**Automatic Twitter Event Summarization: Making Sense of War in Ukraine**

Aishwarya Kale and William Simpson

Dept. of Computer Science & Electrical Engineering, University of Maryland, Baltimore County

DATA 690: Introduction to Natural Language Processing

Dr. Antonio Diana

May 6, 2022

**Introduction**

On Feb. 24, 2022, Russia launched a full-scale military invasion of Ukraine, which to date has resulted in nearly 3000 civilian casualties, killed thousands of soldiers, and displaced approximately 13 million Ukrainians (Matthews et al. 2022; Council on Foreign Relations, 2022). Boycotts, as well as disruptions in access to Russian energy exports, have caused gas and oil prices to soar, tipping the world economy into crisis (Kammer et al., 2022). Information surrounding this global event is surging, particularly on Twitter. For example, the daily number of tweets discussing Russian President Vladimir Putin increased 5-10 times the level prior to the invasion (Pham & Talavera, 2022). The conflict is developing, and its impacts are expanding each day as we see new sanctions, nuclear-threat narrative escalation, Russian army advances, attempts at mediation, etc. Although this complexity makes monitoring this event difficult, doing so is crucial to ensure leaders and decision-makers are well informed about the conflict so they can respond accordingly. Well-informed decisions in turn save lives and stabilize the situation.

Utilizing automatic text summarization, this project presents a web-based analysis tool that is designed to alleviate the difficulty of filtering through the immense volume of data about the war in Ukraine to keep track of major developments. Based on existing research on event summarization, we explore the capabilities of named entity recognition (NER), topic modeling, clustering, and sentiment analysis methods to help identify important sub-events for automatic summarization. Compared to a timeline of actual events, we find that our primarily unsupervised summarization system successfully identified important sub-events within the war in Ukraine.

**Background Research**

Text summarization is the process of condensing one or many documents of text into a shorter version while aiming to preserve the original text’s primary meaning. The dominant approaches to summarization in the literature are extractive and abstractive summarization. The extractive method focuses on identifying the most important and representative sentences in the corpus by ranking them according to a metric, such as a similarity or distance (Sharifi et al., 2010, Rudrapal et al., 2018). On the other hand, abstractive summarization creates a representation of the key information in the corpus (e.g., word embedding) and then leverages text generation techniques to produce a summary that is not bound to the vocabulary that explicitly appears in the source, unlike the extractive approach (Rudrapal et al., 2018).

Research on methods of automatic text summarization traditionally rely on corpora of news articles, such as the TAC (Text Analysis Conference) or CNN/DailyMail datasets (Shen et al., 2013). However, the past decade has seen an increase in utilizing Twitter and other microblog or social media data for summarization. Although the short length, informal grammar, and variety of content in tweets present challenges for summarization tasks, we choose to use Twitter data in this project since it provides a unique perspective to obtain an unfiltered understanding of the war in Ukraine through the eyes of the public. Furthermore, analysis of tweets pertaining to the war allows for fine-grain or near real-time summarization of a global event.

Event summarization is a subset of the broader body of text summarization literature. Distinct from event detection, event summarization centers on determining and monitoring major participants/entities as well as important sub-events during the defined event period (Lee et al., 2021). There are two defining characteristics of existing Twitter event summarization research:

1. A majority of studies aim to summarize sports events (Nichols et al., 2012; Corney et al., 2014; Jai-Andaloussi et al., 2015; Esmin et al., 2014; Huang et al., 2018), or other small-scale events with relatively short timelines (Shen et al., 2013; Alsaedi et al., 2017);
2. Many rely on analysis of spikes in the volume of tweets to pinpoint important sub-events for summarization (Marcus et al., 2011; Shen et al., 2013; Nichols et al., 2012).

