MIE 1624 Course Project

SHAPING POSITIVE PERCEPTIONS:

SENTIMENT ANALYSIS OF RUSSIA'S WAR ON UKRAINE

Group 1

Bavani Bhoopathy - 1008579389

Bill Tang - 1003423921

Joyce Chen - 1003057577

Manoj Machavolu - 1008518506

Sudarsanamurthy Tenneti Prabakaramurthy - 1008628601

EXECUTIVE SUMMARY

Social media is a powerful tool in shaping public perceptions. As such, it can be crucial to understand the sentiments and drivers to effectively improve public opinion. This analysis focuses on understanding the sentiments that are in social media posts and news sites today and identifying key factors and topics contributing to these sentiments. Utilising this public opinion sentiment analysis, the goal is to recommend strategic actions and frame messages in a more positive light to improve Ukraine's image.

With the amount of posts on social media, the analysis limits social media sources to english Twitter, Reddit, CNN, and Fox news posts and headlines within the last three years, mainly with key words like Ukraine, Russia, war and conflict. The methodology uses natural language processing, with a machine learning Logistic Regression model, to output the sentiment classification results.

Overall, negative sentiments consisted of a higher percentage in all datasets, with words like “war, attack, invasion, crime, force and weapon making” having high frequency. For positive sentiments, words such as “support, thanks, peace and love” are prevalent. With the sentiment classification, BERT and LDA Topics modelling was done, resulting in three major themes and topics: Attacks on Cities and Critical Infrastructure, Humanitarian Issues and Aid, Political News & Misinformation.

Using these topics, recommendations can be made under the following narratives: Heroes and Individual Narrative, The Uplifting Humanitarian Efforts and Counteracting the Negative. These themes focus on evoking strong emotional responses by utilising key figureheads and identifiable heroes to relate positive emotions with the public. The recommendation is to also highlight the effect of humanitarian aid to critical areas that are getting negative sentiments, and to ensure transparency in communication so that individuals can see the direct impact of their actions. This transparency can also help combat misinformation, and negative sentiments association. By addressing the negativity with transparency and accountability, trust can be built up and can help lean public opinions towards more positive sentiments.

Overall, the first step to improving Ukraine’s international perception is to understand the sentiments. With the analysis and recommendations provided, Ukraine can work towards addressing these sentiments by shaping messages in a positive light.

TABLE OF CONTENTS

INTRODUCTION **3**
PROBLEM STATEMENT **3**
OBJECTIVES 3
SCOPE 3
METHODOLOGY **4**
SENTIMENT CLASSIFICATION 4
LIMITATIONS 4
RESULTS **5**
SENTIMENT BREAKDOWN 5
REDDIT DATA SOURCE 6
TWITTER DATA SOURCE 6
NEWS (CNN AND FOX) DATA SOURCE 7
KEY THEMES AND TOPICS 8
RECOMMENDATIONS **9**
HEROES AND INDIVIDUAL NARRATIVE 9
THE UPLIFTING HUMANITARIAN EFFORTS 9
COUNTERACTING THE NEGATIVE 10
SHAPING POSITIVE STORIES WITH RHETORIC 11
CONCLUSION AND FUTURE WORK **11**
REFERENCES **12**
APPENDIX **13**

INTRODUCTION

With the popularity and prevalence of social media today for sharing opinions, it is important to understand its effect in shaping and influencing public perception. With the start of Russia's war on Ukraine in 2022, an outpouring of political views, information, and communications have emerged on social media sites such as Twitter and Reddit. This report focuses on these social media sources as a tool to gauge public opinion on a variety of themes and topics and to inform next steps in improving international sentiments towards Ukraine.

PROBLEM STATEMENT

Despite the vast amount of data and opinions that exist on social media and news sites, it is difficult to gain a direct understanding of the key trending topics and issues. This translation of raw opinions and discussions to focused public sentiment data is crucial in policy making, public engagement and strategic action. As such, this analysis aims to address this issue by providing insight on prevalent topics and sentiments towards Ukraine to generate informed and data driven recommendations for improvement of international public perception towards Ukraine.

OBJECTIVES

1. Analyse sentiments of social media posts/tweets and news sites to gain an understanding of general public opinion towards Ukraine
2. Identify key factors and topics that contribute towards these sentiments
3. Recommend strategic actions for the Ukrainian Government and NGOs to improve Ukraine's image based on public opinion data

SCOPE

With the constant influx of new content on social media and of news , it was necessary to reduce the scope of this project. As such, Sentiment Source, Time Frame and Topic constraints were set as the following:

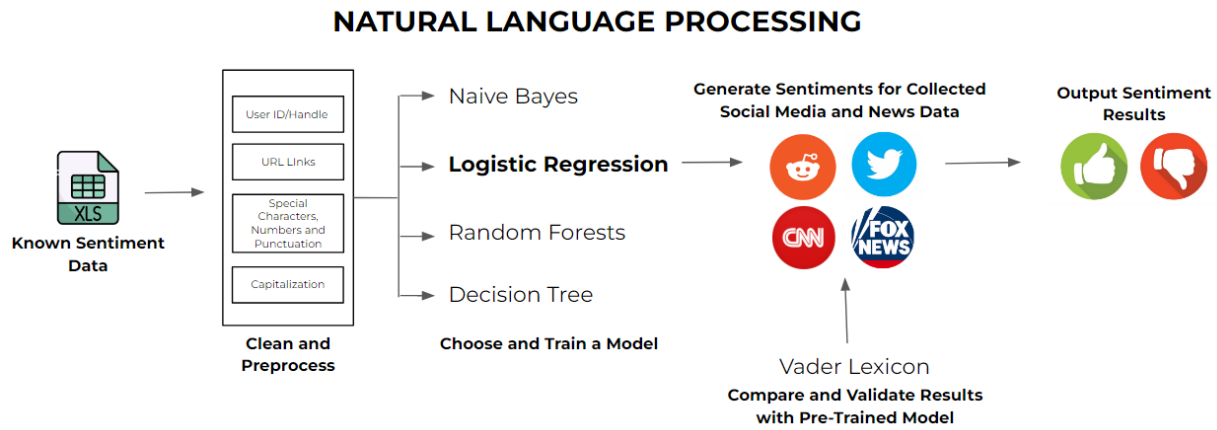
SENTIMENT SOURCES	**Twitter** - real time tweets **Reddit** - online discussions threads **CNN** - "Left" wing news **Fox News** - "Right" wing news
TIMEFRAME	**2021** - Rising Tensions **2022** - Official Instigation **2023** - Current Ongoing Timeline
TOPICS AND LANGUAGE	**Key Words:** Ukraine, Russia, war, conflict **Language:** English

METHODOLOGY

The analysis was done using Natural Language Processing (NLP) and Machine Learning (ML) algorithms. ML algorithms were evaluated and applied as part of the NLP used to generate sentiment results from Reddit, Twitter and official news publications from CNN and Fox News.

