The Evolution of Online Discourse During the Israel-Palestine Conflict: A Reddit Analysis

Federico Annoni, Luca D'Ambrosio, Filippo De Min, Caterina Ruzzi, Federico Ungarelli

Abstract

The Israeli-Palestinian conflict is one of the longest-running territorial conflicts still ongoing in the world. Tensions in the area date back to the first half of the 20^{th} century with the migrations of Jews from Europe and Yemen to Palestine and the first clashes between the two groups after Israel's declaration of statehood. This culminated in the First Intifada and the massive protests that erupted in the Gaza Strip, the West Bank, and Israel among Palestinian-Israelis. On October 7^{th}, a violent attack by Hamas triggered the start of a harrowing conflict that persists to the present day. As the war evolves, controversies and polarization seem to increase among online supporters of each faction. To our knowledge, no previous research studied the linguistic patterns across supporters of the two groups. In this paper, we attempt to uncover underlying structures in the discussion about the conflict through the analysis of Reddit textual data, exploiting topic modeling techniques and large language models to detect the evolution of emotional responses and the differences in prevailing themes between pro-Israel and pro-Palestine content. We find that the pro-Palestine community is more focused on the conflict compared to the Pro-Israel community, which expresses a wider range of topics and concerns. Additionally, throughout the time span of the conflict, an overarching theme of increased resentment and decreased joy emerges in both groups. Lastly, we believe that our research manages to shed light on the community's perception of the conflict, and highlights the main concerns of the supporters of each faction.

1. Introduction

The October 7^{th} attack by Hamas on Israeli civilians provoked a violent response from the Israeli army that decided to invade the Gaza Strip. After 6 months of military intervention, the Israeli forces have taken roughly 35,000 lives in Gaza (Statista (2024)). Our study investigates the reactions of online communities to the recent evolution of the Israel-Palestine conflict. By leveraging 10 months of Reddit data with over 14,000 posts for 3 communities, we explore the evolution of emotional trends and recurring themes within each group.

Our analysis begins with training a classifier using the Israel and Palestine subreddits as labeled datasets. This classifier achieves an 88.6% accuracy score and a 70% F1 score on the evaluation set. Leveraging this tool, we label partisanship on the r/IsraelPalestine subreddit, which hosts both political perspectives.

To quantify emotional changes over time and per community, we output the mean softmax score produced by a DistilRoBERTa fine-tuned on emotions prediction. The study focuses on the evolution of emotions in reaction to major war events, distinguishing between conflict-escalating and mitigating events. Our results show that conflict-escalating events, like the attacks on October 7^{th}, led to a substantial increase in resentment (from 30% to 45%) and a decrease in joy (from 25% to 17%), while fear levels remained volatile. Conversely, mitigation events such as the ceasefire on November 24^{th} resulted in decreased resentment and increased joy. These patterns were consistent across different communities, suggesting a general preference for peace. The findings also show increased polarization in social media posts since the start of the conflict. In addition, our study utilizes a BERT-based topic modeling approach to uncover recurring discussion themes within the two communities, complementing our analysis of emotion. Our model effectively tracks the evolution of topic frequency over time, revealing peaks aligned with major war events and corresponding shifts in topic representation. Eventually, through t-SNE dimensionality reduction, we identify distinct clusters for each faction, facilitating comparison between the two communities. This analysis reveals two distinct clusters, one for each faction. our findings suggest that pro-Israel Reddit content covers a broader range of topics, while the pro-Palestine content is more concentrated, primarily focusing on the conflict.

2. Data Collection

To conduct our analysis, we referred to the Academic Torrents platform and downloaded every monthly Reddit dump from July 2023 to April 2024. The collected datasets contain the majority of Reddit submissions posted in their respective month, totaling an average of 20 GB of data per dump. We then proceeded to filter the data based on our subreddits of interest: r/Israel, r/Palestine, and r/IsraelPalestine. This operation was made possible by an ad-hoc Python script retrieved from the GitHub repository of the user who posted the datasets. After filtering out deleted posts, removed posts, and posts that included an image/poll (no textual data), our final dataset comprises 14,039 total submissions posted across the three subreddits, spanning 10 months. As reported in Table 1, a significant portion of posts was removed by moderators, indi-

Subreddit	Type	Number Of Posts
r/Israel	[deleted]	52
	[removed]	4044
r/IsraelPalestine	[deleted]	22
	[removed]	9273
r/Palestine	[deleted]	15
	[removed]	8773

Table 1: Number of removed and deleted posts per Subreddit

cating a prevalence of offensive or inappropriate content within these subreddits. This behavior is to be expected, given the subject's sensitivity. A full count of monthly posts is reported in Figure 1.