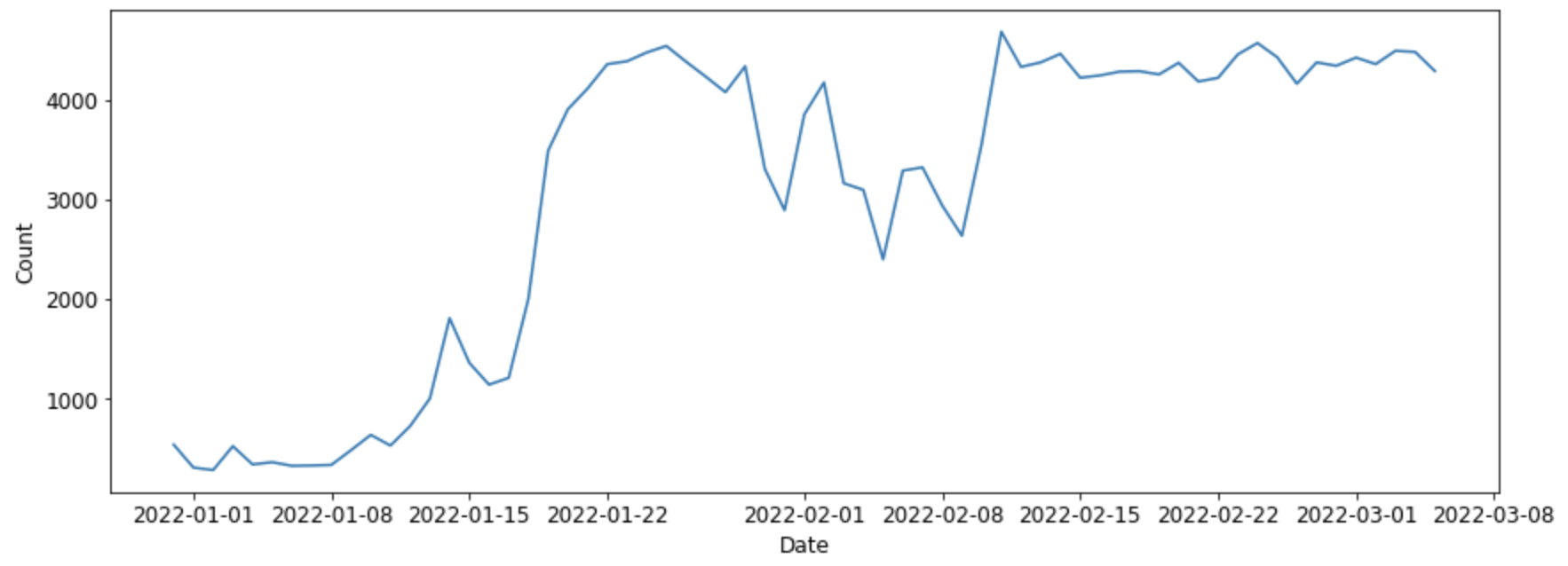
Lee et al. (2021) conducted one of the only studies to attempt summarization of a large-scale event—the COVID-19 pandemic in early 2020. They used a pre-trained T5 (Text-to-Text Transfer Transformer) model to produce a coherent storyline of COVID-19-related events in early 2020. In general, this body of research has seen a shift from extractive statistical techniques to abstractive neural networks and finally to pre-trained transformer models such as T5 or BERT (Bidirectional Encoder Representations from Transformers) (Li & Zhang, 2021; Dusart et al., 2021, Lee et al. 2021). Thus, we also concentrate on pre-trained transformers.

Based on the state of the literature, we selected to analyze Twitter data pertaining to the ongoing war in Ukraine. Applying automatic summarization techniques to such a large-scale global event merits further research because it is understudied and because of its potential as a tool for efficient and concise information sharing. While tweet spike analysis helps summarize small-scale events, tweet volume is consistently high throughout large long-term events such as the Russia-Ukraine War. Therefore, our analysis also aims to investigate the viability of NER, topic modeling, clustering, and sentiment analysis techniques to identify entities and topics within a large event, as alternatives to pinpointing sub-events for summarization.

**Data Description and Preparation**

While a Twitter stream pipeline would enable near real-time analysis of the war in Ukraine, to test our approach, we selected a Kaggle dataset containing 65 days of tweets pertaining to this conflict from Dec. 31, 2021, to March 5, 2022 (Purtova, 2022). The tweets were originally collected using the keyword “Ukraine war” and ranged from 500 to 5000 tweets per day (see figure 1). Features of the data include URL, date, content, user ID, links, hashtags, and counts of replies, retweets, and likes.For data preparation, we first filtered out tweets in languages other than English and dropped duplicate tweets. Data wrangling was performed by cleaning the tweets to remove URLs, Twitter handles, punctuation, and emojis. Tweets were tokenized and lemmatized. For some analyses, tweets are aggregated by 1-2 days.

**Figure 1**

*Volume of Tweets per Day*

**Methodology**

The volume and velocity of the content generated on Twitter necessitates a rigorous manual filtration process and extensive reading effort to understand events and public sentiment. Thus, we propose to leverage NER, topic modeling, clustering, and sentiment analysis in a primarily unsupervised summarization task of Twitter data to achieve the following objectives:

1. Retrieve the most important information, and
2. Generate a robust event summary.

In the first part of this study, we implemented NER using spaCy to identify major players in the war in Ukraine. Topic modeling with Non-negative Matrix Factorization (NMF) and BERTopic illustrates topics pertaining to sub-events within the corpus of tweets. Clustering using network graph analysis and transformers partitions tweets into categories of similar aspects of the war. Sentiment analysis with VADER Sentiment Analyzer and Textblob tracks fluctuation in public emotional response to new developments.

In the second part of our project, traditional extractive and abstractive summarization techniques are supplemented by the information about important sub-events that we gain from the first part. Extractive summaries are generated using Gensim, spaCy, and SUMY which implement TextRank and Latent Semantic Analysis summarization algorithms. The abstractive text summarization of tweets data is generated using a pre-trained T5 model and BART (Bidirectional and Auto-Regressive Transformer). We evaluate our summarization system’s performance with similarity metrics (i.e., cosine similarity) and ROUGE-1 scores. By leveraging methods to determine main sub-events to inform our text summarization techniques, we hope to create a tool to paint a complete picture of the Ukraine-Russia war through the eyes of the public.

