Selective Empathy of Western Media: A comparative of media framing in Ukraine-Russia vs Israel war on Gaza

Sana Saeed[a]

[a]*Nust University, Islamabad, Islamabad, Pakistan*

Abstract

This study investigates whether Western media exhibits selective empathy in its reporting on the Ukraine-Russia and Israel-Palestine conflicts. We employ machine learning and natural language processing (NLP) techniques to analyze responsibility attribution and sentiment in headlines. A rule-based system using dependency parser, SVO (Subject-Verb-Object) structures, and POS tagging is developed to predict responsibility attribution, and its results are compared with labeled data. Additionally, we used Gradient Boosting to predict whether responsibility is attributed or not. These algorithms demonstrated robust performance in handling imbalanced datasets and integrating text-based features with responsibility attribution predictions. Our findings highlight potential patterns of bias in responsibility attribution when responsible actor is "Russia"

Keywords: Media Bias, NLP, Conflict Reporting, Selective Empathy

1. Introduction

Media coverage plays an important and critical role in shaping public perceptions of conflicts around the world. Western media has been accused of displaying selective empathy when reporting on different regions and populations. In media reporting, this phenomenon can manifest in how responsibility is attributed, the tone of coverage, and the prominence given to specific actors. The Ukraine-Russia and Israel-Palestine conflicts represent two high-profile geopolitical disputes. By analyzing headlines from these conflicts, this study seeks to examine whether patterns of responsibility attribution reflect selective empathy.

2. Literature Review

Al-Sarraj and Lubbad (2018) focuses on identifying patterns of media bias in the Palestinian-Israeli conflict using sentiment analysis and supervised machine learning. By leveraging sentiment classifiers like SVM with bi-grams, it achieves up to 91.76% accuracy, emphasizing the role of biased language and narrative structures in perpetuating a pro-Israeli stance during the 2014 Gaza war. The methodology involves analyzing sentiment polarity in Western media coverage using a labeled dataset of headlines, but its reliance on predefined sentiment labels limits its ability to capture nuanced biases.

Lazaridou and Krestel (2016) explores systematic identification of media bias, particularly focusing on political affiliations. The study uses supervised models to predict the slant of newspapers based on coverage and word choices. Its methodology combines text mining and machine learning to detect framing and gatekeeping biases. However, the approach is constrained by its dependence on labeled datasets, which may not comprehensively reflect the full range of political biases in real-world media.

Graber (2017) employs Habermas's critical discourse analysis to examine representations of human shields in U.S. newspapers during the Gaza conflict. It highlights biased narratives that frame Palestinians as aggressors or victims of their own actions while justifying Israeli positions. The methodology involves qualitative discourse analysis of selected headlines and articles, but the subjective nature of qualitative analysis poses a limitation by potentially introducing interpretive biases.

Tzika (2024) compares the framing of these two conflicts, focusing on the dominance of "human interest" and "conflict" frames in reporting. The methodology involves frame analysis across multiple news outlets to identify recurring patterns, revealing how war journalism overshadows peace journalism in the Russia-Ukraine conflict. A limitation of this study is its inability to capture temporal dynamics or shifts in framing over time.

Maulana (2024) examines framing disparities between the two conflicts, revealing differences in the representation of victimhood and responsibility in Western media. It uses textual analysis of headlines and articles to uncover recurring themes and frames. While offering valuable insights, the study's complexity introduces the potential for bias in interpreting textual data.

Slimia and Othman (2022) critiques Western double standards in responding to conflicts, focusing on sanctions against Russia compared to inaction against Israel. The methodology involves a comparative policy analysis of international responses to both conflicts, revealing geopolitical biases and broader interests. Its limitation lies in focusing solely on Western countries, excluding perspectives from non-Western nations.

Kareem and Najm (2024) employs Van Dijk's socio-cognitive framework to reveal discursive strategies that portray Palestinians negatively and Israelis positively. The methodology involves critical discourse analysis of headlines and articles, highlighting the ideological role of Western media in shaping public perception. However, its focus on English-language sources limits its ability to capture biases in non-English media.

