Measuring Public Opinion on Russia-Ukraine War Through Analysis of Tweets

Avish Parmar 112647892
Dukyoung Eom 111309287
Woohyung Lee 110929853
Yejin Lee 115073751

1. Introduction

RUUK's (**RU**ssian **UK**raine research group) research topic is to present a relationship between the Russia-Ukraine war and the levels of discomfort globally through analysis of tweets over time in correlation to the events of the war. Our topic is important because it provides concrete evidence of how the war has affected populations outside of the conflict zone. This evidence, in turn, would give countries a more personal incentive to put an end to the conflict through diplomatic means. One should care about this topic because an overseas conflict may affect your life as it leads to increasing levels of discomfort, creating a more hostile environment for you and your family to live in.

`The SDG goal we are attempting to cover is Peace, Justice, and Strong Institutions, and more specifically targeting the following:

1. Significantly reduce all forms of violence and related death rates everywhere
2. Ensure responsive, inclusive, participatory and representative decision-making at all levels.

We are covering the first SDG goal by identifying spikes in tweets with a negative sentiment, which serve as early warning signs for potential escalations in crime rates from the area where the tweet originated. We are covering the second SDG goal by providing policy-makers with information regarding public opinion through the sentiments of the tweets from different users such that they can better represent their constituents.

A technical challenge that we attempted to solve was filtering out tweets that are not related to the Russia-Ukraine war, such as ads, that use hashtags for publicity. Although not a fix-all strategy, our approach to address this was to randomly sample each of the 13 datasets associated with each month of the conflict to minimize the chances of getting unrelated tweets.

2. Background

Social networks such as Twitter have served as a platform where anyone can express their opinion about current events. With its growing popularity there have been many studies that have used tweets to gauge public opinion regarding an event. One that we particularly drew interest from is Russia-Ukraine War: Modeling and Clustering the Sentiments Trends of Various Countries (2023) which conducted sentiment analysis on the Russia-Ukraine tweet dataset to better understand the public sentiment by country in response to the war. However, there were several limitations to this study:

1. The study covered only the first month of the conflict
2. The dataset that was used was quite small; 140,000 observations of raw data which was reduced further once it was cleaned to include just the tweets with geolocation data.

This is where we saw an opportunity to further the study conducting sentiment analysis on a much larger dataset spanning from the start of the war until April 30, 2023. By doing this we would be able to more efficiently gauge the public sentiment associated with the war given the holistic aspect provided by the dataset.

Since analyzing millions of Twitter comments individually was deemed unrealistic, we utilized the BERT language model to perform sentiment analysis on the preprocessed Twitter data. This allowed us to classify the comments as positive or negative towards the Russia-Ukraine conflict. Additionally, we employed an LDA model to grasp the public's opinion on the progression of the Russia-Ukraine conflict on specific dates. Through this approach, we can plot the support rate for the Russia-Ukraine conflict by country on a world map over time, and obtain opinions on specific time periods (events) using the LDA model.

3. Data

We used the 'Ukraine Conflict Twitter Dataset' and 'Geolocation Data' from Kaggle as our primary datasets and "Twitter Tweets Sentiment Dataset" to conduct a t-test of the model. The 'Ukraine Conflict Twitter Dataset' is in CSV format and has a size of about 40GB. This dataset has well-defined labels such as language, hashtags, and retweet counts. It includes not only comments but also information on the date of the hashtags and the user-defined location where the tweets were posted, enabling us to conduct time-series analysis and analyze the sentiment toward the Russia-Ukraine conflict by country. However, there were some issues with the values in the comment column. For example, there were useless datasets such as spam or advertising comments, and in some cases, the characters were broken and unreadable due to different fonts or unreadable languages, requiring some preprocessing for Twitter comments. We needed to be mindful of these issues.