**Study Outcomes**

**Part 1: Sub-event and Main Idea Detection**

***Named Entity Recognition (NER)***

Identifying entities like organizations and people involved in an event is necessary to intuitively understand the text overall. NER helps us get a gist of the text summary at a glance. Figure 2 uses NER with spaCy to depict relevant organizations and countries from tweets on Dec. 31, 2021. This provides readers with a method for automated information extraction.

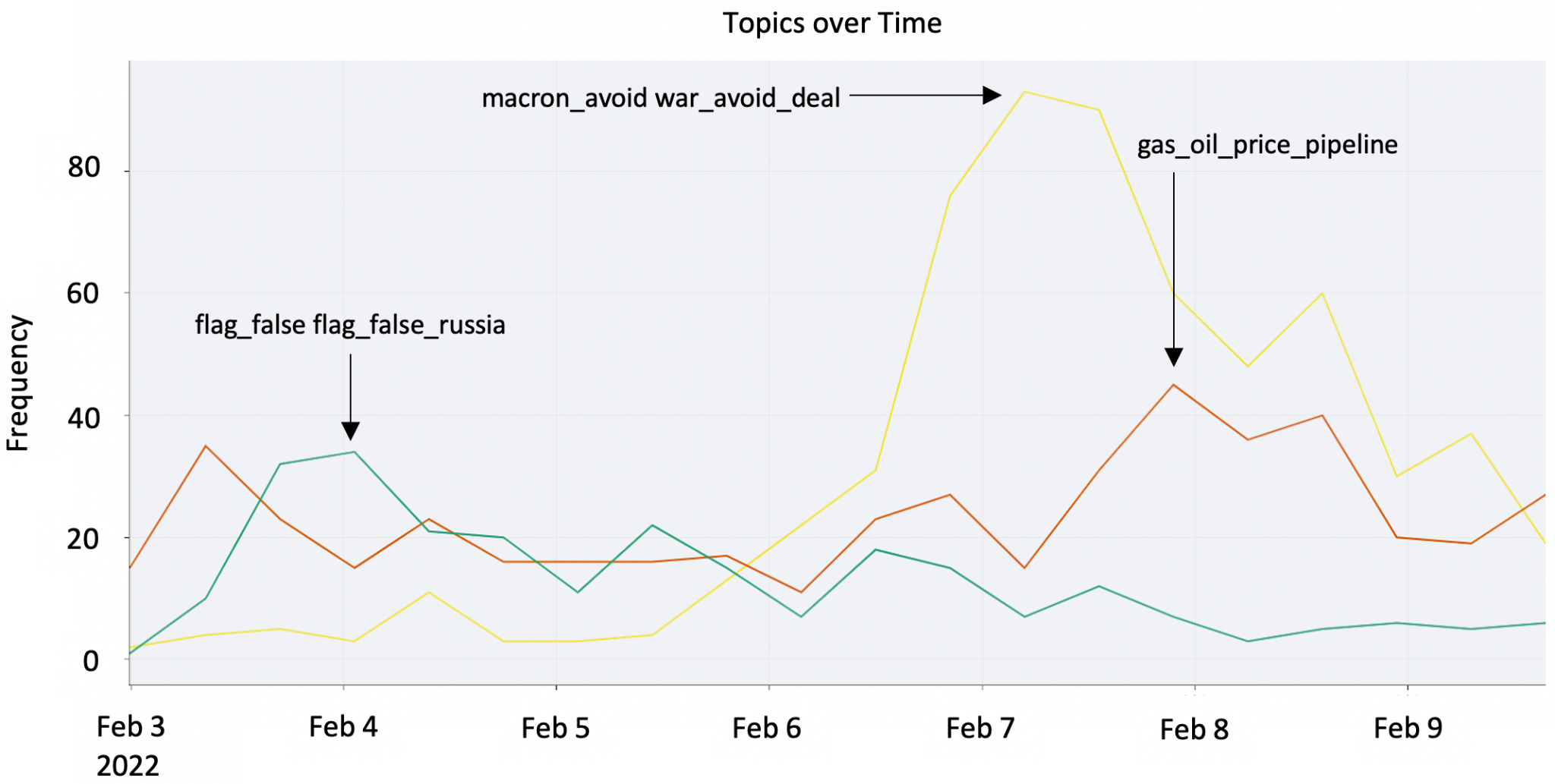
**Figure 2**

*Sample Named Entity Recognition (Dec. 31, 2021)*

***Topic Modeling***

When applied to 1-2 days of tweets, NMF successfully identified relatively fine grain topics that other methods missed. For example, NMF highlighted words in several topics relating to the protests in Kazakhstan leading up to Russia’s invasion of Ukraine (Bilefsky, 2022). However, when analyzing longer timeframes, BERTopic was better able to uncover relevant topics. In figure 3, we plot the frequency of some of the most salient topics in a sample timeline (02/03 - 02/09). This allows us to track fluctuations in the discussion of a particular topic.