ADANE and AMARA (2024) investigates Western media narratives on the Gaza-Palestine conflict, highlighting positive portrayals of Israelis and negative depictions of Palestinians through strategic language and framing. The study employs linguistic analysis of selected articles, emphasizing the need for critical examination of media discourses to uncover underlying biases. A key limitation is its narrow focus on prominent Western media outlets, which may not generalize to the broader media landscape.

Liyih et al. (2024) applies deep learning techniques, such as hybrid CNN and Bi-LSTM models, to analyze sentiment in YouTube comments about the Hamas-Israel war, achieving high classification accuracy. The methodology involves building a labeled dataset of comments and training deep learning models, but its exclusive focus on social media limits its applicability to traditional media narratives.

Finally, Lazaridou et al. (2017) investigates bias through the analysis of reported speech in British newspapers, uncovering systematic patterns in how quotations are framed and which speakers are prioritized. It uses computational linguistics techniques to analyze quotation patterns, revealing selection and framing biases. However, the study focuses solely on speech content and does not account for visual or contextual elements that might influence reader perception.

3. Methodologgy

3.1. Data Collection

The analysis utilized a dataset of 262 headlines sourced from five major Western media outlets: BBC, CNN, NYT, Reuters, and The Guardian. These headlines covered the periods of active conflict in Ukraine-Russia and Israel-Gaza, and were categorized based on the source, date, and framing of responsibility.

3.2. Data Preprocessing

Data preprocessing involved:

1. **Cleaning and Normalizaton**: Removal of duplicates and missing values
2. **Text Tokenization and Lemmatization**: Using SpaCy, headlines were tokenized, lemmatized and stop words were removed
3. **Sentiment Analysis**: Sentiment polarity was assessed using the VADER sentiment analysis tool.
4. **Responsibility Attribution**: Headlines were labeled based on presence of responsibility related terms and structures

4. Exploratory Data Analysis EDA

4.1. Dataset Overview

- Total Headlines: 262
- Conflicts Represented:
  - Ukraine-Russia: 136 headlines
  - Israel-Gaza: 126 headlines