	Unnamed: 0	userid	...	quoted_status_username	extractedts
0	0	529588633	...	NaN	2022-05-01 09:12:42.918515
1	1	835922746416762880	...	NaN	2022-05-01 00:21:48.021996
...	...	...	...	...	...
351885	351885	345096523	...	NaN	2022-05-02 00:18:47.954457
351886	351886	982538740571795456	...	NaN	2022-05-02 00:13:27.035208

The 'Geolocation Data' is a dataset we used to represent the emotional distribution by country. The size of this dataset is approximately 4MB. This dataset had well-defined labels such as latitude and longitude for each country. Additionally, it was well-classified not only by country names but also by states, allowing us to create more detailed plots of emotional distribution when combined with the location column of the 'Ukraine Conflict Twitter Dataset'. We merged cities, states, countries dataset manually so that a single csv file contains information for all the cities, states and countries. We utilized this dataset to identify more accurate location information from user-defined locations.

	city_name	city_latitude	city_longitude	state_code	state_name	state_latitude	state_longitude	country_code	country_name	country_latitude	country_longitude
0	Ashkāsham	36.68333	71.53333	BDS	Badakhshan	36.734772	70.811995	AF	Afghanistan	33.0	65.0
1	Fayzabad	37.11664	70.58002	BDS	Badakhshan	36.734772	70.811995	AF	Afghanistan	33.0	65.0
...	...	...	...	...	...	...	...	...	...	...	...
148064	Shurugwi District	-19.75000	30.16667	MI	Midlands Province	-19.055201	29.603549	ZW	Zimbabwe	-20.0	30.0
148065	Zvishavane District	-20.30345	30.07514	MI	Midlands Province	-19.055201	29.603549	ZW	Zimbabwe	-20.0	30.0

The Twitter Tweets Sentiment Dataset is the dataset we used for hypothesis testing the transformer model. The size of this dataset is almost 4 MB with 27481 observations and 4 features: [textID, text, selected_text, sentiment]. This dataset was labeled as it contained the sentiments positive, negative, and neutral for each text (tweet content).

	textID	text	selected_text	sentiment
0	cb774db0d1	I`d have responded, if I were going	I`d have responded, if I were going	neutral
1	549e992a42	Sooo SAD I will miss you here in San Diego!!!	Sooo SAD	negative
...	...	...	...	...
27479	ed167662a5	But it was worth it ****.	But it was worth it ****.	positive
27480	6f7127d9d7	All this flirting going on - The ATG smiles. Yay. ((hugs))	All this flirting going on - The ATG smiles. Yay. ((hugs)	neutral

4. Methods

4.1 Preprocessing with PySpark

Before we conduct the Sentiment analysis and LDA analysis we had to preprocess the tweet data since it is raw texts with urls and emojis. For bert, we first cleaned up urls, emojis and smileys from the text and removed potential html tags from the text. For LDA, we converted all the text to lowercase, removed urls, mentions, hashtags, numbers, emojis, smileys and punctuations from the text (hashtags are stored in the separate column). Then we removed stop words from the text and lemmatized leftover words. For both preprocessing steps, we commonly applied location matching which matches the user-defined location to the Geolocation data to identify accurate location information. Also based on the followers, followings, favorites and retweets, we computed popularity score and reach score of each tweet.

4.2 Sentiment Analysis with BERT using PyTorch

We used PyTorch for this part of our project. For sentiment analysis of the tweet content we had initially planned on using bert-base-multilingual-cased to account for tweets but upon closer examination we found that a significant majority of the dataset was in english. Therefore as a more efficient approach we decided to filter out non-english tweets and focused on analyzing only the english tweets. This also allowed us to use a more time-efficient model, namely distilbert-base-uncased-finetuned-sst-2-english. This particular version of DistilBERT was trained on the Stanford Sentiment Treebank dataset which consists of single sentences extracted from movie reviews.

To make use of the model we used the pipeline API which allows us to abstract out the complex code associated with transformers. The hyperparameters of distilbert-base-uncased-finetuned-sst-2-english are the following: [Number of layers : 6] [Number of hidden layers : 768] [Number of head : 12] [Number of parameters : 66M] [Number of workers : 8] [Batch size : 32] [Learning rate : 1e-5 (0.00001)] [Warmup : 600 steps] [Max sequence length : 128] [Number of train epochs : 3] [Activation function: softmax] We executed the code in google colab with T4 GPU.

We first instantiated the classifier by using the pipeline function to specify the analysis type and model, and then ran it on 500,000 randomly sampled tweets in over 13 datasets to determine the sentiment of the tweet and its sentiment score. The tweets were randomly sampled only if the dataset had more than 500,000 tweets. The sentiments were then grouped together based on country with the use of the geolocation dataset and plotted into a line graph to showcase the public sentiment through each month. We also created a heat map using the geolocation dataset which showcased the overall public sentiment spanning the entire time interval covered by the dataset.

