MIE 1624 Final Report

Public Perceptions of the Russia-Ukraine Conflict: Recommendations for Ukraine's Approach

Submitted by

Group 9 for group report

Department of Mechanical & Industrial Engineering, University of Toronto

Date of Submission: April 2, 2023

Course Professor: Oleksandr Romanko

Honour Pledge

By electronically submitting this report I/we pledge that I/we am/are the sole contributor to the written lab report and the work contained in the report is our own group's work. I /we fully understood the consequences of academic misconducts including plagiarism, cheating, copying, sharing etc. as stipulated in UofT Policy.

Name	ID	Date
Mengxuan Chen	1000415482	April 2, 2023
Jianping Zhou	1007702238	April 2, 2023
Chaojun Chen	1004176795	April 2, 2023
Chuyi (Sky) Hou	1004197834	April 2, 2023
Ayesha Patnaik	1008681696	April 2, 2023
Haoyu Li	1009507598	April 2, 2023

Table of Contents

Table of Figures

1 Introduction

In 2021, the Ukraine-Russia conflict continued to intensify with increased hostilities in eastern Ukraine's Donetsk and Luhansk regions. Reports of an elevated Russian military presence near the Ukrainian border raised concerns about a potential large-scale invasion, resulting in urgent diplomatic efforts and increased support for Ukraine from Western countries. Although an all-out war did not ensue, the situation persisted with ongoing skirmishes and casualties as the international community closely monitored the potential for further escalation. [1]

Various nations and individuals have expressed differing opinions on this matter. To gain a deeper understanding of public opinion and the viewpoints of influencers, we employed a data-driven approach that involved analyzing individuals' speeches, posts, and words related to Ukraine, Russia, and the war, as well as influencers' posts and speeches on these subjects [2].

Our project aims to conduct sentiment analysis to help Ukraine’s international image as the invaded country through accessing the public and influencers' sentiments and emotional reactions toward the Ukraine-Russia War. We utilized various techniques such as sentiment modelling, sentiment classification, machine learning-based factor and topic identification, and data visualization to accomplish this goal.

To ensure the credibility of the findings, we limited our research to public opinion and influencers' opinions while disregarding news articles that may be influenced by governments. This approach can potentially increase the odds that the results obtained are a genuine representation of individuals' thoughts and feelings, rather than the official narratives of governments or media entities [3].

Through the implementation of these data science methodologies, we can gain a more in-depth comprehension of the intricate perspectives and emotions of both the public and influencers regarding the Ukraine-Russia War. This information can inform diplomatic efforts, policy-making, and public relations strategies, enabling Ukraine to tailor its communication strategies effectively and garner international attention and support. While sentiment analysis alone cannot guarantee victory in the war, it plays a critical role in informing Ukraine's approach and rallying global support.

2 Background and Methodology

Sentiment Analysis is a tool that helps us determine if a statement or document would be perceived by people as positive or negative in terms of tone, emotion, etc. It's becoming more important in our society because it allows us to understand public opinion by analyzing social media content. During times of political instability, like the recent war in Ukraine, this technology can be especially useful for politicians, media outlets, and organizations to adjust their strategies based on public sentiment.

In this section, our methodology of conducting the study on the Russia-Ukraine war is provided along with a brief technical background for the sentiment analysis.

2.1 Methodology Overview

We have divided the study into the following four sections:

1. **Data Collection and Cleaning:** This is a standard procedure for any data analysis. Text data relating to Russia-Ukraine conflict will be collected and pre-processed before applying any data science methods.
2. **Sentiment Modeling:** This involves creating our sentiment analysis model to establish the sentiment polarity of the text data gathered from various sources.
3. **Sentiment Classification:** The collected text data is classified and quantified according to their sentiment polarity, categorized as either positive, negative, or neutral sentiments.
4. **Factor and Topic Identification via Machine Learning:** Utilizing unsupervised learning techniques such as topic modelling and clustering to identify the key factors and topics that drive the sentiment of the text data.

2.2 Data Collection and Cleaning

The data collection process for this project is straightforward as we are analyzing individuals' tweets and online posts. Additionally, this type of research has been done before, pre-collected datasets can be easily obtained from certified platforms like Kaggle.