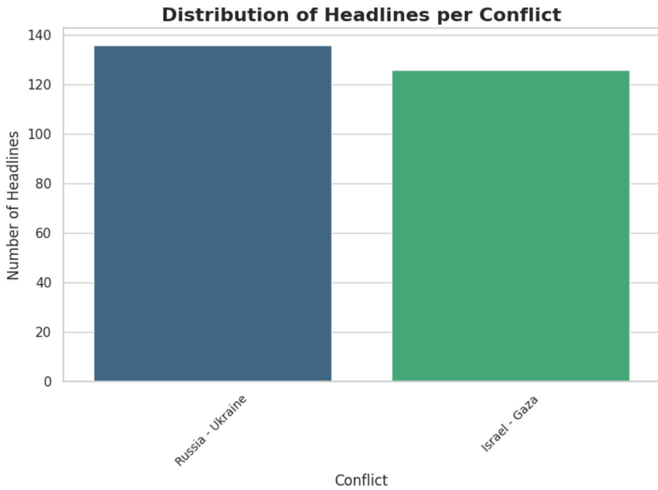


Figure 1: Distribution of headlines per conflict

4.2. Source Distribution

- NYT: 60
- CNN: 50
- BBC: 50
- Reuters: 50
- The Guardian: 52

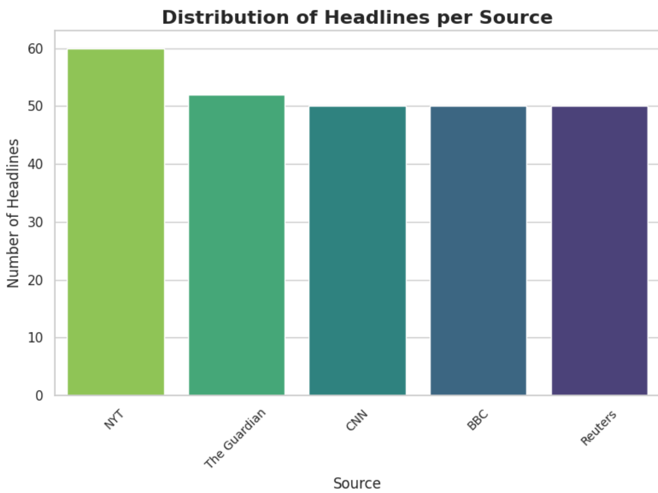


Figure 2: Distribution of headlines per source

4.3. Responsibility Attribution

Responsibility attribution is categorized into three original classes: **Direct**, **Indirect**, and **No** Responsibility, to explore how responsibility is framed in media headlines. To enhance analytical flexibility, a new column, `responsibility_merged`, has been introduced, merging some of these classes to create two broader categories:

1. Yes: Combines ”Direct” and ”Indirect” responsibility.
2. No: Corresponds to ”No Responsibility.”

The data initially contained the following distribution

- Direct Responsibility: 143
- No Responsibility: 89
- Indirect Responsibility: 30

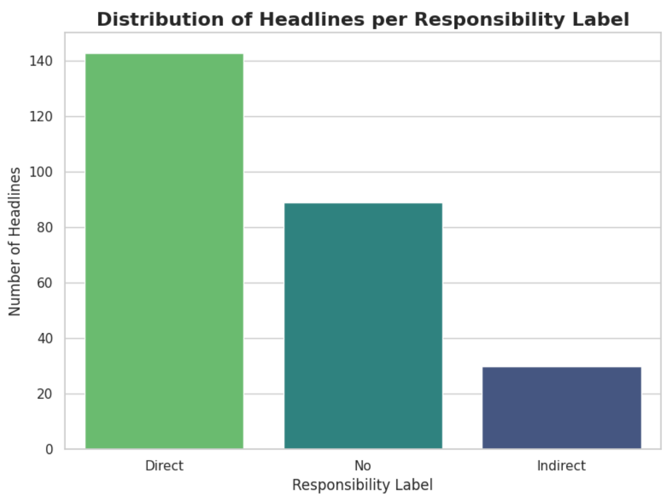
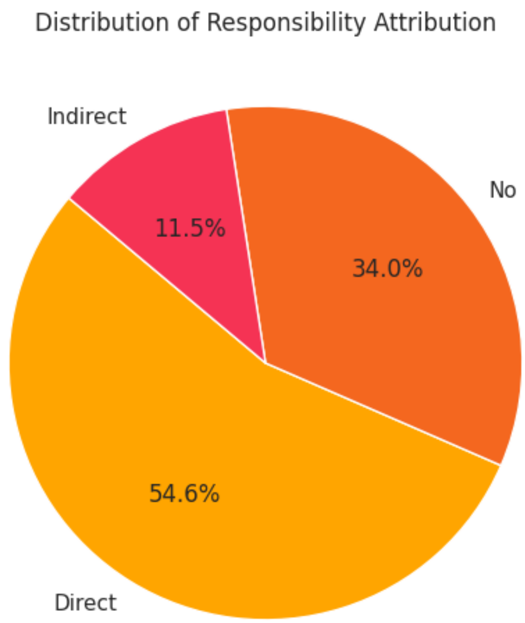


Figure 3: Distribution of headlines per Responsibility Label



Responsibility attribution in terms of ”Yes” and ”No”. The distribution is as follows