4.3 LDA(Latent Dirichlet Allocation) Topic modeling

We divided data monthly, and extracted topics using one of the topic modeling tools, LDA, to identify trends of public opinion throughout time. We ran this model using a CPU of max 13GB RAM. LDA(Latent Dirichlet Allocation) extracts ‘hidden topics’ or themes from a large corpus of text data. It is an unsupervised learning method that identifies patterns and clusters of words that often co-occur in a set of documents. Topic modeling algorithms aim to discover latent topics in text by analyzing the distribution of words within documents, as well as the co-occurrence patterns of words across documents. The output is a set of topics, where each topic represents a cluster of words that frequently occur together in the document. Using this, we extracted 10 topics from each month, and examined the variation of it, and associated it with major events during war.

5. Evaluation/Results

After conducting the sentiment and LDA analysis we were able to determine the overall sentiment associated with each country as well as the topic that was trending for each month. The positive sentiment accuracy was 0.9390526953200334 and the negative sentiment accuracy was 0.9600455731667898. Through our analysis, we were able to determine that most countries showcased a negative sentiment. Upon analysis of the results we found that there were some irrelevant tweets such as ads that were included as part of the random sample. Thus there were some instances where these tweets were marked as positive/negative although they did not relate to the war. The general pattern observed was that there were spikes of negative sentiment for each country at the start of the war after which the sentiments began to bottom-out, all while the negative sentiment rate being above the positive sentiment rate. An example output of the sentiment analysis is the following:

	id	date	text	hashtags	sentiment	sentiment_score
0	2	11/1/2022 0:00	@apmassaro3 @IlvesToomas France Spain and Italy... you what? RussiaIsLosing russiaisateroriststate	['RussiaIsLosing', 'russiaisateroriststate']	NEGATIVE	0.983655
1	3	11/1/2022 0:00	@RishiSunak we are NOT being invaded. Apologise to us and more importantly, apologise to President Zelensky ; the brave people of Ukraine - who ARE being invaded. SackBraverman HateSpeech Conservatives	['Zelensky', 'Ukraine', 'SackBraverman', 'HateSpeech', 'Conservatives']	POSITIVE	0.985423

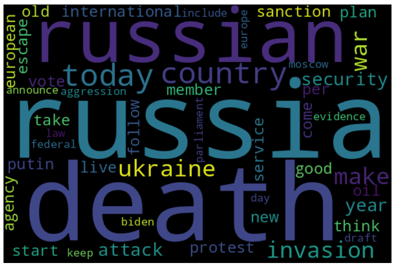
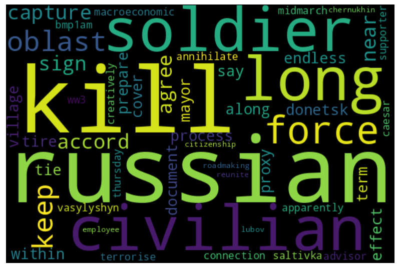
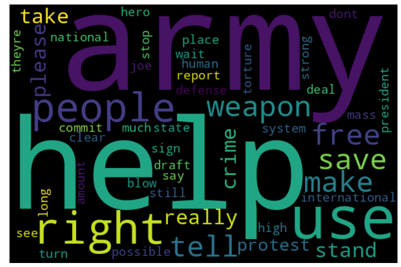
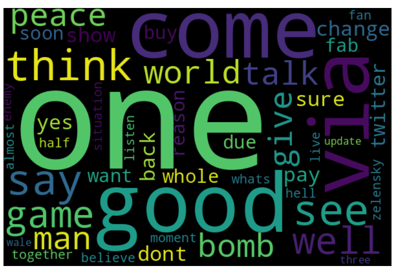
Some unique aspects of the project such as a global sentiment on the war is showcased through the heat map as well as the line graph, along with a case-by-case sentiment breakdown of some countries.

For the result, we conducted two hypothesis tests to evaluate the performance of our sentiment analysis model and to assess the world's sentiment regarding the Russia-Ukraine conflict. For the first test, our objective was to determine the accuracy of our sentiment analysis model. After sentiment analysis, we compared the model's accuracy using 5-fold cross-validation with the sentiment results obtained from our model. In addition, we compared the average accuracy of our sentiment analysis model with the accuracy of randomly labeled values. We assumed a null hypothesis stating that "there is no difference in accuracy between our model and a randomly labeled model" and conducted a t-test with a significance level of 0.05. The result of the t-test was the t-statistic value being higher than the critical value, leading to the rejection of the null hypothesis; confirming that our model performs better than random classification.