Next, we utilized NLTK to process the human language data. NLTK is a well-known platform for creating Python programs that can effectively process our datasets. We used it to tokenize the input text, eliminate stop words, lemmatize the words, and convert text to a bag of words. Additionally, we applied standard text preprocessing techniques using standard Python libraries, including converting text to lowercase, removing URLs, user mentions, numbers, and punctuations. We carried out these procedures to eliminate text features that are irrelevant for sentiment analysis purposes. For instance, numbers do not really reflect emotions, and URLs are also not meaningful.

After these text preprocessing steps, we can train our sentiment models as well as conducting further factor and topic identifications.

2.3 Sentiment Model Training

We have picked four machine learning models for training our sentiment model. They are Logistic Regression, Decision Tree, Random Forest, and XGBoost which are commonly used algorithms for classification tasks.

- Logistic Regression is a statistical model that is used for binary classification problems. It is a simple and efficient algorithm that models the probability of the target variable based on the input features. It estimates the probability of a binary outcome using a logistic function. The logistic function maps any input value to a probability value between 0 and 1. It's a popular algorithm as it is easy to implement and interpret, and it works well with datasets with any size.
- Decision Tree is another popular algorithm that is easy to interpret and visualize. They work by recursively splitting the input feature space into smaller regions based on the values of the input features. The splits are chosen to maximize the information gain or minimize the impurity of the target variable. The final result is a tree structure where each leaf node represents a classification decision. Decision Trees can handle non-linear relationships between the input features and the target variable.
- Random Forest: Random Forest is an ensemble algorithm that combines multiple Decision Trees to improve the classification accuracy and reduce overfitting. Random Forest works by constructing multiple Decision Trees on different subsets of the input feature space and the training data. Each Decision Tree is trained on a randomly selected subset of the input features and a randomly sampled subset of the training data. The final classification decision is made by combining the predictions of all the Decision Trees. Random Forests can handle high-dimensional input spaces, noisy data, and missing values.
- XGBoost is another popular ensemble algorithm that combines multiple decision trees to improve classification accuracy. It works by iteratively constructing decision trees on subsets of the training data and input features. Each new tree is built to correct the errors of the previous trees. XGBoost uses a gradient boosting technique that minimizes the loss function between the predicted and actual values of the target variable. It's a powerful algorithm that can handle high-dimensional data and noisy data.

2.4 Sentiment Classification

The best model out of the four options mentioned previously will be implemented on additional datasets related to the Russia-Ukraine conflict for sentiment classification. Furthermore, we have chosen the Hugging Face Transformer, a pre-existing model, for comparative analysis. Since we have discussed the details on our models, here, we will focus providing the background regarding the Hugging Face Transformer.

2.4.1 Hugging Face Transformers

Hugging Face transformers is an open-source software library that provides advanced language processing capabilities to businesses. Essentially, it is a set of tools that can help businesses to work with large volumes of text data, such as customer feedback, social media posts, news articles, and other unstructured data. Pipeline function will be used to load our sentiment analysis model and sentiment_classifier object will be used to analyze the comments from our dataset and obtain the classified sentiments for public opinion on the Russia and Ukraine conflict [4][5].

2.5 Factor and Topic Identification

As sentiments alone can only provide limited insight, we must identify the factors and topics that influence these sentiments to gain a more comprehensive understanding of the different aspects of the Russia-Ukraine conflict.

The model we have chosen to use to extract the factors and topics is the Genism LDA. It will also help us create easy-to-interpret visualizations such as word cloud.

2.5.1 Gensim library

Gensim is a widely used open-source Python library that is essential for natural language processing tasks. It has become popular for tasks such as document similarity, topic modeling, and information retrieval. The Latent Dirichlet Allocation (LDA) algorithm is one of its key features for topic modeling. LDA is a probabilistic generative model that identifies hidden topics and factors within a large corpus of text. With its advanced NLP capabilities, Gensim has become an important tool in the arsenal of businesses that need to work with textual data [6].

2.5.2 Word Cloud

A word cloud is a data visualization tool used to represent text data and the most frequent words in our dataset could be shown in a larger font size in the graph. In this project, word cloud will be used to highlight the keywords from topics after the top words from each topic is extracted by the LDA model [7].

3 Result and Discussion

This section presents the results of our sentiment analysis for each part of the study detailed in the previous section. Additionally, we offer a comprehensive and candid discussion of our findings.