SENTIMENT CLASSIFICATION

Logistic Regression and NLP allowed text to be divided into positive, neutral and negative sentiments. The rest of the report will focus notably on the positive and negative sentiments. Detailed ML and NLP procedures and code can be found in included IPYNB notebook.



LIMITATIONS

It is to note that the results are representative of only a subset of social media as narrowed by the scope and thus bias may exist in the results. Additionally, tonal context, colloquialism and language nuances such as sarcasm, idioms and rhetoric may not be fully captured into sentiments.

RESULTS

After running the model, topic analysis and visualisations as listed in Table 1 were done on the resulting sentiments. The following section depicts the key takeaways from the sentiments data gleaned out of each source. The full list of word cloud charts can be found in Appendix A.

Table 1: Analysis and Visualisations Done with Sentiments

Chart Type	What Was Done?	What is it?	How to Interpret?
Pie Chart	Scrape source and split into positive, neutral and negative	Percentage breakdown of sentiments	Overall sentiment breakdown percentage in a source
WordCloud	Split text into positives and negatives	Word frequency	The bigger the font, the more frequently the word appears
Unigram/ BiGram	Split text individual “token” words	Sequence of single words or pairs of words	Frequency of certain tokens in context
Network Diagram	Plotted based on unigram data	Distance and frequency correlation between words	Nodes are together if they are contextually close in headlines. The darker the line, the more often they appear in headlines.
LDA (Latent Dirichlet Allocation)	Applied this topic model with the positive and negative data for each source and ask for a number of topics to generate	Identifies probability of a topic relating frequencies and words appearance	Each topic is grouped together to see common trends.
BERT	Applied this topic model with the positive and negative data for each source	Automatically generates top topics based on context and text	Each topic is grouped together to see common trends based on context.

SENTIMENT BREAKDOWN

Overall, negative sentiments consisted of a higher percentage in all datasets, with Twitter Webscrapped Dataset 2 and Fox News holding the highest percentage overall, as seen in Figure 1.

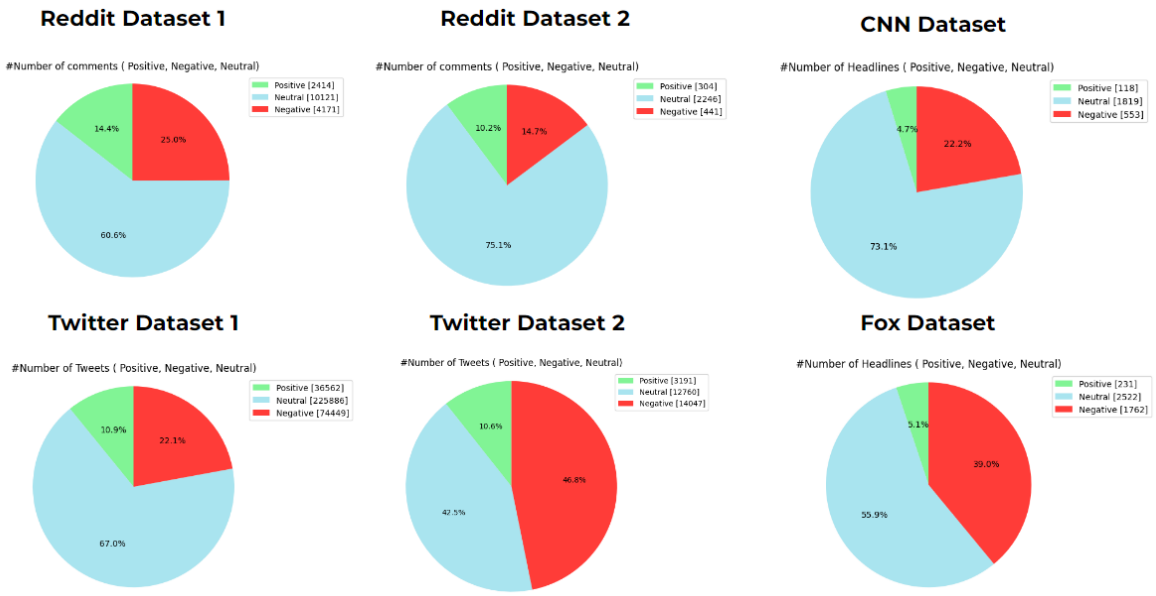


Figure 1: Positive, Negative and Neutral Break Down Among Data Sources

REDDIT DATA SOURCE

From the word cloud and unigrams, words like support, thanks, peace and friends are labelled as positive sentiments. Similarly, words like war, attack, missile, enemy, forces, fire and artillery are major contributors towards the negative sentiment. In the network graph for negative sentiments, as shown in Figure 2, it is clearly indicated that every word in the headline is strongly connected to "war", with war being the centre of the centre of the cluster. The full untruncated diagram can be found in Appendix B.

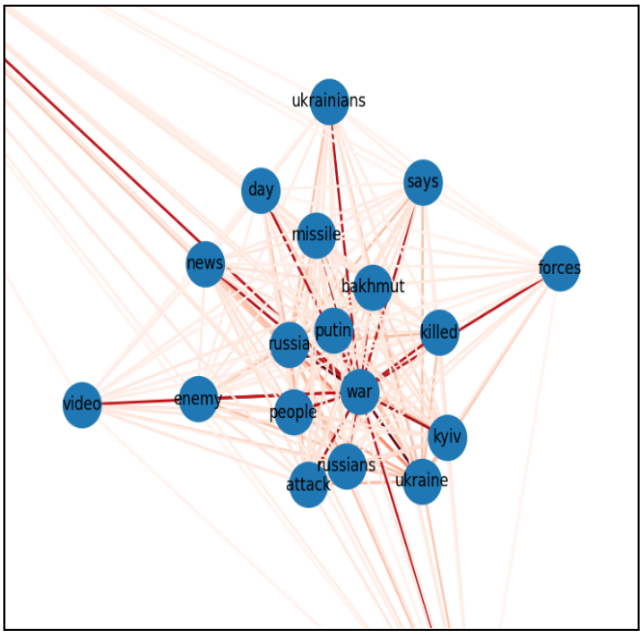


Figure 2: Negative Network Graph

TWITTER DATA SOURCE

From the word clouds and unigrams shown in Figure 3 and Appendix C, words like support, thanks, peace and love lean towards the positive sentiment. Similarly, words like war, attack, invasion, crime, force and weapon are contributors towards the negative sentiment. While words like Russia and war do appear on the positive unigram, this can be inferred to be caused by the high frequency of these words, along with the model's limitation to pick up sarcasm from posts (e.g. "Wow, great, another war!").