**Figure 3**

*Selected Topics visualized over one week (Feb. 3 to Feb. 9)*

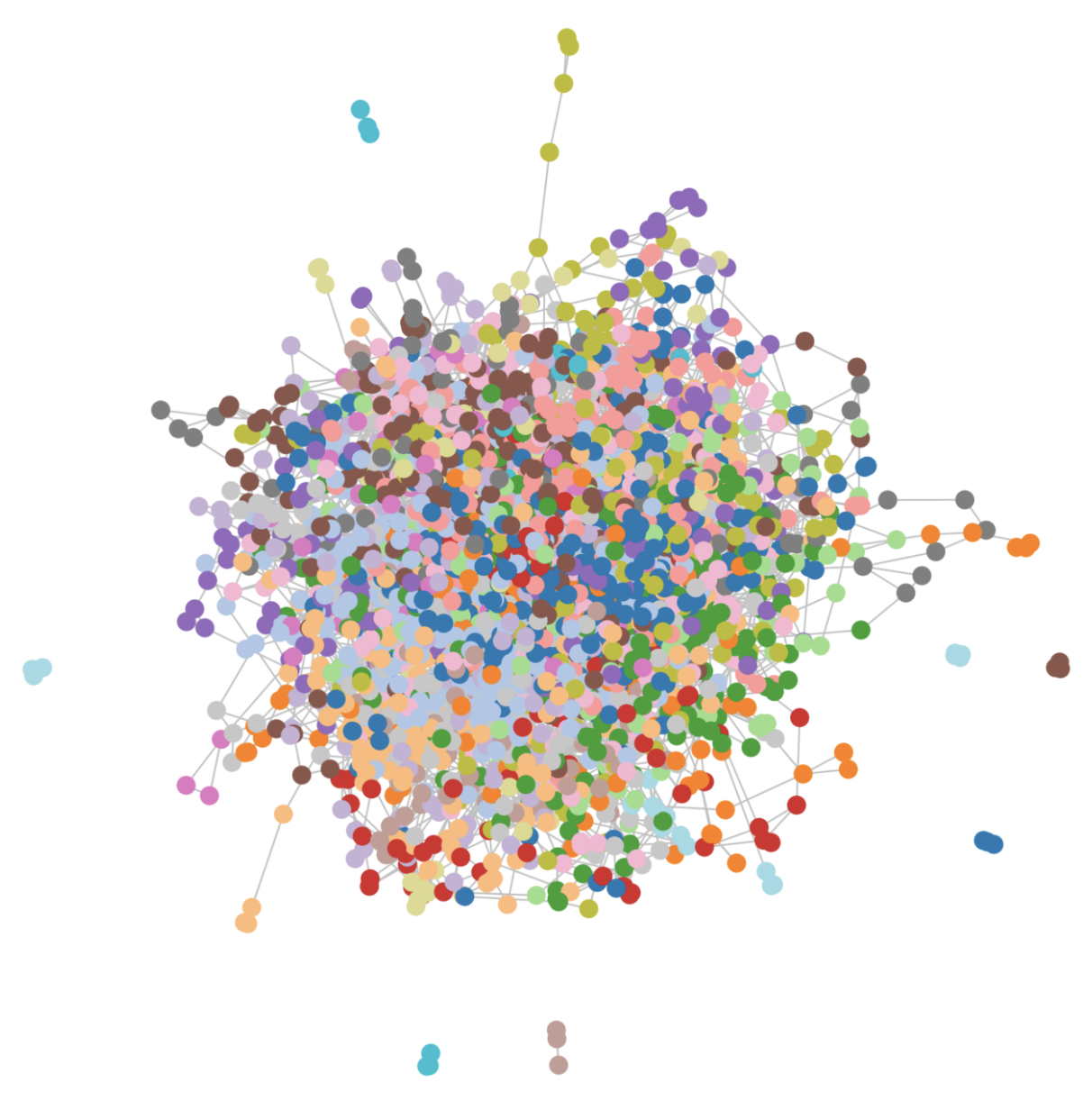
Notably, peaks in topic frequency accurately pinpoint the timeframe of sub-events. The *flag\_false flag\_false\_russia* topic peaks on Feb. 4 which coincides with the date that US officials reported Russia may attempt to conduct a false flag operation as a pretext for invading Ukraine (Nakashima et al., 2022). The spike on Feb. 7 identifies the date French President Emmanuel Macron traveled to Moscow to meet with President Putin in hopes of avoiding war (Rankin, 2022). The Feb. 8 spike in the *oil\_gas\_price\_pipeline* topic aligns with Putin’s threats to close its main gas pipeline to Germany if Western countries banned Russian oil (Thomas & Race, 2022).