- Yes: 173
- No: 89

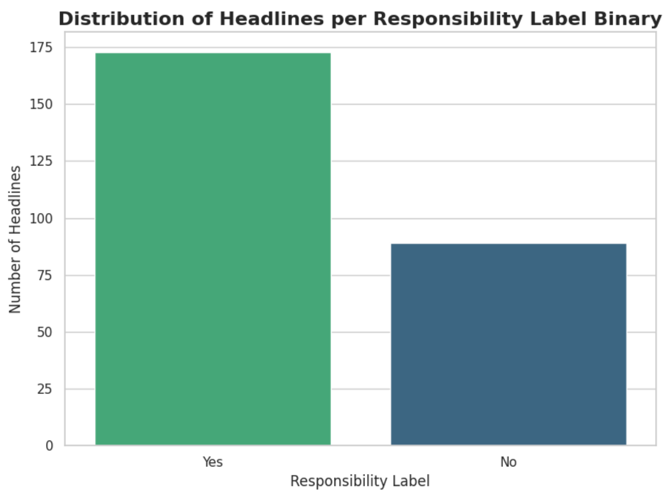


Figure 4: Distribution of headlines per Responsibility Label (Binary)

Both the original `responsibility` column and the merged `responsibility_merged` column will be used in the analysis:

1. Original Responsibility Attribution:
   - Allows for a more nuanced understanding of explicit and implicit attribution patterns.
   - Facilitates the exploration of subtle biases in media narratives.

2. Merged Responsibility Attribution:

   - Simplifies comparisons across sources and conflicts.
   - Enables robust statistical and machine learning analyses by reducing target class complexity.

By utilizing both the original and merged attribution columns, this approach balances granularity with simplicity, ensuring comprehensive insights while maintaining clarity in visualizations and model outputs.

This dual framework provides the flexibility to investigate detailed attribution patterns while enabling broader trend analyses.

Responsibility attribution according to the source can be seen in figure 5. This depicts the distribution of responsibility attribution labels (Direct, Indirect, No) across five major news outlets. The analysis reveals a dominance of Direct Responsibility attribution in most outlets

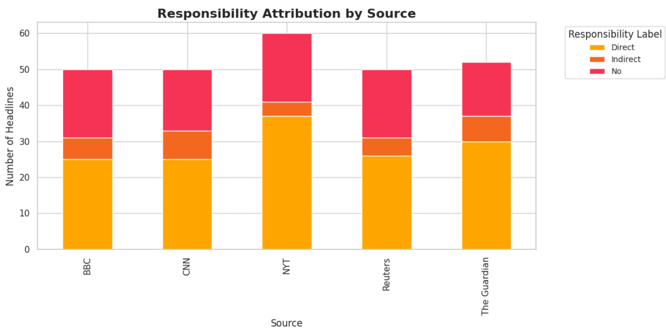


Figure 5: Responsibility Attribution by Source

Figure 6 illustrates the distribution of responsibility attribution labels (Direct, Indirect, and No) for headlines covering the Israel-Gaza and Russia-Ukraine conflicts. The data shows that headlines related to the Russia-Ukraine conflict predominantly assign direct responsibility, reflecting a clear framing of accountability. In contrast, Israel-Gaza headlines display a more balanced distribution, with a significant number of headlines avoiding responsibility attribution altogether. This contrast highlights potential differences in media framing between the two conflicts

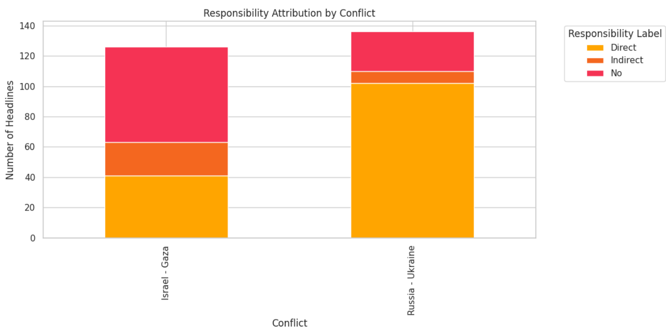


Figure 6: Responsibility attribution by Conflict

4.4. *Temporal Trends*

Data is collected from 2022 to 2024. The figure 7 shows significant fluctuations in media attention, with pronounced peaks corresponding to major events in the Ukraine-Russia and Israel-Palestine conflicts. Periods of low activity are attributed to limitations in the dataset collection process.