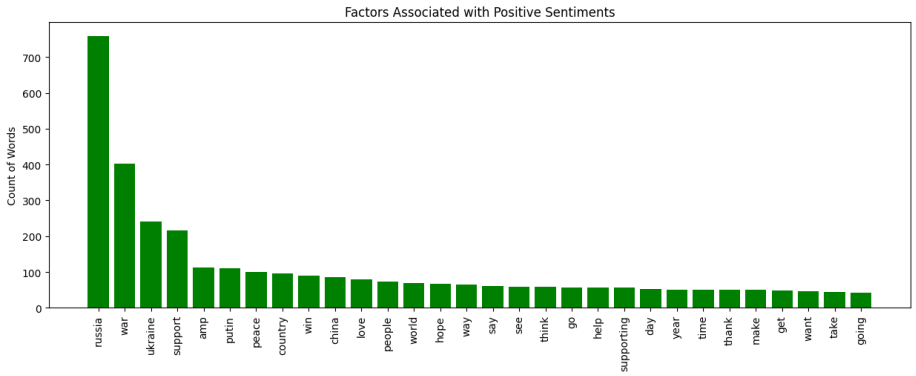


Figure 3: Twitter Positive Sentiment Unigrams

NEWS (CNN AND FOX) DATA SOURCE

The main results for the news sites are shown in the timeline, which largely followed the events in Ukraine. As seen in Figure 4, the frequency of negative tweets relate to the timeline of Russia's war on Ukraine.

In general , the word cloud, unigram and bigram in Appendix A and Appendix D reflected the events as well. Notably, words like support, thanks, peace, military assistance, support Ukraine, US support, and Zelensky are major contributors towards the positive sentiment. Similarly, words like war, attack, missile, invasion, forces, strike, weapons, Russian forces, crimes, death and Russian shelling are negative. These words can be seen to allude to the events that occurred throughout 2022 to 2023 as follows [1] :

Feb 24th: Official Instigation by Russian President Putin

Apr - May 2022: Maternity Hospital Bombing, war crimes in Bucha

Oct - Nov 2022: Critical Energy Plant in Kyiv is attacked with missiles

Dec 2022 - Feb 2023: General rising in positive sentiments as aid is approved from international countries and Kherson is liberated

March 2023: President Biden visits Kyiv

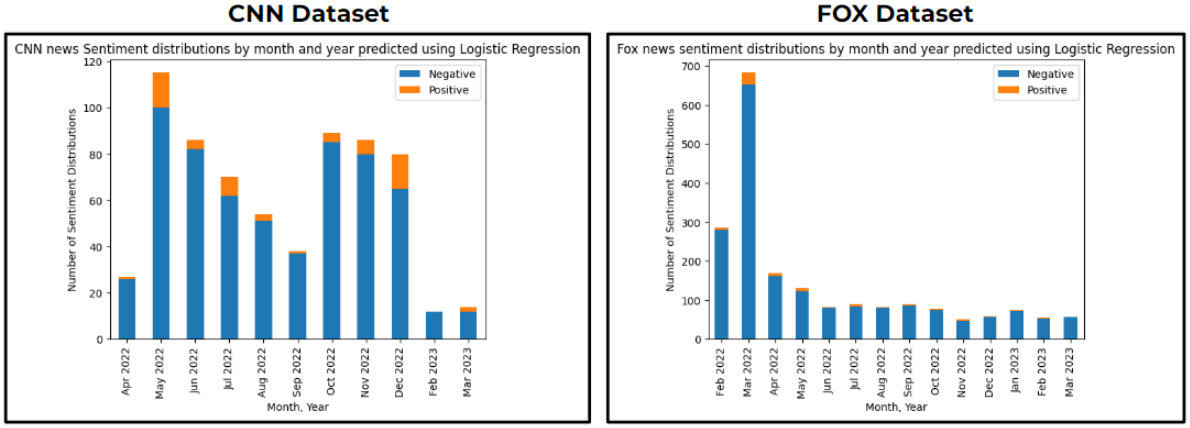


Figure 4: Sentiment Analysis Over Time

KEY THEMES AND TOPICS

Using BERT and LDA with sentiment analysis, topics were grouped up into three main themes as seen in Figure 5. Samples of a full BERT and LDA analysis can be found in Appendix E.

1. **Attacks on Cities and Critical Infrastructure**
    a. Examples: Bakhmut, Kharkiv, Kherson, Donetsk/Azovstal
    b. Topics on missiles, bombings, killings, battles and assaults
2. **Humanitarian Issues and Aid**
    a. Global food crisis
    b. Refugees, families and civilian situations
    c. Support from Biden, Germany, NATO alliance
3. **Political News & Misinformation**
    a. Fake Referendums
    b. Russian media and opposition narrative
    c. ICC arrest warrants

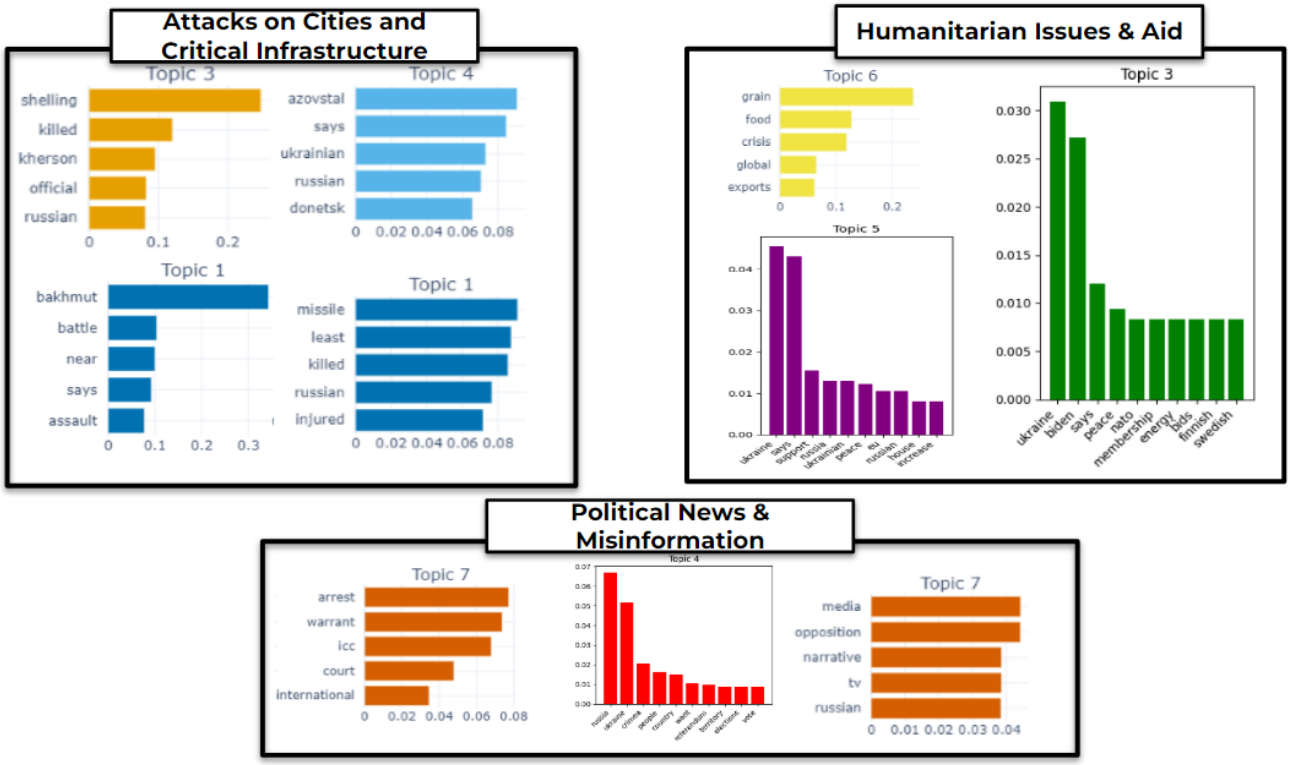


Figure 5: Themed BERT and LDA Topic Outputs

With the distinction of these key topics that are prevalent on social media, the following section will focus on addressing shaping these issues in a positive light.