3.1 Sentiment Model Training

We first preprocess the data to remove unwanted characters such as hashtags, then we use feature engineering techniques to process the data. Refer to Figure 2 Sentiment model training data samples and distributions for the overview of the dataset. Four classification algorithms, including logistic regression, XGBoost, decision trees, and Random Forest, were trained on the training data. Lastly, we evaluated each model on the test data to obtain accuracy measures and performing hyperparameter tuning and cross-validation.

The Logistic Regression has the best and surprisingly good performance, 97% overall accuracy, amongst the four training results, see Table 3.Table 2 Sentiment classification results obtained by deploying our trained logistic regression model and Hugging Face Transformer to the chosen datasets. We will use this trained sentiment model to move forward into the sentiment analysis of their opinion on the Russia Ukraine War.

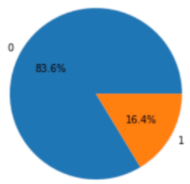
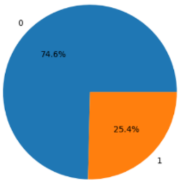
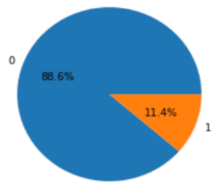
3.2 Sentiment Classification

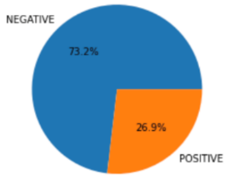
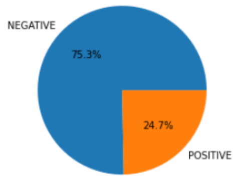
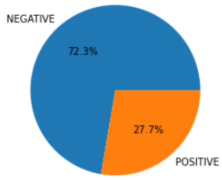
In this part, we deployed our trained Logistic Regression model onto three new datasets which contain mainly posts and tweets from the public. Meanwhile, we used the Hugging Face Transformer [4][5] to classify the same dataset for comparison. The three datasets we selected are listed in the table below.

Table 1 Descriptions of the chosen datasets.

Datasets	Contents
Public response to Elon Musk poll on the subject matter	291348 responses from the public to Elon Musk Tweets
Public tweets on the subject matter	47994 original, non-retweet tweets about Ukraine Russia conflict
Reddit comments on the subject matter	16707 reddit comments about Ukraine Russia conflict

Table 2 Sentiment classification results obtained by deploying our trained logistic regression model and Hugging Face Transformer to the chosen datasets.

	Elon Musk Dataset	Public Tweets	Reddit Comments
Logistic Regression Model			

Hugging Face Transformer			

Based on the above sentiment analysis results, we observed that over 70% of the Russia-Ukraine conflict-related data had a negative sentiment, while roughly 30% had a positive sentiment. We did not verify the consistency across the models, i.e., whether the negative texts classified by one model are also classified as negative by the others, and vice versa. This is because we intend to use only one model to classify public opinions, ensuring consistency is unnecessary as long as the model serves our purpose well. Based on these findings, several potential conclusions can be drawn:

- The conflict between Ukraine and Russia is having a significant negative impact on public's emotions and opinions. The high percentage of negative sentiment indicates that people are likely feeling angry, frustrated, sad, or fearful about the situation.
- The positive sentiment could reflect the public's support for one side or the other. In some cases, people may express positive sentiment towards one side of the conflict because they believe that side is justified or because they have a personal connection to that side.

However, this result is not enough to provide us with a clear picture of how Ukraine is perceived internationally. Hence, we will extract factors and identify topics that are causing the tweets/posts to be positive or negative. We will provide a clearer discussion after we obtain the result in the next section.

3.3 Factor and Topic Identification

Table 4 in appendix summarizes the top topics extracted by LDA model on the three datasets. It shows a clear divide in opinions on this subject matter.

The three datasets are: Elon Musk dataset, Public Tweets dataset, and Reddit dataset. The Elon Musk dataset contains the public replies to the Elon Musk's poll on October 3rd, 2022, shown in the screenshot on the right. The poll has shown that more than 2.7 million people participated in Musk's online poll, and 59.1% of

them were against his peace initiative. The top 5 positive topic from this dataset indicates that people want peace and Ukrainian has the right to make its own choice. Since more than half of the people disagreed with Elon Musk, they argued by a negative language, which is clearly showed in the top 5 negative topics, such as "wtf", "stop the vote".