***Clustering***

To visualize the complexity of sub-events in this global event we generated a network graph representing the degree of cosine similarity between tweets in a specified range of dates. This graph for tweets on Feb. 4 in figure 4, presented alongside the largest three clusters (on right), supports that this conflict has a significant amount of highly interconnected components.

**Figure 4**

*Network graph of similar tweets from Feb. 4, 2022*

By clustering tweets using a pre-trained BERT model from the SentenceTransformers library, we could determine individual sub-events in a more interpretable way. Furthermore, we find these clusters produced with transformers to be highly accurate when compared to a timeline of events (Figure 5). As an example, the word clouds in figure 5 showcase the power of transformers in separating tweets into specific and atomic clusters. For each date range, we selected three representative clusters that detected specific sub-events. As expected, words like “Ukraine”, “War”, and “Russia” appear throughout the clusters, so examining the set of words slightly less frequent (i.e., smaller font size) than these is most illuminating.

**Figure 5**

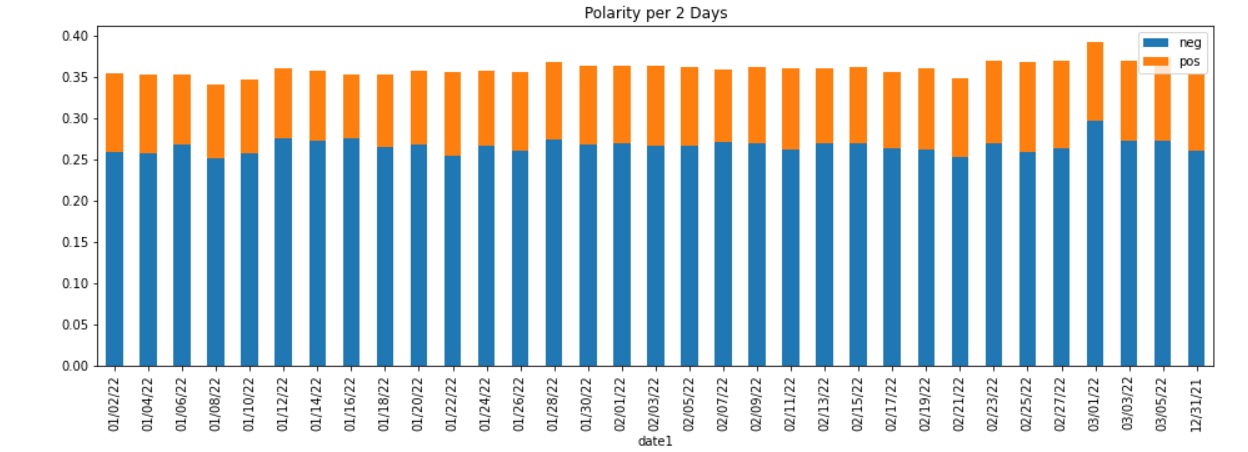
*Word Clouds of Sample Clusters across 3 Time Intervals of the War in Ukraine*

|  |  |  |
| --- | --- | --- |
| Jan. 27 - Feb. 6 | Feb. 19 - Feb. 24 | Feb. 26 - March 5 |
| President Joe Biden deployed US troops to Eastern Europe to reassure NATO allies (Feb. 2). | Putin recognizes the two separatist regions of Ukraine as independent nations (Feb. 21). | Russia blocks Facebook and Twitter (March 4) |
| Russia-Belarus joint military drills suggest Belarus may help Russia in an invasion (Feb. 3-10). | Putin sends Russian troops into eastern Ukraine (Feb. 24). | US announces Vice President Kamala Harris to visit Poland to show US support (March 4). |
| The Biden Administration warns of a possible Russian false flag attack (Feb. 4). | Anti-war protests erupt in Russian cities leading to thousands of arrests (Feb. 25-26). | Putin saying Western sanctions are akin to an act of war (March 5). |

***Sentiment Analysis***

Sentiment analysis is an important type of text analysis that aims to support decision-making by extracting and analyzing opinions, identifying positive and negative sentiments, and measuring how positively or negatively an entity is regarded. As more people use Twitter to communicate their political and social ideas, tweets become significant sources of information for sentiment analysis. This study proposes sentiment analysis as a text summarizing technique for detecting societal interest and public opinion about a social event. We analyzed the sentiment of the Ukraine war tweets to evaluate the change in sentiment over time. Figure 6 shows the sentiment for 1-2 days intervals of tweets using VADER (Valence Aware Dictionary for Sentiment Reasoning) Sentiment Analyzer, which uses a dictionary to map lexical elements to sentiment scores (Beri, 2020). Unsurprisingly, the tweets are highly negative throughout the war, but interestingly, sentiment scores remain relatively constant. Thus, fluctuations in sentiment could not be used to help identify the timeframe of key sub-events, but an analysis of sentiment does provide important overall context to interpret the results of other techniques.

**Figure 6**

*Sentiment by 2-day Intervals*

**Part 2: Event Summarization and Performance Evaluation**

Table 1 below shows summaries generated from a sample set of 304 tweets from Jan. 2 using abstractive and extractive methods. Both summaries give an overview of President Biden’s early efforts to avert war in Ukraine.