RECOMMENDATIONS

Shaping positive perceptions:	How do we do this effectively?
1. Heroes and Individual Narrative	1. Relatability: Reflecting words back to the readers
2. The Uplifting Humanitarian Efforts	2. Positive and transparent Language
3. Counteracting the Negative	3. Humour

HEROES AND INDIVIDUAL NARRATIVE

There is a well known psychological principle of people being more empathetic to an identifiable individual compared to a large group that is suffering similarly. This psychological principle, called the "identifiable victim effect" or "singularity effect", partially explains why charity campaigns often feature individual stories while soliciting donations [2]. An extension of this principle is a hero's story, as it also focuses on an individual, but in addition to their suffering, also features virtues like courage, strength, resilience and sacrifice.

In the case of the Russian Ukraine war, there were many mentions of president Zelensky in public conversation. His name is highly correlated to posts with positive sentiments. This suggests that he is well regarded on the international stage and has a positive impact on Ukraine's international image. Another example of a hero would be the soldier who was tragically killed by the Russian force because he refused to say "slava Russia '' but instead chose "slava Ukraini '' as his last word [3]. The vast majority of comments under the video had sentiment that is positive towards Ukraine or negative towards Russia.

To positively influence Ukraine's international image, the recommendation would be to share personal stories, images, or testimonials to create an emotional connection and make the issue more relatable and compelling to the general public. Large-scale suffering can be overwhelming and may lead to a sense of helplessness or compassion fatigue. In contrast, focusing on a single individual may feel more manageable and motivate people to take action.

THE UPLIFTING HUMANITARIAN EFFORTS

Sentiment analysis in social media can be especially useful for Humanitarian Aid. The association of positive sentiments in social media with words such as "refugees, children, people, help, aid" can be leveraged to further communicate motivations for action. This strategy will appeal to the general public's compassion in aiding suffering innocents (refugees, children, families and civilians etc.) and their ability to see the impact they

are making. One of the most prominent ways of framing sentiments would be to show the effect of aid to critical areas that are getting negative sentiments, such as Bakhmut and Kharkiv, which were both words that have shown up in the sentiment analysis.

This transparency and communication prioritises the issues that come up most frequently for sentiments, such as the food crisis. For instance, projects on investment into agriculture, and NGOs that are helping to provide food security can be promoted. Similarly, establishment of shelters, safe zones, and healthcare in Kherson, where bombings, attacks and missiles were particularly rampant as seen in the negative word sentiments. Notably, showcasing of international support, such as from President Biden and NATO on a positive humanitarian narrative, such as housing refugees and those affected by the war, can help propagate the positivity on an international scale.

COUNTERACTING THE NEGATIVE

Cognitive biases refer to systemic patterns of deviation and rationality in judgement. In the context of the Ukraine-Russian war, these biases can play a significant role in the way people perceive Ukraine. It can lead to inaccurate judgement and can distort public opinion towards Ukraine.
Due to Russia's war against Ukraine, the media coverage around Ukraine tends to have a more negative connotation. This causes people to have a negative view of Ukraine than they otherwise would. The anti-Ukrainian propaganda machine has been systematically using these biases to promote false perceptions to blur ground-level truths, create a divide and undermine the support Ukraine receives. Therefore, it is extremely important to counteract the negative image by improving accountability, building trust and controlling the spread of disinformation.
Transparency is essential for changing perceptions towards Ukraine as it addresses some common sources of negative opinions such as mistrust and misinformation. The Ukrainian government must understand the various narratives being used to spread disinformation and identify the most vulnerable audience to foster resilience against manipulative narratives. To facilitate this, sentiment analysis can act as a useful tool. Words from the analysis such as referendum, warrants and opposition were often found in negatively inclined posts, tweets and headlines. To combat this, the impact of aid on Ukrainian lives must be highlighted. Being transparent about aid allocation statistics and reporting the progress of other reconstruction processes is crucial for swaying public opinion in Ukraine's favour.

SHAPING POSITIVE STORIES WITH RHETORIC

Delivery is a critical aspect of making powerful, convincing and positive narratives. With the power that social media has on driving change and awareness, it is essential that the messages be sent with the appropriate framing. This can be done in three main ways:

1. **Relatability:** With sentiment analysis, hot topics and words can be taken and used to reflect words back to the readers to create content that will resonate. This ties into the concept of choosing the appropriate figureheads and evaluating topics of concern within social media.
2. **Positive Language:** Critically, with humanitarian efforts, it can be beneficial to avoid words with negative or violent connotations such as “weapons, missiles, nukes etc.”. This will help frame the context in a more positive manner and redirect attention to the positives of the humanitarian efforts rather than the violence and tragedies occurring.
3. **Humour:** Humour and satire can be very powerful tools with the right audience. Humour is especially effective for the social media audience which are generally younger, who grew up communicating in memes and jokes. These audiences can be suitable for more uplifting, humorous content which can evoke a more positive and enthusiastic response. An example would be the use of memes, or satirical comics as they are easily digestible, shareable and relatable. They bring laughter to people, which is emotionally appealing and bypasses many barriers against new information.

CONCLUSION AND FUTURE WORK

To sum up, improving Ukraine’s international perception requires a multifaceted approach. Ukraine can improve its reputation by using relatable, humorous and positive language. With the recommendations provided in this report, Ukraine can build stronger relationships globally and make good progress in rebuilding after the war.

While this report provides several insights into the sentiment surrounding Ukraine and the conflict, there is room for further analysis. A granular analysis could be done to identify specific regions or countries that hold unique views. Further exploration into local and international sentiment can also be insightful. Finally, analysing social media posts in other languages, particularly those spoken in countries with a significant stake in the conflict will enhance the quality of insights gained.

REFERENCES

[1] “Russian invasion of Ukraine: A timeline of key events on the 1st anniversary of the War,” CNN. [Online]. Available: https://www.cnn.com/interactive/2023/02/europe/russia-ukraine-war-timeline/index.html.

[2] Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. Organizational Behavior and Human Decision Processes, 102(2), 143-153.