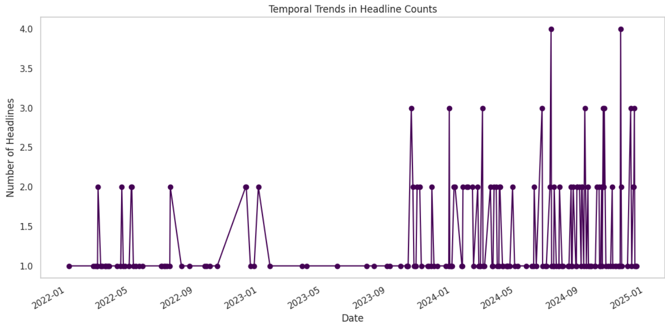


Figure 7: Temporal Trends in Headlines count

4.5. *Correlation Matrix*

Figure 8 shows the correlation matrix for the encoded features: source, conflict, and responsibility. The analysis reveals a weak relationship between the source of the headline and both conflict type and responsibility attribution. However, a moderate negative correlation between conflict type and responsibility attribution suggests that the framing of responsibility differs significantly between conflicts. This underscores the influence of conflict-specific dynamics over media sources in shaping responsibility narratives.

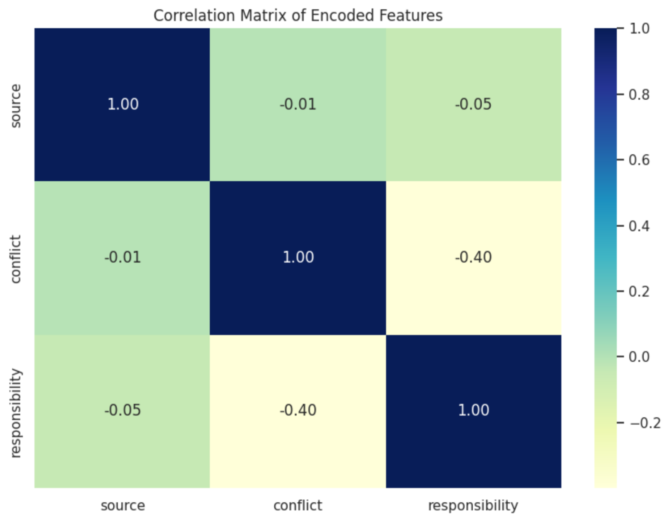


Figure 8: Correlation Matrix for Encoded Features

5. Results

5.1. *Sentiment Analysis*

Sentiment distribution by conflict:

- Israel-Gaza:

- Negative: 106
- Neutral: 14
- Positive: 6

• Ukriane-Russia:

- Negative: 122
- Neutral: 8
- Positive: 6

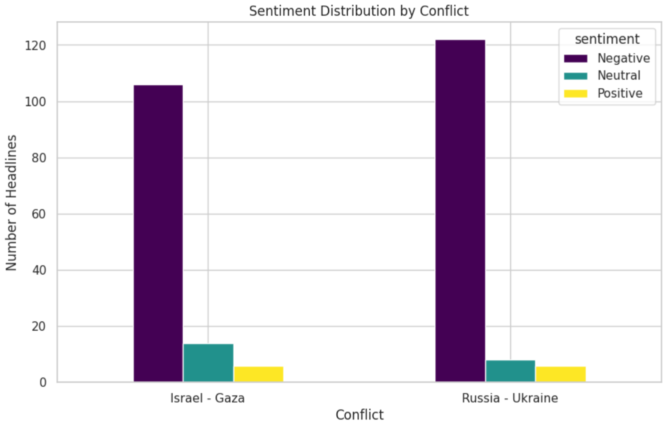


Figure 9: Sentiment Distribution by Conflict

The sentiment analysis indicated a predominantly negative portrayal in both conflicts, but with a more pronounced negativity in Ukraine-Russia. Now, this aligns with our finding that direct attribution is more often made in Ukraine-Russia rather than Israel-Gaza, therefore the prevalence of more negative sentiment.