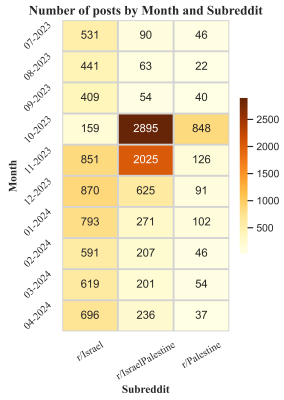

Figure 1: Heatmap of subreddit activity by month

3. Methodology

3.1 Partisanship Classifier

To extract the maximum amount of information from the available data, we fine-tuned a roBERTa model (Conneau et al. (2019)), aimed at classifying text as either pro-Israel or pro-Palestine and leveraged it to make inference on the r/IsraelPalestine subreddit. We use posts from the r/Israel and r/Palestine subreddits to train our classifier, attributing a different label to each of them. One inherent limitation of this approach is the clear class imbalance that emerges from the data (Table 2). This concern could not be addressed due to the lack of additional labeled pro-Palestine content on the platform. Regarding the relevant model parameters, we opted for 5 training epochs with a batch size of 16 samples. The model achieved an accuracy of 88.6% on the evaluation set. However, it exhibited a clear bias towards classifying text as pro-Israel, attributable to the aforementioned class imbalance. This bias is evident from the evaluation f1-score of roughly 70%. Thus, to retain the dataset structure and attain meaningful metrics we applied stratified sampling when building our evaluation set. The final version of the model was trained on the entire dataset and out-of-sample predictions were performed for the r/IsraelPalestine subreddit. After a manual inspection of a random portion of the data, the same bias emerged to some extent, but overall, the results appeared promising.

Label	Number of Posts
Pro Israel	5960
Pro Palestine	1412

Table 2: Number of Posts per Subreddit

3.2 Emotions' Evolution

To detect emotions in Social media text, we employed a pre-trained Transformer DistilRoBERTa (Hartmann (2022)) fine-tuned to predict Ekman's 6 basic emotions (anger, disgust, fear, joy, sadness, surprise), plus a neutral class. The authors fine-tuned the model on an ensemble of six datasets, including Reddit data, cumulating 2,811 observations per label. The model has achieved 66% accuracy on the evaluation set. We decided to exclude 'surprise' due to its irrelevance to our analysis. Additionally, the 'disgust' and 'neutral' emotions are not displayed because of their low mean predicted scores over the period we considered (1% and 5%, respectively). We chose to group 'sadness' and 'anger' because both arise from perceived injustice or following tragic events. When combined, they manifest as resentment, a complex emotional response to feeling wronged; 'fear' on the other hand, has been confined in its own separate category as, contrary to sadness and anger, it arises from a sense of threat or danger. Subsequently, we proceeded to investigate the detected emotions' evolution through time and across subreddits, emphasizing their reaction to major war events (Washington Post (2024); Reuters (2024)). We distinguished between mitigation events and conflict-escalating events; the latter kind of event often entailed human losses, hence we expect an increased level of 'resentment' and 'fear'.