[3] Sky News. (2023, March 7). Ukraine War: 'executed' soldier video could be a war crime. YouTube. Retrieved March 29, 2023, from https://youtu.be/CCB_L9ayodg

APPENDIX

APPENDIX A: WORD CLOUD

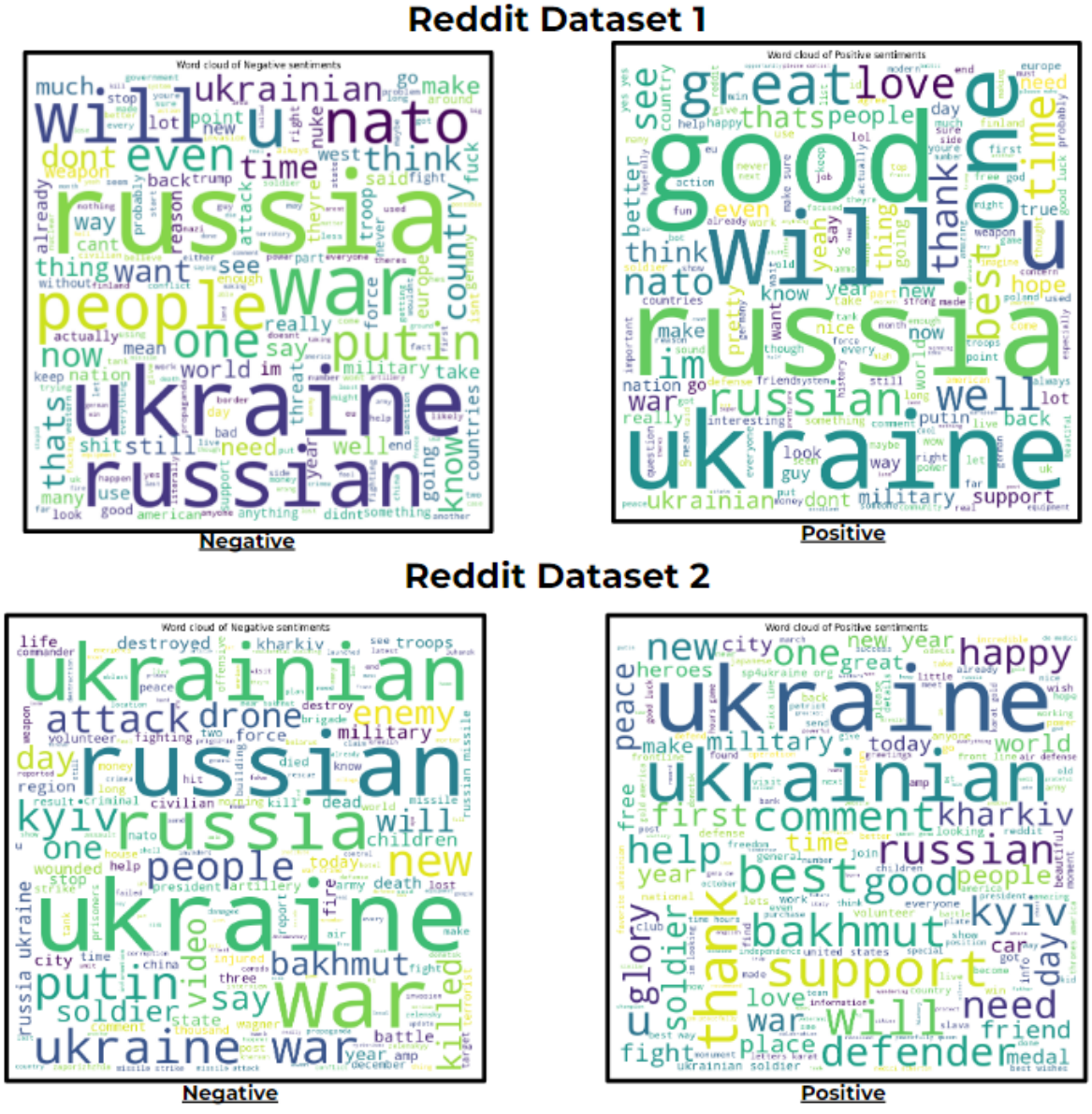
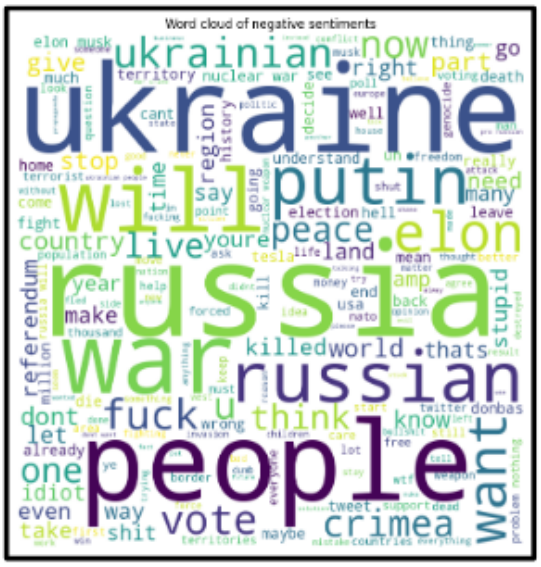