5.2. *Framing and Responsibility Attribution*

5.2.1. *Key Findings*

1. Israel-Gaza conflict:
   - About 50% of headlines assigning responsibility
   - Gaza and Israel prominently mentioned entities
2. Ukraine-Russia conflict:
   - Higher responsibility attribution at 81%
   - Ukraine and Russia as dominant entities

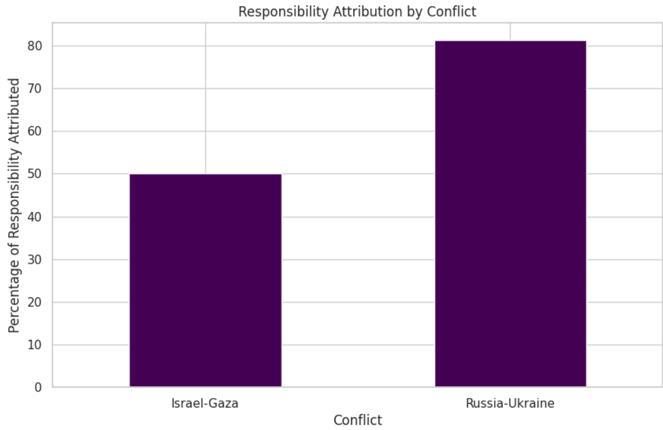


Figure 10: Responsibility Attribution by Conflict

5.2.2. *Entity Analysis*

- Israel-Gaza: Gaza (98 mentions), Israel (24 mentions)
- Ukraine-Russia: Ukraine (61 mentions), Russia (34 mentions)

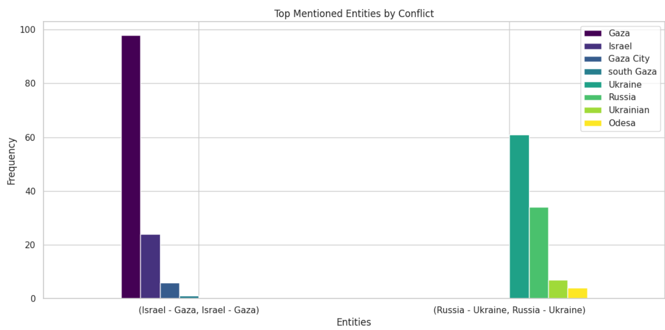


Figure 11: Top Mentioned Entities by Conflict

5.3. *Statistical and Machine Learning Analysis*

5.4. *Rule-Based Model*

A rule-based heuristic model was implemented, leveraging subject-verb-object (SVO) extraction and the presence of responsibility-related keywords. This model achieved an accuracy of 64%, highlighting its potential in scenarios with limited data but also its limitations in handling complex linguistic nuances. Rule-based system works on simple sentences more efficiently

5.4.1. *Classification Models*

The selection of specific models, such as Gradient Boosting, was driven by their ability to handle imbalanced datasets and capture complex, non-linear patterns in the data. Gradient Boosting, in particular, is well-suited for classification tasks that involve intricate relationships between features, as it builds strong predictive models by combining multiple weak learners. This capability made it the preferred choice for achieving high accuracy in responsibility attribution.

The classification of responsibility attribution was initially conducted using three target classes: Direct, Indirect, and No Responsibility. Each model's performance was evaluated on this multi-class framework:

- Gradient Boosting Model: Achieved 88% accuracy in classifying responsibility attribution
- Logistic Regression Model: Achieved 75% accuracy in classifying responsibility attribution
- Random Forest Model: Achieved 81% accuracy in classifying responsibility attribution
- Support Vector Machine SVM: Achieved 44% accuracy in classifying responsibility attribution

After evaluating the performance of these models, the Gradient Boosting Model emerged as the best-performing model. To further refine the results, fine-tuning was applied to this model, which enhanced its accuracy and overall performance.

