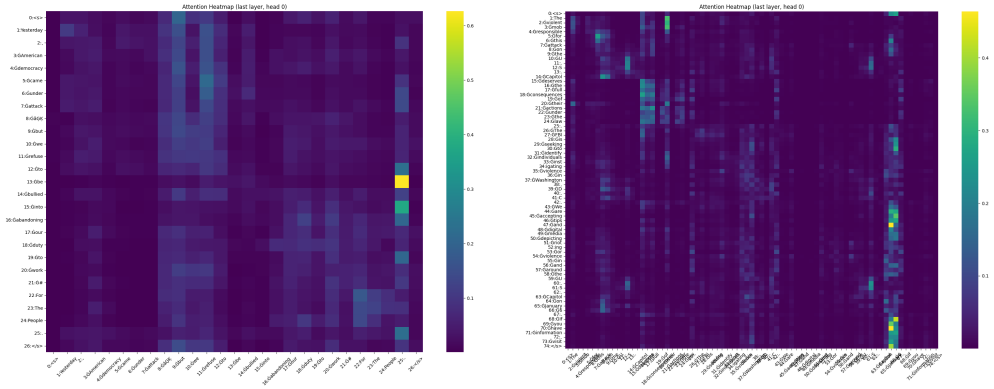
The attention heatmap (Figure 9a) shows moderate internal focus among the tokens "but", "we", "refuse" and "to" suggesting the model has captured both syntactic and semantic cohesion across these terms. Increased attention at the sentence ending (e.g., "#ForThePeople.") reflects standard final-layer aggregation behavior, but does not shift the model's emotion prediction. Overall, the model interprets Pelosi's statement as a firm expression of disapproval rooted in agency and institutional resolve, rather than overt emotionality.

In contrast, McCarthy's post-riot message is classified as overwhelmingly *neutral* (0.953), with only faint traces of evaluative tone. As shown in Figure 10b, the term *deserves* introduces a slight moral judgment, subtly lowering the neutrality score. However, most tokens, including those that might appear emotionally charged, such as *violent mob*, *attack*, and *riot*, are embedded within a legal and procedural framing. As a result, their emotional contribution is minimal.

The attention heatmap (Figure 9b) shows that final-layer concentrates on a small set of factual anchors. Slightly brights vertical bands occur at the first mention of *Capitol* and at *Washington*; and most strongly at the late-sequence date tokens *January* and *6*. By contrast, columns at ostensibly emotive words such as *violent*, *mob*, and *attack* are relatively dim and diffuse. Together with the SHAP result, this suggests that the model's final-layer representation is grounded in entities, place, and time rather than affective vocabulary, yielding an overall *neutral* classification.

5 Conclusions

This study examined the emotional framing of U.S. congressional tweets during and after the January 6, 2021, Capitol Riot using transformer-based emotion classification



(a) Pelosi — attention heatmap (b) McCarthy — attention heatmap

Fig. 9: Attention heatmaps after the riot for Pelosi (left) and McCarthy (right).

(a) Pelosi — SHAP attribution (b) McCarthy — SHAP attribution

Fig. 10: SHAP token attributions after the riot for Pelosi (left) and McCarthy (right).

combined with token-level explainability. By analyzing both multi-day and hourly intervals, the approach captured both gradual trends and rapid shifts in rhetorical tone.

At the broader temporal scale, both parties moved away from celebratory and partisan language toward more restrained and somber communication, marked by increases in sadness, moral appeals, and neutral statements. Democrats exhibited a stronger shift toward neutral, procedural language intended to project institutional stability, while Republicans displayed greater increases in moral appeals and anger, particularly at the height of the riot. Although anger declined the following day, it did not fully return to pre-riot levels.

The case studies of Speaker Pelosi and Minority Leader McCarthy reflected these broader partisan patterns. During the riot, Pelosi emphasized procedural authority and urgency, adopting a tone of institutional control, whereas McCarthy expressed moral condemnation and emotionally charged criticism. In the aftermath, Pelosi's rhetoric combined disapproval with reaffirmations of democratic resilience, while McCarthy's response centered on law, justice, and accountability within a factual and institutional register.

Overall, this work demonstrates how integrating transformer-based emotion detection with interpretability tools such as attention visualization and SHAP attributions can move beyond coarse sentiment analysis to reveal the rhetorical functions of

political language. The framework presented here can be extended to other political crises, media ecosystems, or cross-national contexts to illuminate how emotional framing—ranging from institutional neutrality to overt moral appeal—shapes public discourse, political alignment, and collective memory.

Author's Declaration on Use of AI Tools

The ChatGPT-4o language model (OpenAI) assisted in parts of the analysis, drafting, and code refinement. Its outputs, used for summarization and wording suggestions, were reviewed and edited by the author, who retains full responsibility for all content and conclusions.