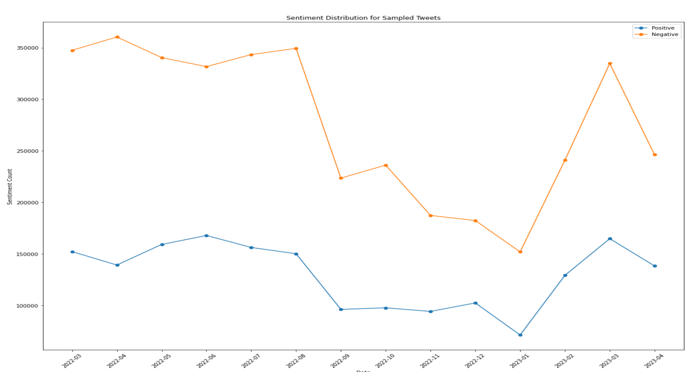
For the second test, we aimed to determine the world's sentiment concerning the Russia-Ukraine conflict. We selected 20 countries with the largest data volume and assigned a value of 1 for positive opinions and -1 for negative opinions. We established a null hypothesis stating that "users in these countries are not negative about the Russia-Ukraine conflict". With a significance level of 0.05, we first obtained the sentiment index for each country. We then conducted a t-test to determine if the sentiment index was not higher than 0.5 (the minimum level for positive sentiment). The null hypothesis was rejected in 70% of the 20 countries, allowing us to confirm that the majority of the world holds a negative opinion about the Russia-Ukraine conflict.

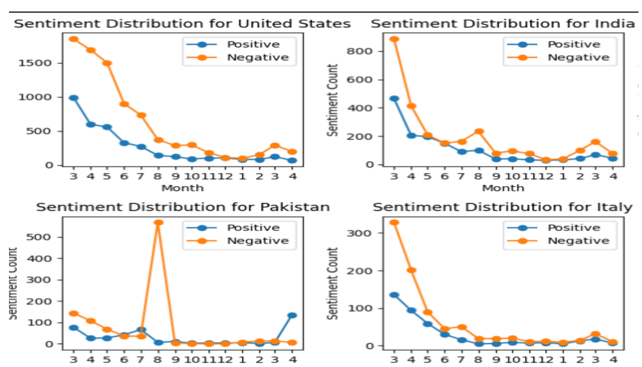
We were able to identify that topics from LDA analysis have a correlation with the results from sentiment analysis. For example, from March 2022 to June 2022, the Russian army was predominant and their invasion brought the largest number of casualties. During this time, word clouds generated from LDA topic output emphasizes ‘death’, ‘kill’, ‘russia’. Also, negative sentiment during this time was at its zenith. However, as the Ukrainian army fired back to Russia and started to retrieve their region, word clouds showed ‘help’, ‘one’, ‘good’ and the negative sentiment dropped drastically. Furthermore, as Ukraine got support from their nation and other countries, Ukraine's diplomatic standing has improved and tweets about the war show these; ‘light’, ‘mean’, ‘remember’. Along with that, negative sentiment aroused but the positive sentiment put a definite point. Regarding the increase of negative figures could be derived from the increased interest towards the war itself(considering the amount of tweets tripled after February than previous 5 months; when Biden visited Ukraine), we can conclude that the topics are correlated with sentiment.

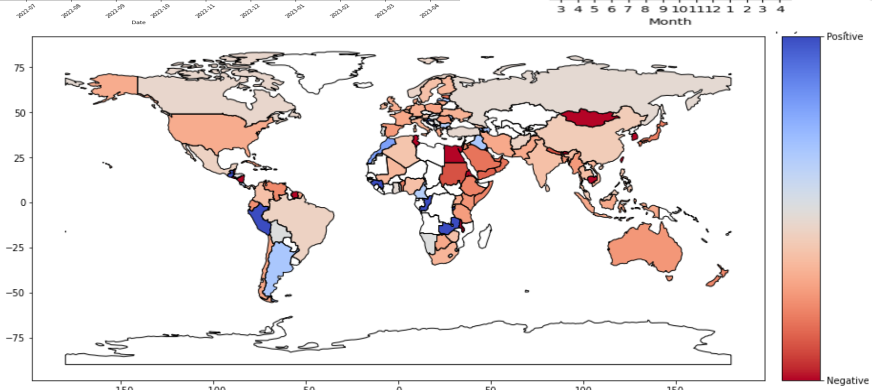
Below is the 1st topic of each month; which has the biggest proportion and thus is the most representative.