Subsequently, the target labels were simplified by merging the Direct and Indirect classes into a single class labeled ”Yes,” representing responsibility attribution. The classes were thus reduced to ”Yes” and ”No” responsibility labels. With this binary classification, the Gradient Boosting Model achieved a significantly improved accuracy of 94

Cross-validation was conducted to ensure the robustness of the Gradient Boosting Model's performance. The model demonstrated consistent accuracy across multiple validation folds, further establishing its reliability in handling the classification task. This progression highlights the importance of iterative model refinement and class restructuring to optimize performance.

- Gradient Boosting Model: Achieved 94% accuracy in classifying responsibility attribution

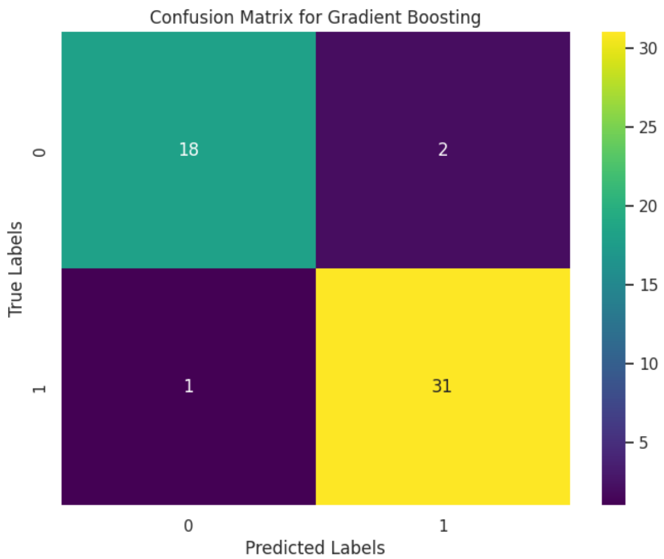


Figure 12: Confusion Matrix for Gradient Boosting

6. Discussion

The findings underscore the selective empathy in Western media's framing of conflicts. The Ukraine-Russia conflict was portrayed with a higher degree of responsibility attribution and frequent mentions of victimization. In contrast, the Israel-Gaza conflict showed a more balanced or ambiguous portrayal. These disparities reveal biases that influence public perception and policy narratives.

Media framing shapes not only public opinion but also policy decisions, potentially influencing international relations and humanitarian aid. Biased reporting may perpetuate stereotypes, exacerbate tensions, or lead to unequal resource allocation between conflicts. These insights emphasize the importance of fostering balanced reporting to promote equitable awareness and action across global conflicts.

The findings underscore the selective empathy in Western media's framing of conflicts. The Ukraine-Russia conflict was portrayed with a higher degree of responsibility attribution and frequent mentions of victimization. In contrast, the Israel-Gaza conflict showed a more balanced or ambiguous portrayal. These disparities reveal biases that influence public perception and policy narratives.

7. Conclusions

This study highlights the disparities in media framing of the Ukraine-Russia and Israel-Gaza conflicts, using both statistical analysis and machine learning techniques to uncover patterns in responsibility attribution, sentiment, and entity prominence. The dataset comprised 262 headlines from five major Western media outlets, revealing that the Ukraine-Russia conflict received a higher level of responsibility attribution (81%) compared to the Israel-Gaza conflict (50%). Additionally, sentiment analysis showed predominantly negative portrayals for both conflicts, with a slightly more negative emphasis in Ukraine-Russia headlines.

Advanced analyses using machine learning models demonstrated the effectiveness of computational approaches in classifying responsibility attribution. The Gradient Boosting Model achieved the highest accuracy (94%), followed by Logistic Regression (75%) and the rule-based model (64%). Support Vector Machine (SVM) performed the poorest, with an accuracy of 44%, highlighting the variability in model performance depending on the complexity of the classification task.