Moreover, the topics extracted from the public tweet dataset contain some meaningless words such as "el" and "das" since people like using acronyms, slang, and different languages when they post on twitter. Additionally, Reddit dataset top topics show that most people are discussing about military and political rather than express their views on this war. Therefore, we will mainly focus on identifying factors and topics that cause the different sentiments inside the Elon Musk dataset to provide a more accurate analysis.

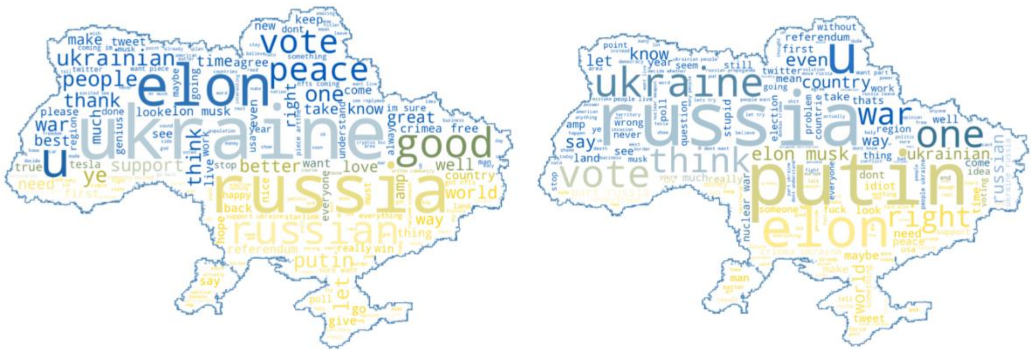


Figure 1 Word Clouds generated from positive (left) and negative (right) posts from the Elon Musk Dataset

From Figure 1, the positive posts contain words such as "peace", "support", and "love", suggesting that the writers are advocating for a peaceful resolution to the conflict and expressing solidarity with Ukraine. Below are two original posts.

"People at war are killing each other every day, on both sides. War is loose-loose. Any peaceful solution is a good solution. All lives matter!"

"Peace is the only sane road."

On the other hand, the negative posts contain words like "Russia", "Putin", "problem", and "war", indicating that the writers are expressing negative sentiments towards Russia and the conflict. The mention of "vote" in negative posts could suggest that these writers believe in the importance of democratic processes and the right of the Ukrainian people to decide their own fate. Here are two posts against the vote by Elon Musk.

"Any "vote" with a gun to your head isn't a vote."

"Russians tortured and killed Ukrainians in the occupied territories. How to vote

in a referendum cynically killed?"

It is worth noting that the words "Ukraine" and "right" appear frequently in both the positive and negative posts, suggesting that the conflict is central to discussions about Ukraine's rights and sovereignty.

Overall, the identified words in the positive and negative posts about the Russia-Ukraine conflict provide some insight into the key themes and public opinions associated with this issue. Some are advocating for peace and support for Ukraine, while others express negative sentiments towards Russia and the conflict.

Note: We have also generated similar word clouds shown above for the other two datasets we collected. Due to the complexity analyzing these figures inside the body of the report, we move the discussions to the Appendix. In addition, we have done other topic importance using LDA model and topic frequencies using bag-of-words. However, the result is not meaningful enough for the analysis. Hence, we put them in the Appendix for reference.

To sum up, sentiment analysis is a valuable tool for understanding people's attitudes and perceptions towards a particular entity, including countries. However, relying solely on a sentiment model may not provide a complete picture of a country's image.

To gain a more detailed understanding of Ukraine's image, it is essential to supplement sentiment models with more direct quantitative data sources. For example, financial data can provide insights into a its economic stability, while data on charity and food can offer a glimpse into the social and cultural values of Ukraine.

Moreover, it is worth noting that sentiment models may not capture more nuanced forms of communication such as sarcasm and irony. These forms of communication are often context-dependent and heavily influenced by cultural nuances, making it difficult for sentiment models to accurately detect and interpret them.

Thus, it is crucial to recognize the limitations of sentiment models and to use a combination of sentiment analysis and direct quantitative data to provide a more comprehensive understanding of Ukraine's image. This approach can aid policymakers, and organizations in making more informed decisions about the country and the public.

4 Recommendation

Based on the results produced by the algorithms, we proposed the following recommendations to the Ukrainian government and NGOs to enhance Ukraine's international presence. First of all, we would like to acknowledge that the existing digital marketing efforts done by the Ukrainian government have been quite successful in the Western countries. From the word cloud, we can see many people are advocating for Ukraine and the word “peace”, “support”, and “love” appear frequently.