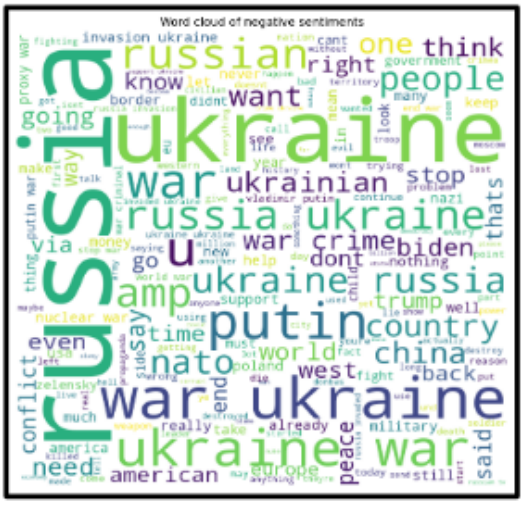
Figure A.1: Reddit Word Cloud

Twitter Dataset 1



Negative **Positive**

Twitter Dataset 2



Negative **Positive**

Figure A.2: Twitter Word Cloud

CNN Dataset

Negative

Positive

Fox Dataset

Negative

Positive

Figure A.3: News Word Cloud

APPENDIX B: REDDIT NEGATIVE NETWORK DIAGRAM

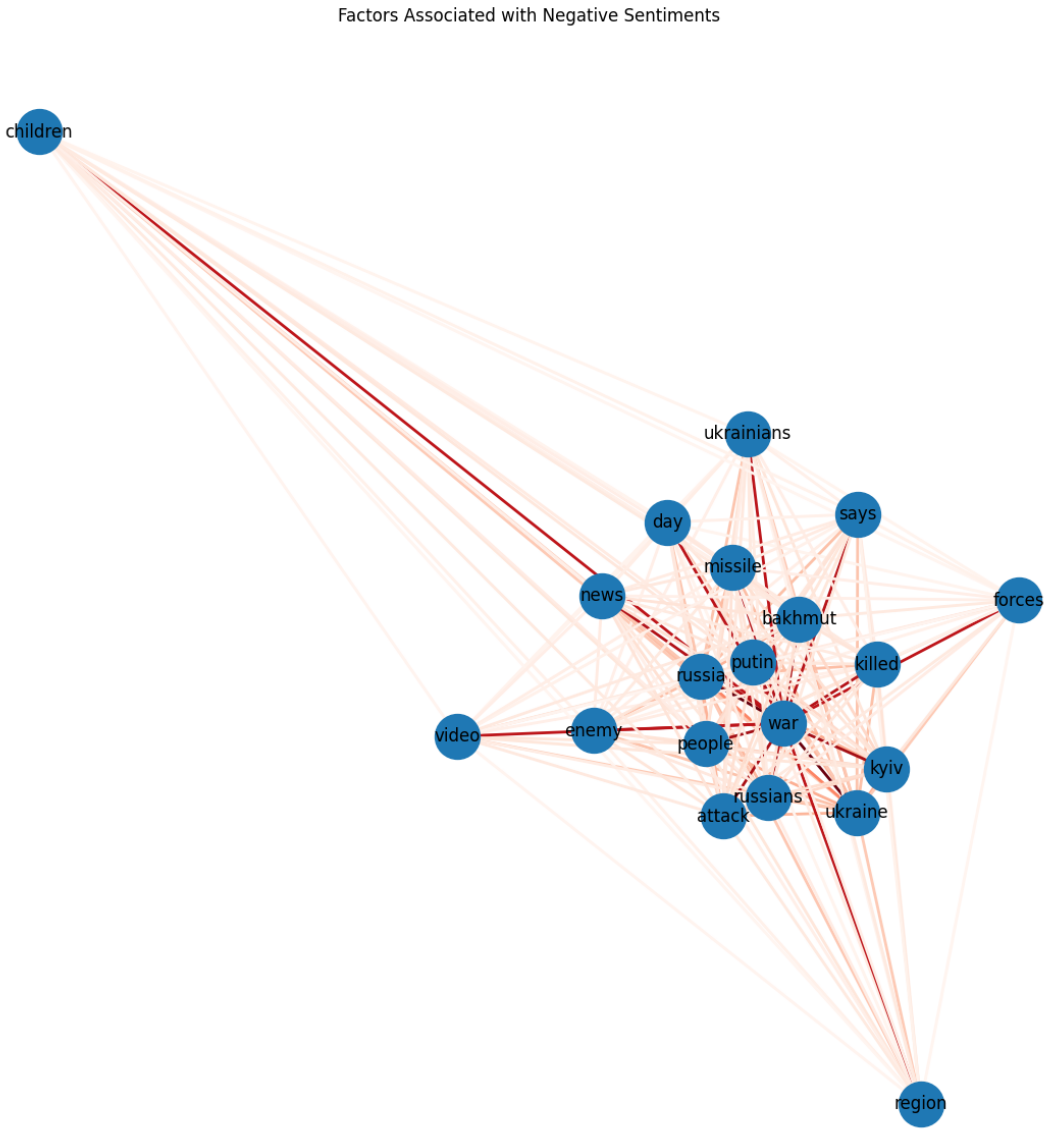


Figure B.1: Full Negative Sentiment Network Diagram

APPENDIX C: TWITTER POSITIVE AND NEGATIVE WORD SENTIMENTS