The rule-based heuristic approach, while achieving a moderate accuracy of 64%, was instrumental in extracting linguistic patterns such as subject-verb-object structures and keywords related to responsibility. The findings suggest that media framing significantly shapes public perception and policy discussions, often reflecting biases in responsibility attribution and sentiment portrayal. This study's multi-method approach underscores the importance of integrating rule-based heuristics and machine learning to provide comprehensive insights into media behavior

8. Limitations and Future Work

8.1. Limitations

This study has several limitations that should be acknowledged:

1. Dataset Size: The analysis was based on a dataset of 262 headlines, which, while sufficient for identifying general trends, limits the granularity and statistical power of some findings. A larger dataset would allow for more robust conclusions and deeper analysis of nuanced patterns.
2. Language Bias: The study relied solely on English-language media from Western outlets. This inherently excludes perspectives from non-Western and non-English-speaking media, which could provide contrasting narratives and framing techniques.
3. Temporal Coverage: While the dataset includes headlines from key periods of conflict, it does not encompass a continuous timeline, potentially missing shifts in media framing over time.
4. Focus on Textual Data: The study was limited to textual analysis of headlines, excluding other forms of media content such as images, videos, or full articles, which could provide a more comprehensive understanding of framing.

8.2. Future Work

To address these limitations and expand the scope of this research, several extensions are proposed:

1. Larger and More Diverse Dataset: Future studies should aim to collect a larger dataset spanning multiple languages and media regions. This would enable cross-cultural comparisons and a more global perspective on media framing.
2. Efficient Data Collection: Leveraging automated tools for web scraping and natural language processing could streamline the process of collecting and pre-processing data. Incorporating continuous data collection pipelines would allow for real-time analysis of media framing trends.
3. Multimodal Analysis: Future research could integrate textual analysis with visual content analysis to explore how images, videos, and infographics contribute to media framing.
4. Temporal Trends: Expanding the dataset to cover continuous timelines would facilitate a more dynamic analysis of how media narratives evolve in response to events.

References

ADANE, H.B., AMARA, Y., 2024. Analyzing Western Media Narratives on Israel Attack on Gaza-Palestine: a Critical Discourse Analysis. Ph.D. thesis. ibn khaldoun university-Tiaret.

Al-Sarraj, W.F., Lubbad, H.M., 2018. Bias detection of palestinian/israeli conflict in western media: A sentiment analysis experimental study, in: 2018 International Conference on Promising Electronic Technologies (ICPET), IEEE. pp. 98–103.

Graber, S.M., 2017. War of perception: a habermasian discourse analysis of human shield newspaper reporting during the 2014 gaza war. Critical Studies in Media Communication 34, 293–307.

Kareem, A.H., Najm, Y.M., 2024. A critical discourse analysis of the biased role of western media in the israeli-palestinian conflict. journal of Language Studies 8, 200–215.

Lazaridou, K., Krestel, R., 2016. Identifying political bias in news articles. Bulletin of the IEEE TCDL 12, 1–12.

Lazaridou, K., Krestel, R., Naumann, F., 2017. Identifying media bias by analyzing reported speech, in: 2017 IEEE International Conference on Data Mining (ICDM), IEEE. pp. 943–948.

Liyih, A., Anagaw, S., Yibeyin, M., Tehone, Y., 2024. Sentiment analysis of the hamas-israel war on youtube comments using deep learning. Scientific Reports 14, 13647.

Maulana, M.A., 2024. A comparative analysis of western nations' actions in russia-ukraine and israel-palestine conflicts. Nation State: Journal of International Studies 7, 29–52.

Slimia, A., Othman, M.F., 2022. The double standards of western countries toward ukraine and palestine "western hypocrisy". Central European Management Journal 30, 476–485.

Tzika, S., 2024. A comparative analysis of media framing: The israel-hamas and the russia-ukraine wars .