Since the beginning of the war, the Ukrainian government has implemented several marketing strategies in the Western world, including digital marketing campaigns and public relations efforts. Detailed information of the war is visible on official websites launched by the Ukrainian government. Additionally, Ukraine has been active on social media and participated in various international events to raise awareness of the ongoing conflict. For example, Zelenskyy participated in an event hosted by University of Toronto President Meric Gertler and the Munk School of Global Affairs & Public Policy to discuss how Canada – and Canadian universities in particular – can support Ukraine’s fight for survival as Russia’s invasion. Zelenskyy also took questions from students following his address. These efforts have helped to improve Ukraine's international presence and attract support from Western countries in economic and military affairs.

In addition to providing war updates, we encourage the Ukrainian government and NGO’s to continue promote Ukraine's culture, economy, and history to the global audience. Ukraine has an official website that features articles, photos, and videos highlighting Ukraine's history, traditions, and modern developments. Ukraine can continue maintaining this website and set up online shop that sell cultural related items such art works, traditional cloths and food. This could help generate funding and support Ukrainian businesses, thereby stimulating the country's economy and supporting local communities.

To improve Ukraine's publicity in the Eastern part of the world, where the general public is less familiar with Ukraine's culture and history, the Ukrainian government and NGOs should increase their digital presence by increase marketing Ukrainian culture, such as its history, traditions, and people. As well as raising awareness about the ongoing conflict and its impact on the country and its people through sharing news articles and social media posts about the Russian invasion in different languages such as Chinese, Indian, and Japanese. Most people in the Eastern countries are non-English speakers, thus, creating contents in their local languages will help raise awareness.

Ukraine already has an Instagram account with good contents and these contents can be translated into different languages to increase coverage. Ukrainian government can also consider producing documentaries, short films, and promotional advertisements that portray the history and culture of Ukraine and engaging social media influencers to produce content showcasing Ukrainian stories. By implementing these recommendations, Ukraine can build a stronger and more positive image on the global stage, promote its culture and values, and secure humanitarian and economic support from individuals across Eastern and Western countries.

The Ukrainian government should also be vigilant against misinformation on social media. With the proliferation of social media, it's become increasingly difficult to distinguish between accurate and misleading information, particularly about the ongoing war in Ukraine. Misinformation can have serious consequences, especially when it comes to shaping public opinion and decision-making. To address this problem, it is important to identify and call out instances of misinformation on social media. One effective strategy is to create posts that specifically target misinformation, using factual evidence and credible sources to debunk false claims. These posts can also include information about where to find accurate information about the conflict in Ukraine. By proactively identifying and addressing misinformation on social media, individuals can help to ensure that accurate information is disseminated and help to promote greater understanding of complex issues.

5 Conclusion

Our study aimed to analyze public opinion on the Russia-Ukraine conflict using sentiment analysis and topic modeling techniques. We collected a large dataset of social media posts and comments related to the conflict and trained a sentiment classifier using machine learning algorithms. The sentiment analysis showed that the majority of the posts had a negative sentiment towards the Russia-Ukraine conflict. Additionally, we used topic modeling to identify the key factors and topics that were discussed in the posts. Our findings revealed that the main topics discussed were political tensions, military actions, and economic sanctions. The study provides valuable insights into public opinion on the conflict, which can be useful for businesses and policymakers in making informed decisions.

Based on our studies, we encourage the Ukrainian government and NGOs' to promote Ukrainian's culture, arts, and music through participating in international events and setting up online shops to market Ukrainian products. Additionally, the government could invest in creating documentaries and promotional videos to raise awareness and distribute these contents in multiple languages. We also recommend Ukraine to form mutually beneficial trade relations with eastern countries and combat misinformation on social media regarding the conflict.

Overall, the sentiment analysis results provide valuable insights into people's emotions and opinions related to the conflict. However, it is important to keep in mind that sentiment analysis has its limitations and should be used in conjunction with other data sources and analysis methods to gain a more complete understanding of the situation.