3.3 Topic Modeling

We leveraged the BERTopic Python library to uncover topics. BERTopic (Grootendorst (2022)) is a topic modeling technique that offers significant advantages over standard approaches like LDA. It uses BERT embeddings to capture semantic relationships. In addition, BERTopic employs a class-based TF-IDF approach, making it more robust to noisy data. Given their continuous nature, BERTopic topics are flexible and can evolve over time. For dynamic topic modeling, BERTopic splits the corpus into time slices, runs independently on each slice, and aligns topics across slices using c-TF-IDF, allowing for an effective track of topic evolution. To effectively represent the evolution of topics, we utilized the visualization features of the BERtopic package. Specifically, we plot the trends of topics over time by measuring their normalized frequency at each timestamp. Frequency is normalized to more easily compare the lines with each other.
Finally, we employed a t-SNE algorithm to reduce the dimensionality of topic embeddings to two dimensions. The perplexity parameter of the model is set to 30 to strike a balance between capturing both local and global structures in visualization.

4. Results and Discussion

4.1 Emotions' Evolution

Figure 2 displays the 7-day rolling average of the softmax score predicted by the emotion classifier, with vertical lines marking major war events. After the events of October 7^{th}, resentment levels spiked from 30% to 45% (a 50% increase),

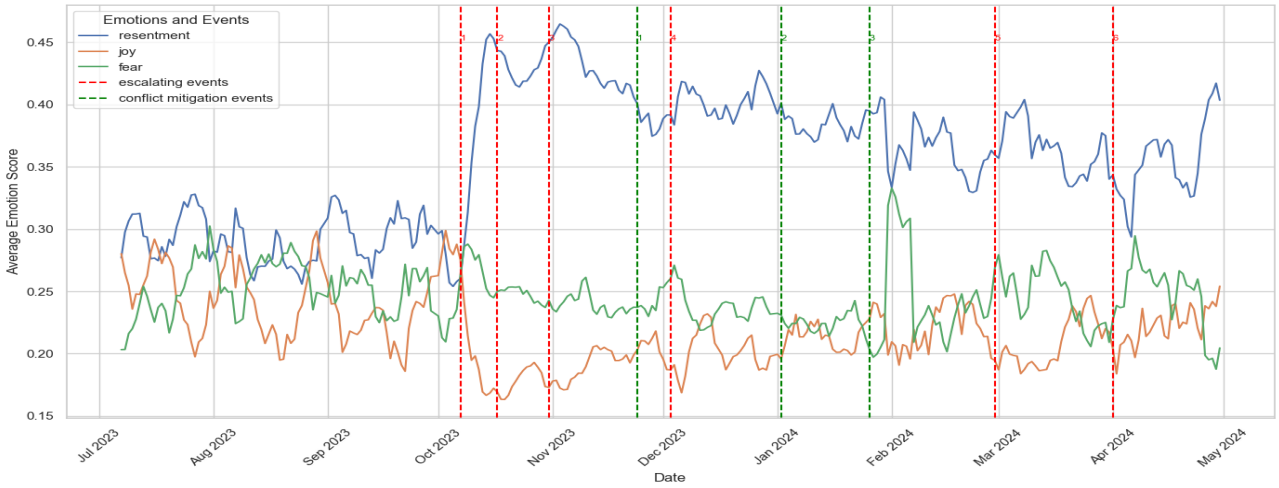

Figure 2: 7-day rolling average of Emotions for all subreddits

while joy levels drop from 25% to 17% (a 32% decrease). Fear levels, however, showed high volatility without a clear trend, suggesting either that fear was not significantly affected by the war or that the classifier struggled to distinguish between fear, sadness, and anger, leading to confounded measures.

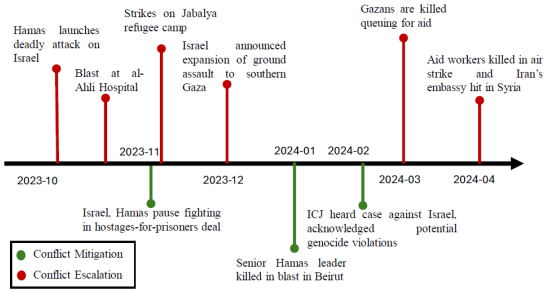

Figure 3: Timeline of Major Events

During the first month of the conflict, resentment levels remained steady around 45%, indicating that Hamas' terrorist attacks had lasting emotional impacts. The bombings in Gaza causing significant civilian casualties likely contributed to this sustained resentment peak.