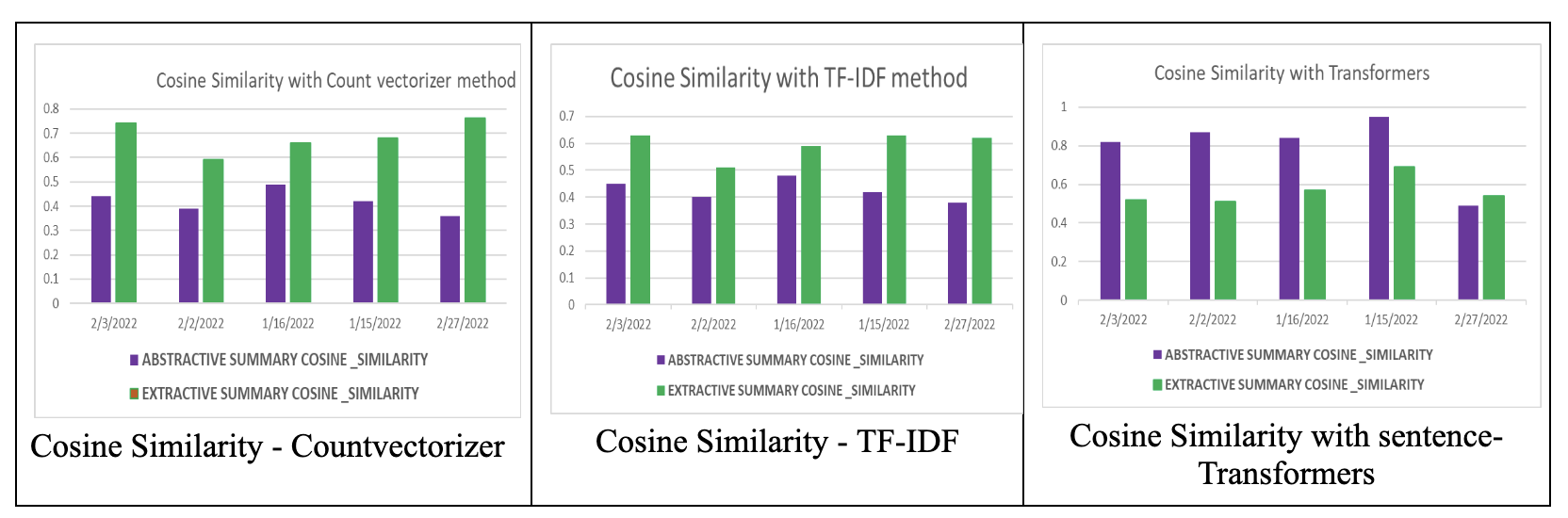
**Table 1**

*Comparing Sample Abstractive and Extractive Summaries*

|  |  |
| --- | --- |
| **Abstractive Summary** | **Extractive Summary** |
| biden vows us to act decisively if russia invades ukraine . russia and china know it and they lift a finger when they feel like it biden cannot afford another blunder . putin will not invade ukraine but to stifle it creating a tense atmosphere for a considerable long duration, biden will not risk a nuclear war . | the 17th-century Great Northern War made Russia a European power, and it ended with victory in Ukraine .“If Putin manages to keep NATO out of Ukraine, Georgia, and Moldova, and U.S. intermediate-range missiles out of Europe, he thinks he could repair part of the damage Russia’s security sustained after the Cold War ended.”.politics government Biden to talk with Ukraine's Zelenskyy to counter Russian intimidation: President Biden was back at work trying to prevent a new war in Europe Sunday, planning a phone call with Ukrainian Prime Minister Volody Radzikhovsky pens a provocative rejection of Putin’s supposed war intentions in Ukraine, arguing that he’s more Chaplin than Hitler. |

***Similarity Metrics***

For the purpose of evaluation, we applied similarity measures to find relations between a set of tweets (the input) and their generated summaries (the output). To calculate relatedness, we used metrics of cosine similarity, calculated with the count vectorizer method and TF-IDF. The similarity was also calculated using transformers (Sentence-BERT) cosine similarity. Here we select a random sample of tweets from 5 days to create extractive (Textrank) and abstractive (transformers) summaries. We found that input tweets and their generated summaries had an average similarity of about 65% for either extractive or abstractive methods. Cosine similarity with count vectorizer and TF-IDF was higher between the tweets and extractive summaries compared to abstractive ones for the same day. However, the similarity score generated by using sentence-transformers was higher between the tweets and abstractive summaries. Overall, we expect extractive techniques to score high with this method because they contain exact language from the input text, but they do not encode similarity in semantic information. In the third figure in table 2, we see the result of using sentence embeddings to compare tweets, which incorporate semantic similarity. Four days in this sample show the abstractive summary achieved greater than 80% similarity with the input text, suggesting that abstractive techniques often succeeded in distilling the most important information from the input tweets.