6 Reference

1. United Nations. (2023) As Russian Federation's Invasion of Ukraine Approaches One Year Mark, Speaker in Security Council Trade Accusations for Failure of Minsk Peace Accords
https://press.un.org/en/2023/sc15202.doc.htm

2. Justina, M., & Tomas, K. (2018) Ukrainian Conflict in Media: Two Approaches to Narrative Analysis
https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-178/MP-IST-178-04.pdf

3. Hamed, V., Mohammad, G., Fatemah, S., Faezeh, A., & Hamid-Reza, N. (2023) Russian-Ukraine war: Modeling and Clustering the Sentiments Trends of Various Countries
https://arxiv.org/pdf/2301.00604.pdf

4. Jain, S. M. (2022). Hugging Face. In Introduction to Transformers for NLP. Apress L. P. https://doi.org/10.1007/978-1-4842-8844-3_4

5. Hugging Face. (n.d.). Pipelines. Retrieved March 31, 2023, from https://huggingface.co/docs/transformers/main_classes/pipelines

6. Dwivedi, P. (2019, March 27). NLP: Extracting the main topics from your dataset using LDA in minutes. Medium. Retrieved March 31, 2023, from https://towardsdatascience.com/nlp-extracting-the-main-topics-from-your-dataset-using-lda-in-minutes-21486f5aa925

7. Kibe, K. (2023, February 21). Topic modeling using latent Dirichlet allocation (LDA). Analytics Vidhya. Retrieved April 1, 2023, from https://www.analyticsvidhya.com/blog/2023/02/topic-modeling-using-latent-dirichlet-allocation-lda/

7 Appendix

	ID	text	label
550386	8.046170e+17	@goddesses_o I can't stop watching her...mm. M...	0
550387	8.046180e+17	Poor old Tom Odell doesn't look like he would ...	0
550388	8.046180e+17	#antsmasher I smashed 7 ants in this awesome ...	1
550389	8.046180e+17	@LizHudston @KymWyllie @Evasmiless @meanBok @l...	1
550390	8.046190e+17	Bixbeat Mixtape Vol.2 is here with great artis...	1

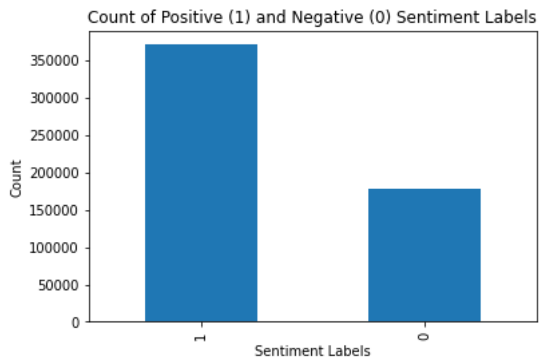


Figure 2 Sentiment model training data samples and distributions

Table 3 Sentiment model training results

Models	Training Results	precision	recall	f1-score	support
Random Forest	0	0.78	0.93	0.85	35854
	1	0.96	0.87	0.92	74225
	accuracy			0.89	110079
	macro avg	0.87	0.90	0.88	110079
	weighted avg	0.90	0.89	0.90	110079
Decision Tree	0	0.53	0.89	0.67	35854
	1	0.92	0.62	0.74	74225
	accuracy			0.71	110079
	macro avg	0.73	0.76	0.70	110079
	weighted avg	0.79	0.71	0.72	110079
XGBoost	0	0.95	0.19	0.31	35854
	1	0.72	1.00	0.83	74225
	accuracy			0.73	110079
	macro avg	0.83	0.59	0.57	110079
	weighted avg	0.79	0.73	0.66	110079
Logistic Regression	0	0.95	0.96	0.96	35854
	1	0.98	0.98	0.98	74225
	accuracy			0.97	110079
	macro avg	0.97	0.97	0.97	110079
	weighted avg	0.97	0.97	0.97	110079

Table 4 Topic importance using LDA Model on the chosen datasets

<table>
<tr><td>Replies to Elon Musk Poll</td><td>Top 5 topics that drive positive sentiment:
Topic 1: russia | want | ukraine | peace | go | see | think | sure | wtf | elon
Topic 2: people | ukraine | russian | russia | get | ukrainian | good | vote | like | would
Topic 3: elon | money | russia | ukraine | make | war | like | right | dosa | poll
Topic 4: crimea | russia | ukraine | vote | russian | donbas | idiot | return | war | fair
Topic 5: people | ukrainians | russian | right | ukraine | best | peace | propaganda | see | us