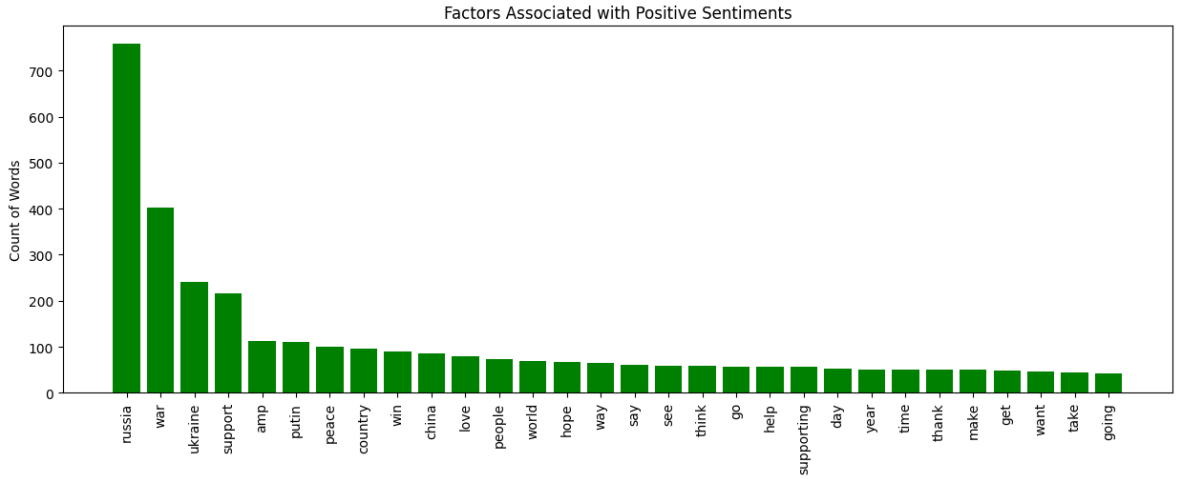


Figure B.1: Twitter Positive Sentiment Words

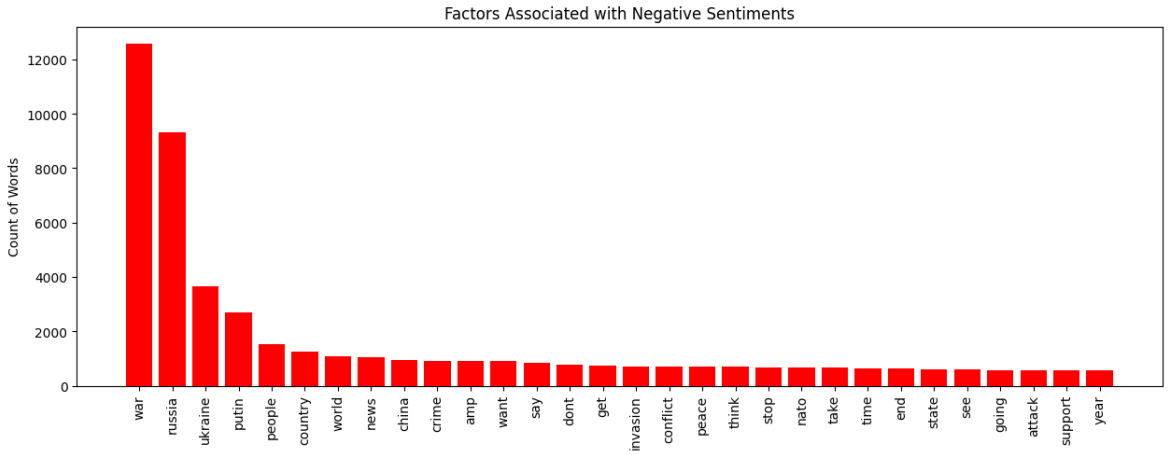


Figure C.2: Twitter Negative Sentiment Words

APPENDIX D: CNN AND FOX POSITIVE AND NEGATIVE WORD SENTIMENTS

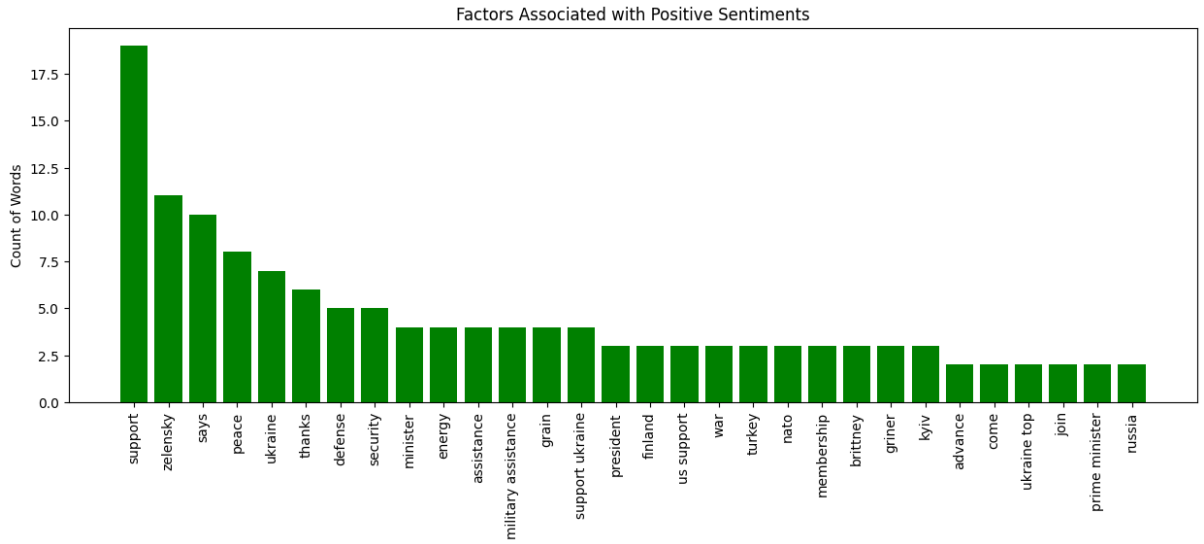


Figure D.1: CNN Positive Sentiment Words

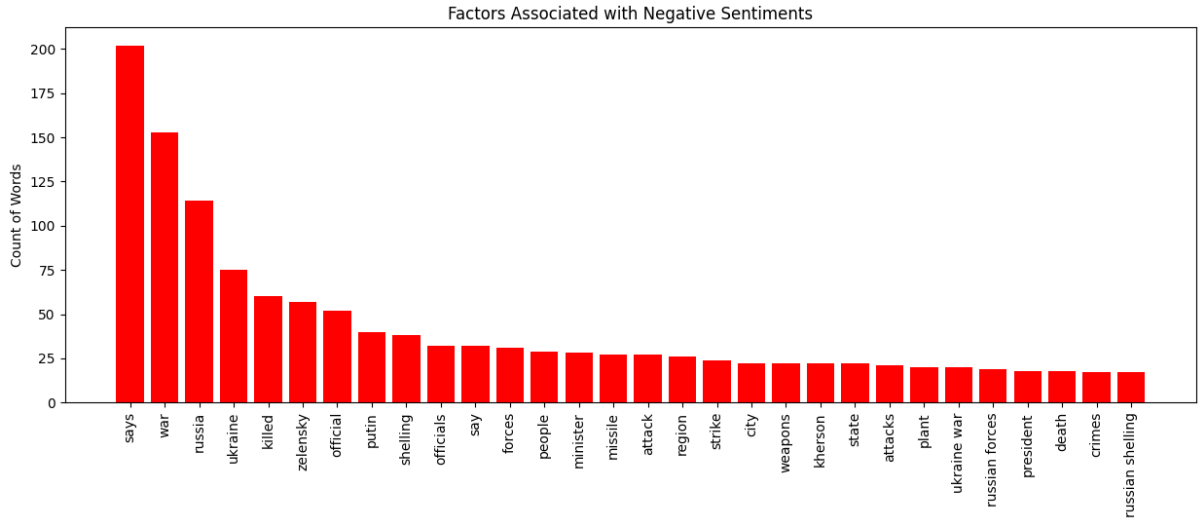


Figure D.2: CNN Negative Sentiment Words