<table>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>March 2022
Russian invasion started. Largest number of casualties in a month</td>
<td>June 2022
Russia's fierce invasion continued for 4 months; worst public order in ukraine</td>
<td>September 2022
Ukraine launched a counteroffensive and retook some region</td>
</tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>November 2022
Significant victory for Ukraine.
Russian forces withdrew from Kherson to the eastern side of the Dnipro River.</td>
<td>February 2023
US President Joe Biden made an unannounced visit to Kyiv, followed by other country's presidents</td>
<td>April 2023
Global efforts toward a diplomatic solution to the war in Ukraine</td>
</tr>
</table>

Figures below are Sentiment distribution from random 500000 samples each month and Sentiment distribution across the globe







6. Conclusion

Our project aimed to measure the public opinion and perception through the analysis of tweets regarding the ongoing Russia-Ukraine conflicts. Based on the preprocessed datasets, we conducted sentiment analysis with a pre-trained DistilBERT model to observe public sentiment on this issue. Among 15M English tweets, we randomly sampled about 6M tweets (500,000 from each month). Through our analysis we found strong evidence that the majority of the public on twitter has a negative aspect and increasing anxiety towards the conflict. The rate of negative sentiment varied on each country but, the sentiment distribution across the globe indicated that the population in the majority of countries had a negative viewpoint on this matter. With LDA analysis based on monthly tweet dataset we were able to demonstrate the trending topics and words in time series. Combining this with the results of sentiment analysis we were able to showcase topics that were driving the sentiment of the public, allowing us to find a direct correlation between the events of the war and public sentiment. Thus, our analysis provides concrete evidence that the Russia-Ukraine conflict has had global ramifications and that policy-makers should push for diplomatic solutions to put an end to the conflict.

7. References:

1. Batterywala, Alifiya. "Ukraine vs Russian Topic Modeling." Kaggle, www.kaggle.com/code/alifiyabatterywala/ukraine-vs-russian-topic-modeling.
2. Bwandowando. "Ukraine-Russian Crisis Twitter Dataset (1.2 M Rows)." Kaggle, www.kaggle.com/datasets/bwandowando/ukraine-russian-crisis-twitter-dataset-1-2-m-rows.
3. Brown, Eva & Schwartz, Lindsey & Huang, Richard & Weber, Nicholas. (2023). Soft-Search: Two Datasets to Study the Identification and Production of Research Software. 10.48550/arXiv.2302.14177.
4. Liew, You Sheng. "Geolocation." Kaggle, www.kaggle.com/datasets/liewyousheng/geolocation?select=countries.csv.
5. Odeyemi, G. "Twitter Data Mining‐Measuring Users' Influence." Towards Data Science, Medium, 27 July 2020, towardsdatascience.com/twitter-data-mining-measuring-users-influence-ef76c9badfc0.
6. Schmid, Phil. "K-Fold as Cross-Validation with a BERT Text Classification Example." Phil Schmid's Blog, www.philschmid.de/k-fold-as-cross-validation-with-a-bert-text-classification-example.
7. Socher, Richard, et al. "Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank." ACL Anthology, aclanthology.org/D13-1170/.
8. "DistilBERT Base Uncased Finetuned SST-2 English." Hugging Face, huggingface.co/distilbert-base-uncased-finetuned-sst-2-english. Accessed 14 May 2023.
9. "Main Classes - Pipelines." Hugging Face, huggingface.co/docs/transformers/v4.29.1/en/main_classes/pipelines#transformers.TextClassificationPipeline.
10. Vahdat-Nejad, H., Akbari, M. G., Salmani, F., Azizi, F., & Nili-Sani, H.-R. (2023, January 2). Russia-ukraine war: Modeling and clustering the sentiments trends of various countries. [2301.00604] Russia-Ukraine war: Modeling and Clustering the Sentiments Trends of Various Countries. https://arxiv.org/ftp/arxiv/papers/2301/2301.00604.pdf
11. Yasser, H. "Twitter Tweets Sentiment Dataset." Kaggle, www.kaggle.com/datasets/yasserh/twitter-tweets-sentiment-dataset?resource=download.