Top 5 topics that drive negative sentiment:
Topic 1: ukraine | votes | people | stop | russia | russian | back | let | like | putin
Topic 2: russia | people | vote | stop | war | years | russian | dont | elon | would
Topic 3: fuck | elon | go | russia | ukraine | russian | people | also | want | putin
Topic 4: ukraine | crimea | would | elon | one | russia | alaska | people | vote | give
Topic 5: russia | let | part | country | give | elon | get | us | would | people</td></tr>
<tr><td>Public Tweets</td><td>Top 5 topics that drive positive sentiment:
Topic 1: affiliate | business | twitter | deal | ad | shop | shopping | gift | doge | blame
Topic 2: russia | war | people | like | us | day | one | get | would | china
Topic 3: de | la | en | di | un | el | e | con | se | che
Topic 4: russia | gas | via | standwithukraine | china | video | finland | government | free | west
Topic 5: long | create | term | income | awesome | hire | passive | affiliate | autopilot | developer

Top 5 topics that drive negative sentiment:
Topic 1: military | us | world | region | russia | ukraine | aid | training | air | ground
Topic 2: ukraine | ukrainerussiawar | russian | ukrainewar | standwithukraine | front | august | ukrainian | russiaisaterroriststate | russia
Topic 3: die | ist | das | den | es | der | hat | und | wir | sie
Topic 4: russia | china | nuclear | ukraine | war | news | plant | president | power | world
Topic 5: china | russiaisaterroriststate | u | xi | na | seen | join | done | life | plan</td></tr>
<tr><td>Reddit Comments</td><td>Top 5 topics that drive positive sentiment:
Topic 1: twitter | status | artillery | news | russia | last | border | like | air | live
Topic 2: r | https | please | com | watch | post | message | www | bot | video
Topic 3: russia | would | even | nuclear | also | war | could | military | like | think
Topic 4: trump | us | would | like | used | support | also | even | actually | ready
Topic 5: war | like | na | agree | gon | us | go | people | cold | stay

Top 5 topics that drive negative sentiment:
Topic 1: soviet | hear | people | genocide | union | love | action | day | killing | comment
Topic 2: one | like | removed | war | time | got | see | saying | enemy | know
Topic 3: russia | war | people | us | like | world | think | know | one | want
Topic 4: must | good | russia | country | one | give | nothing | language | like | black
Topic 5: ukraine | russia | russian | nato | putin | us | eu | help | military | russians</td></tr>
</table>

7.1 Additional Word Clouds

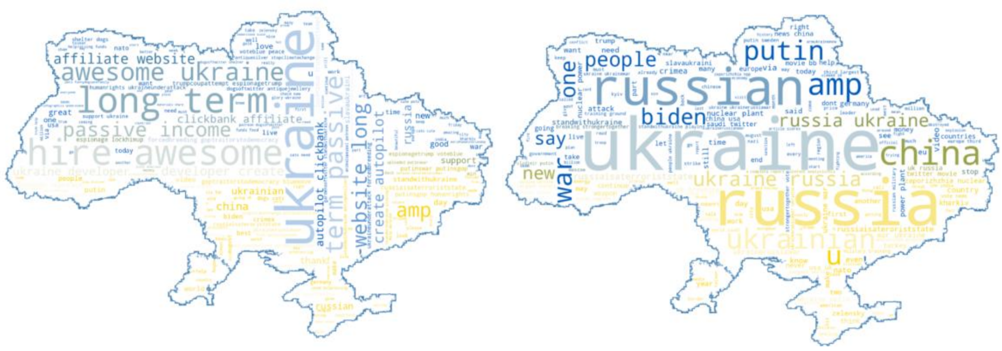


Figure 3 Word Clouds generated from positive (left) and negative (right) posts from the Public Tweets about Russia-Ukraine War

Figure 4 Word Clouds generated from positive (left) and negative (right) posts from the Reddit Comments about Russia-Ukraine War

This section provides additional explanations on the word clouds generated for the two datasets that were not included in the factor and topic identification section. The rationale for their exclusion is that these results offer a lesser degree of information than those included in the body of the report.

Moreover, the public tweets about the Russia-Ukraine war dataset contains a considerable number of advertisements, which have significantly skewed the analysis. Consequently, we decided against their inclusion in the analysis. However, it is noteworthy that Ukraine and Russia remained the central topic in all three datasets, and we could still identify some familiar themes that appeared in our analysis in these word clouds on the appropriate sentiment side.