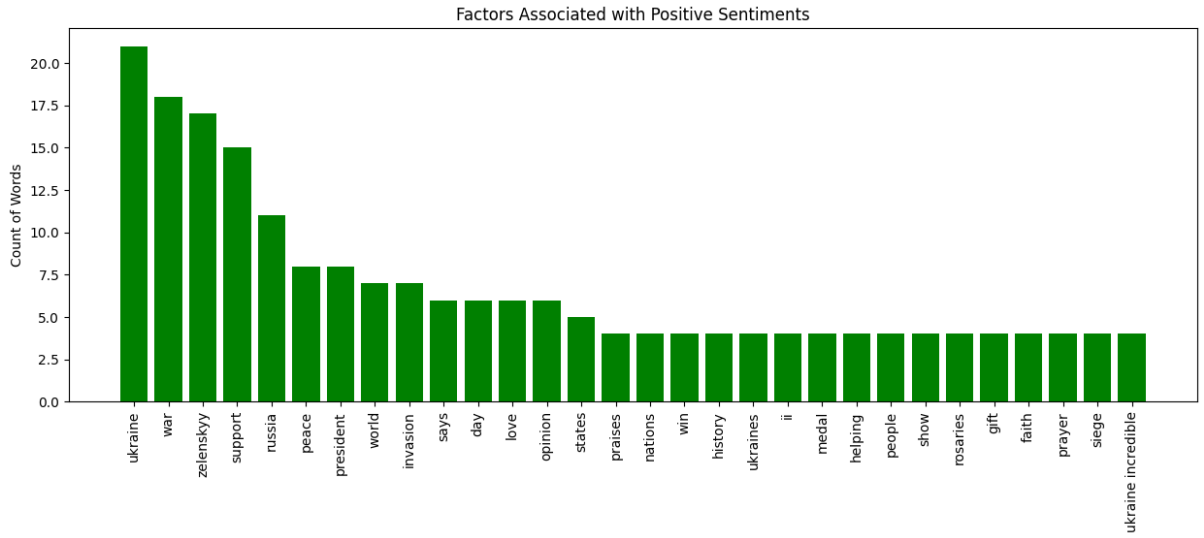


Figure D.3: Fox News Positive Sentiment Words

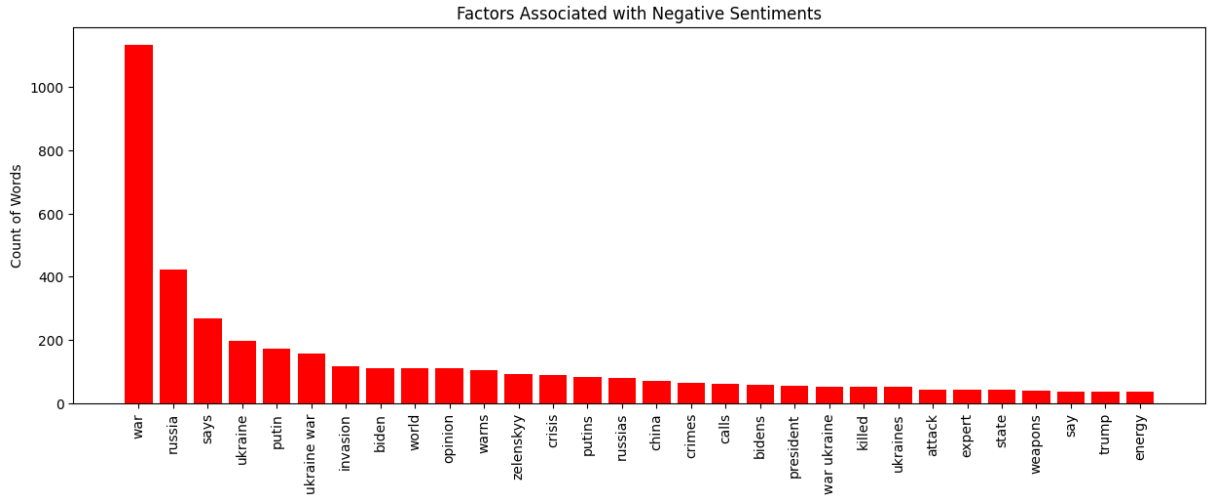


Figure D.4: Fox News Negative Sentiment Words

APPENDIX E: CNN AND FOX BERT AND LDA ANALYSIS

Figure E.1: CNN BERT Topics

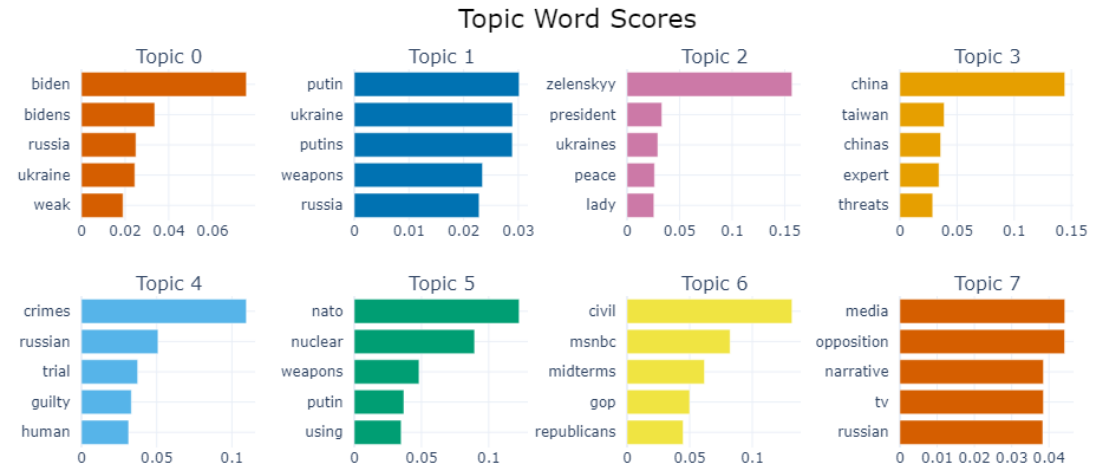


Figure E.2: Fox News BERT Topics

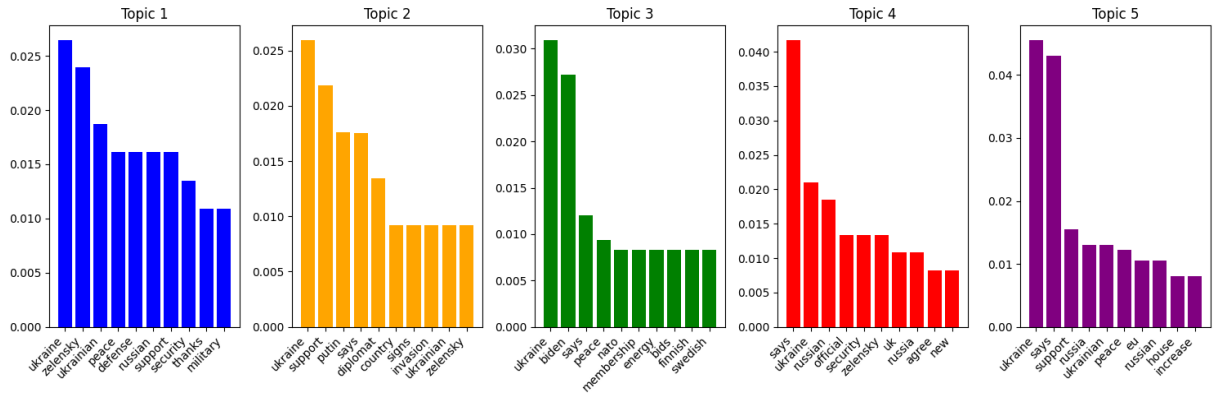


Figure E.3: CNN LDA Topics

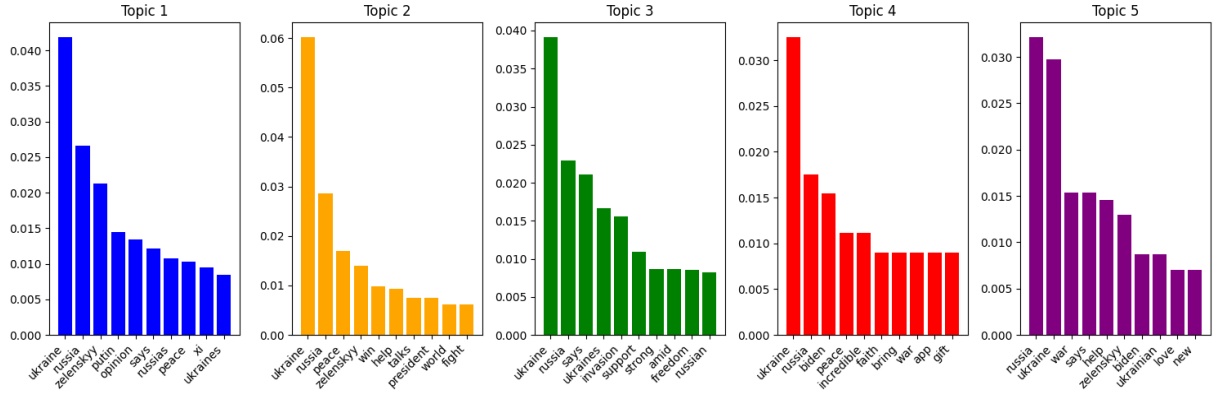


Figure E.4: Fox News LDA Topics