The main pattern in the data shows that conflict-escalating events are followed by increases in resentment and decreases in joy for about a week. Conversely, conflict-mitigating events are associated with opposite trends. These results, consistent when analyzing all communities independently, imply that communities react negatively to war escalation regardless of the target, suggesting a preference for peace over victory.

For instance, the ceasefire and hostage deal on November 24^{th} were followed by declines in resentment and increases in joy. The International Court of Justice's (ICJ) acknowledgment of potential genocide by Israel on January 26^{th} caused a notable spike in fear, though this could be due to noise or genuine community reactions.

Although not shown in the figure, the neutral category averaged 7.5% before October 7^{th} but dropped to 5% afterward. This decline in neutral content is statistically significant at the 0.1% level suggesting strong polarization in posts' content since the beginning of the war.

It is important to note the imbalance in the number of observations, with fewer data points in the first and last month, leading to less reliable measures. Additionally, the removal of most posts from r/Israel in October 2023 may invalidate our results for that month.

4.2 Topic Modeling

Once we incorporated the previously classified posts for r/IsraelPalestine (Section 3.1) into the r/Israel and r/Palestine datasets to enhance the final corpus for analysis, we found that the final number of pro-Palestine submissions was not sufficiently sparse over time to perform an accurate trend analysis. Hence, our study, as outlined in Figure 4, presents the trends of the four main topics of discussion within the pro-Israel community from August 2023 to April 2024:

1. *Zionism and Antisemitism*: this topic presents a significant increase in normalized frequency from early October to early November, rising from 0.1 to nearly 0.5. These peaks likely correspond to events that heightened the tension, increasing discussions around Zionism and antisemitism. Results from the first growing phase are supported by findings in 4.1, in which Hamas terrorist attacks and Gaza bombings raised resentment in users' posts.
2. *Hostage and Prisoner Release*: this topic began to be discussed with the outbreak of the war on October 7^{th}, 2023, and its frequency reached a major peak, around 0.7, in mid-November 2023 followed by a gradual decline. We noticed that the growth in terms of frequency appeared to start around the end of October. Interestingly around the same period, precisely on October 21^{st}, Hamas released the first two hostages to the International Red Cross. The peak likely occurred on November 22^{nd}, when Israel and Hamas announced their first agreement on a ceasefire

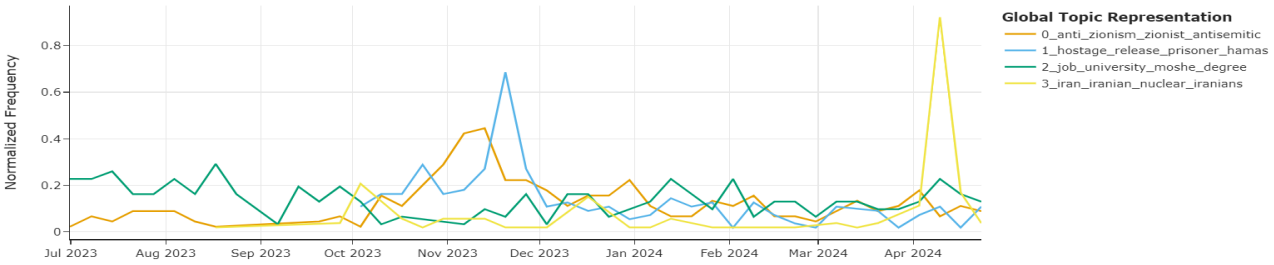

Figure 4: Normalized frequency of Topics over time for Pro-Israel content

and the release of 50 hostages, with Israel reciprocating by releasing 150 Palestinian prisoners.

3. *Job and University*: this topic maintained a low and relatively consistent level of discussion over time. It is interesting to notice that discussion on this topic appears to decrease following the outbreak of the war highlighting the increasing focus towards the conflict. Yet the level of discussion on Jobs and Universities eventually returned to pre-war levels. We interpret this as a possible sign that the Israeli community has stabilized back in response to the surrounding situation.
4. *Geopolitics and Iranian Nuclear Threat*: the frequency of discussions on this topic shot up at the beginning of April 2024, reaching their peak, 0.8, in the middle of the month. We explain this trend with the occurrence of significant geopolitical events: on April 1^{st}, Israel bombed the Iranian consulate in Damascus, escalating tensions. This situation culminated on April 13^{rd}, when Iranian forces responded by launching drones and ballistic missiles on Israel. Eventually, following the aforementioned events, we observed a shift in the representative words associated with this topic. Initially, the terms were more general, such as 'Russia', 'Iran', and 'Weapon'. However, post-attack, more specific terms like 'embassy' and 'missile' began to emerge.