**Table 2:** *Similarity Scores for Extractive and Abstractive Summarization Methods* 

***ROUGE-1 Scores***

The classic method for evaluating automatically generated summaries is the ROUGE score (Recall-Oriented Understudy for Gisting Evaluation). Rudrapal et al. (2018) conducted a review of literature on Twitter event summarization and observed that ROUGE metrics are the most common method for evaluating tweet summarization. Furthermore, they note that “ROUGE-1 score is the most reliable” in this task, thus we also select ROUGE-1 to evaluate our summarization system (Rudrapal et al., 2018). The downside to this method is it requires a human-written reference summary to measure the unigram overlap with a system summary. This is time and labor-intensive, so for the purposes of this study we selected ten reference summaries from news articles pertaining to ten sub-events in the war (e.g., Russia blocking Facebook, possible false flag operations, halting the Nord Stream pipeline, etc.). These events were identified using our sub-event detection methods from part 1, thus we expected the summarization system to perform well on them. Two keywords per sub-event were used to generate ten system summaries with a pre-trained T5 model. For each pair, we then calculated ROUGE-1 scores and obtained an average score of 38.85. This means that on average 38.85% of unigrams in a reference summary were present in the system summary. See table 3 for a sample.

**Table 3**

*ROUGE-1 Calculations for Two of the 10 Sub-events Selected for Evaluation*

|  |  |  |
| --- | --- | --- |
| Reference Summary | Keywords and System Summary | ROUGE-1 |
| “Russian oligarch Roman Abramovich said he will sell the Chelsea soccer club in England, a dramatic move that comes in response to outrage over Russia’s invasion of Ukraine. Abramovich at the same time announced that he has instructed the renowned team to set up a foundation that will use all net proceeds from the sale of Chelsea to benefit all victims of the war in Ukraine.” (Mangan, 2022). | Keywords: ['abramovich', 'chelsea']  System Summary: russian billionaire roman abramovich will sell chelsea with all net proceeds for the benefit of all victims of the war in ukraine. he said proceeds from the premier league team sale will benefit all victims. the club will be sold to victims of russia's war and all proceeds will go to ukrainian war victims, he says. | 0.489 |
| “Pope Francis calls the war in Ukraine an act of violent aggression and a sacrilege without justification, and appeals for an end to the violence and long-term welcome of Ukrainian refugees.” (Watkins, 2022) | Keywords: ['pope', 'francis']  System Summary: pope francis calls for a worldwide day of prayer due to the threat of war in ukraine. he condemns actions destabilizing coexistence among nations and discrediting international law. pope francis decries folly of war. pope says war would be madness and voiced hope that tensions with russia can be overcome with serious dialogue. | 0.462 |

For comparison, Lee et al. (2021) summarized COVID-19 related tweets with a pre-trained T5 model and achieved a ROUGE-1 score of 38.13. Therefore, we conclude that for large-scale event summarization, our approach produces comparable results to established methods. Lee et al. (2021) found retraining the T5 model on social media data improved its performance, thus we recommend a similar procedure for future iterations of this study.

**Conclusion**

As the war in Ukraine develops, the approach we present here to condensing the main aspects of this event can be adapted to analyze tweets in near real-time, empowering decision-makers with timely and relevant information. Our study shows that automatic text summarization goes beyond simple extractive and abstractive summarization techniques. While sentiment analysis provides context for the event, NER, topic modeling, and clustering enable accurate identification of sub-events and the timeframe in which they occurred. This can then be used to produce relevant summaries, to sum up any real-time event from Twitter.

**References**

Alsaedi, N., Burnap, P., & Rana, O. (2017). Can we predict a riot? Disruptive event detection using Twitter. ACM Transactions on Internet Technology (TOIT), 17(2), 1-26. <https://doi.org/10.1145/2996183>

Beri, A. (2020, May 27). SENTIMENTAL ANALYSIS USING VADER. Medium. <https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>

Bilefsky, D. (2022, January 7). How the Kazakhstan Protests Started and Why They Matter—The New York Times. New York Times. <https://www.nytimes.com/2022/01/05/world/asia/kazakhstan-protests.html>

Corney, D. P., Martin, C., & Göker, A. (2014, April). Two Sides to Every Story: Subjective Event Summarization of Sports Events using Twitter. In SoMuS@ ICMR. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.661.5146>

Council on Foreign Relations. (2022, April 29). Conflict in Ukraine | Global Conflict Tracker. Global Conflict Tracker. <https://cfr.org/global-conflict-tracker/conflict/conflict-ukraine>

Dusart, A., Pinel-Sauvagnat, K., & Hubert, G. (2021). TSSuBERT: Tweet Stream Summarization Using BERT. arXiv preprint <https://doi.org/10.48550/arXiv.2106.08770>