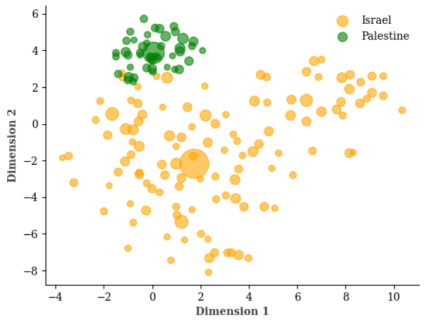

Figure 5: t-SNE Visualization of Topic Embeddings for Pro-Israel and Pro-Palestine Content

To conclude our analysis we present a scatter plot of topic embedding after t-SNE was applied (Figure 5). Each point in the graph represents a topic and the size of each point is proportional to the frequency of the topic, meaning larger points indicate more frequently discussed topics within the subreddit. Pro-Israel posts display more dispersed topics, suggesting a broader range of discussions. Hence, we conclude that the pro-Israel discussion covers a wider variety of themes compared to pro-Palestine submissions, which are predominantly focused on a narrow set of issues, mainly related to the conflict. Despite observing slight variability in the outcomes and sensitivity to model parameters, our confidence in the findings is strengthened by the consistent outcomes obtained through alternative methodologies, such as PCA and Cosine similarity between embeddings.

5. Conclusion

In this study we investigated the emotional and thematic responses of Reddit communities to the Israel-Palestine conflict that escalated on October 7^{th}, 2023. Using a DistilRoBERTa model fine-tuned for emotion detection, we quantified shifts in sentiment over time, focusing on resentment (anger and sadness), fear, and joy. Our findings reveal a significant increase in resentment and a decline in joy following major conflict-escalating events, such as the initial attacks by Hamas and subsequent Israeli military actions. Fear, on the other hand, exhibited a volatile yet stable pattern, without a clear upward or downward trend. Interestingly, these emotional trends were consistent across both Pro-Israel and Pro-Palestine communities, indicating a shared response to the escalation of violence. This analysis underscores the heightened polarization and the emotional toll of the conflict on online discourse.

Additionally, we applied a BERT-based topic modeling approach to track the evolution of discussion themes. Our results show that Pro-Israel content covers a broader range of topics, including geopolitics and societal issues, compared to Pro-Palestine content, which focuses more narrowly on the conflict. Significant peaks in specific topics align with major conflict events, reflecting the dynamic nature of online discussions in response to real-world occurrences. To conclude, our research reveals the complex dynamics of online sentiment and topic evolution in the context of ongoing geopolitical conflicts, providing valuable insights into the digital manifestation of public opinion and emotional response.

References

Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer & Veselin Stoyanov. 2019. Unsupervised cross-lingual representation learning at scale. *CoRR* abs/1911.02116.

Grootendorst, Maarten. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*

Hartmann, Jochen. 2022. Emotion english distilroberta-base. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/

Reuters. 2024. Chronology of israel's war in gaza. https://www.reuters.com/world/middle-east/major-events-during-war-gaza-2024-02-27/

Statista. 2024. https://www.statista.com/statistics/1422308/palestinian-territories-israel-number-fatalities-and-injuries-caused-by-the-israel-and-hamas-war/

Washington Post. 2024. Six months of the israel-gaza war: A timeline of key moments. https://www.washingtonpost.com/world/2024/04/07/israel-hamas-gaza-war-timeline-anniversary/