Esmin, A.A.A., Júnior, R.S.C., Santos, W.S., Botaro, C.O., Nobre, T.P. (2014). Real-Time Summarization of Scheduled Soccer Games from Twitter Stream. In: Métais, E., Roche, M., Teisseire, M. (eds) Natural Language Processing and Information Systems. NLDB 2014. Lecture Notes in Computer Science, vol 8455. Springer, Cham. <https://doi.org/10.1007/978-3-319-07983-7_29>

Huang, Y., Shen, C., & Li, T. (2018). Event summarization for sports games using twitter streams. World Wide Web, 21(3), 609-627. <https://doi.org/10.1007/s11280-017-0477-6>

Jai-Andaloussi, S., El Mourabit, I., Madrane, N., Chaouni, S. B., & Sekkaki, A. (2015, December). Soccer events summarization by using sentiment analysis. In 2015 international conference on computational science and computational intelligence (csci) (pp. 398-403). IEEE.

Kammer, A., Azour, J., Selassie, A. A., Goldfajn, Ii., & Rhee, C. (2022, March 15). How War in Ukraine Is Reverberating Across World’s Regions. IMF Blog. <https://blogs.imf.org/2022/03/15/how-war-in-ukraine-is-reverberating-across-worlds-regions/>

Lee, C. H., Yang, H. C., Chen, Y. J., & Chuang, Y. L. (2021). Event Monitoring and Intelligence Gathering Using Twitter Based Real-Time Event Summarization and Pre-Trained Model Techniques. Applied Sciences, 11(22), 10596. <https://doi.org/10.3390/app112210596>

Li, Q., & Zhang, Q. (2021, May). Twitter event summarization by exploiting semantic terms and graph network. In Proceedings of the AAAI conference on artificial intelligence (Vol. 35, pp. 15347-15354).

Mangan, D. (2022, March 2). Russian oligarch Roman Abramovich says he will sell Chelsea soccer club amid the Ukraine war. CNBC. <https://www.cnbc.com/2022/03/02/russian-oligarch-abramovich-says-he-will-sell-chelsea-soccer-club-amid-ukraine-furor.html>

Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., & Miller, R. C. (2011, May). Twitinfo: aggregating and visualizing microblogs for event exploration. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 227-236). <https://doi.org/10.1145/1978942.1978975>

Matthews, A. L., Stiles, M., Nagorski, T., & Rood, J. (2022, April 28). The war in data: 2,000 Russian armored vehicles destroyed. Grid News. <https://www.grid.news/story/global/2022/03/24/the-war-in-data-tracking-the-toll-in-ukraine/>

Nakashima, E., Harris, S., Parker, A., Hudson, J., & Sonne, P. (2022, February 3). U.S. accuses Russia of planning to film false attack as pretext for Ukraine invasion. Washington Post. <https://www.washingtonpost.com/national-security/2022/02/03/russia-ukraine-staged-attack/>

Nichols, J., Mahmud, J., & Drews, C. (2012, February). Summarizing sporting events using twitter. In Proceedings of the 2012 ACM international conference on Intelligent User Interfaces (pp. 189-198).

Pham, T., & Talavera, O. (2022, March 30). Putin, Путин: A descriptive analysis of the Twitter universe. <https://voxukraine.org/en/putin-putyn-a-descriptive-analysis-of-the-twitter-universe>

Purtova, D. (2022, March 6). Russia-Ukraine war—Tweets Dataset (65 days). Kaggle. <https://www.kaggle.com/foklacu/ukraine-war-tweets-dataset-65-days>

Rankin, J. (2022, February 7). Macron plays down expectations as he arrives for Ukraine talks with Putin – as it happened. The Guardian. <https://www.theguardian.com/world/live/2022/feb/07/ukraine-russia-crisis-macron-heads-for-talks-with-putin-while-scholz-and-biden-meet-in-dc-live-coverage>

Rudrapal, D., Das, A., & Bhattacharya, B. (2018). A survey on automatic Twitter event summarization. Journal of Information Processing Systems, 14(1), 79-100. <https://doi.org/10.3745/JIPS.02.0079>

Sharifi, B., Hutton, M. A., & Kalita, J. (2010, June). Summarizing microblogs automatically. In Human language technologies: The 2010 annual conference of the north american chapter of the association for computational linguistics (pp. 685-688). <https://moam.info/summarizing-microblogs-automatically-acl-anthology-association-_5bbc2b9a097c4792598b4596.html>

Shen, C., Liu, F., Weng, F., & Li, T. (2013, June). A participant-based approach for event summarization using twitter streams. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 1152-1162). <https://aclanthology.org/N13-1135>

Thomas, D., & Race, M. (2022, March 8). War in Ukraine: Russia says it may cut gas supplies if oil ban goes ahead. BBC News. <https://www.bbc.com/news/business-60656673>

Watkins, D. (2022, March). Pope: ‘War of aggression against Ukraine is inhuman and sacrilegious’—Vatican News. Vatican News. <https://www.vaticannews.va/en/pope/news/2022-03/pope-francis-ukraine-war-inhuman-sacrilegious.html>