A. Appendix

A.1 Image

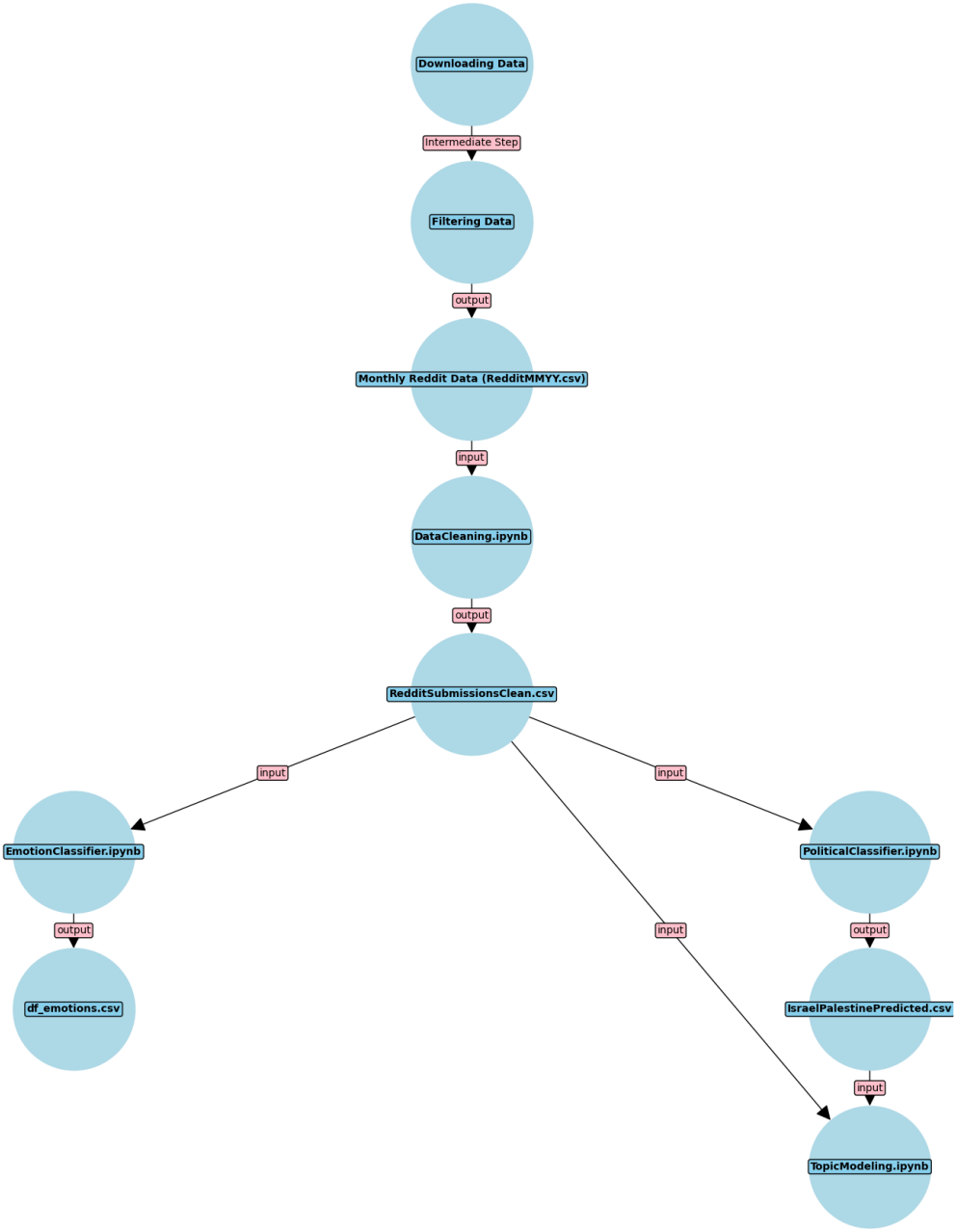

Figure 6: Comprehensive Workflow for Data Processing and Analysis, Following the Logical Steps Outlined in Our Research Process

A.2 DRIVE DIRECTORY

With this link, one can access the Google Drive folder containing the models and the code implemented in section 3.1. https://drive.google.com/drive/folders/1DvbzRUNOv7nct9bfFufKV4IXWkZXhD-w?usp=sharing