U.S. Congressional Tweeting About the Russo-Ukrainian War: A Textual Analysis

Ben Berkman and David Trakhtenberg

May 11, 2022

1 Introduction

The Russo-Ukrainian War began in 2014 with the Russian annexation of Crimea and battles for land in the Donbas. The war escalated precipitously when Russia invaded Ukraine on February 24, 2022 and has been raging on since—the results of which have been devastating. Europe's largest refugee crisis since World War II has ensued with more than 5 million Ukrainians leaving the country and a quarter of the population displaced (Doherty, 2022).

Consequently, the outcry from the international community has been persistent. From the UN, there have been calls for a full withdrawal of Russian troops as well as initial attempts at trying high-ranking Russian officials as war criminals. In response, many nations are lending their support financially and emotionally, from accepting refugees into their country to pledging hundreds of millions of dollars in security assistance (Jeong et al., 2022). Most notably, America has lent a big helping hand, leading the charge in many of these efforts. Republican support of Ukraine has increased from 42% to 58%, pre- and post-invasion, while Democratic support has similarly jumped from 58% to 70% (Viala-Gaudefroy, 2022). While a simple analysis of these figures lead us to believe in a general unity in attitude towards the war between the two parties, we find it a useful exercise to analyze the differences in views between the right and the left on this far-reaching war. Twitter is a suitable medium for a comprehensive dive into the textual communications of American politicians – in this case congresspeople.

There are a number of methods we can employ to study Twitter data. One approach is to apply descriptive inference (DI) which provides methods to reduce dimensionality – important for inherently high-dimensional text. DI also yields a model to represent text as vectors and measure their similarity, perfect for understanding differences in congresspeoples' beliefs. Another approach is traditional supervised classification where we learn a relationship between inputs (Tweets) and a labeled set of outputs. And one final approach is unsupervised learning where we learn latent structures in an unlabeled version of our data to interpret the sentiments, topics, and other hidden representations.

By utilizing these text analysis techniques, we seek to answer three research questions. Namely, (1) can we predict Tweet virality solely from a congressperson's Tweets about the Russo-Ukrainian War, (2) do the right and left differ in sentiment when tweeting about this war, and (3) what topics can we derive from these Tweets to better understand overarching themes and how do these differ by party lines? We find that words are not predictive of virality as shown through the training of various supervised learning models. In terms of sentiment, Republicans have slightly more negative messaging and focus on individuals, while Democrats espouse a more policy-driven and community-oriented agenda. We discuss the six distinct topics formed from these Tweets, and similarly observe through these topics that Republicans are more focused on individualism while unity powers Democratic Tweets.

Past research largely focuses on using modern language models (e.g., RoBERTa) to predict tweet virality (Maldonado-Sifuentes et al., 2021), assessing Twitter use by American Congress members broadly (Golbeck et al., 2010), and curating Tweets about the Russo-Ukrainian War (Chen and Ferrara, 2022). Our paper is influenced by all of these works in an effort to analyze virality, topics, and sentiments via Tweets by Congress about the Russo-Ukrainian War.

2 Data

Data is collected from Tweets of congresspersons using Python and Twitter's API via Tweepy (Roesslein, 2020). We use a repository of Congressional Twitter accounts curated by UC San Diego to form a list of Twitter usernames for all members of Congress (House of Representatives and Senate). This list includes each congressperson's state and party affiliation. Using Tweepy, we collect the 200 most recent Tweets from each individual. We exclude any retweets, quoted retweets, or replies to ensure we get the most original text from each congressperson. URLs and non alpha-numeric characters are removed, and the text is lowercased per standard text preparation practices.

This returns about 75,000 Tweets. However, we are only concerned with Tweets about the Russo-Ukrainian War, and specifically about how Congress Tweets about Ukraine. We filter the 75,000 Tweets to only include Tweets that include "ukrain" (case insensitive and stemmed to include both "Ukraine" and "Ukranian"). Because this study also includes a discussion of virality, we include only Tweets that are old enough to have gone viral. In this case, we select only Tweets that are greater than three days old.

4,377 Tweets remain after these filters are applied; they are used for the analyses in this paper. Tweets range from Sept. 9, 2021 to April 11, 2022 with a strong left skew (the median date is March 10, 2022). The most number of Tweets on any day (374) occurs on Feb. 24, 2022, the day on which Russia launched its attack on Ukraine. The second-highest day, March 16, 2022, (357 Tweets) marked the beginning of Russian and Ukrainian negotiators discussing neutrality, and Russian president Vladimir Putin's crackdown on dissenters. And March 12, 2022 (223 Tweets), the third most-tweeted day, coincided with an Associated Press team filming Russian tanks leveling an apartment building in Mariupol, an event that went viral on Twitter.

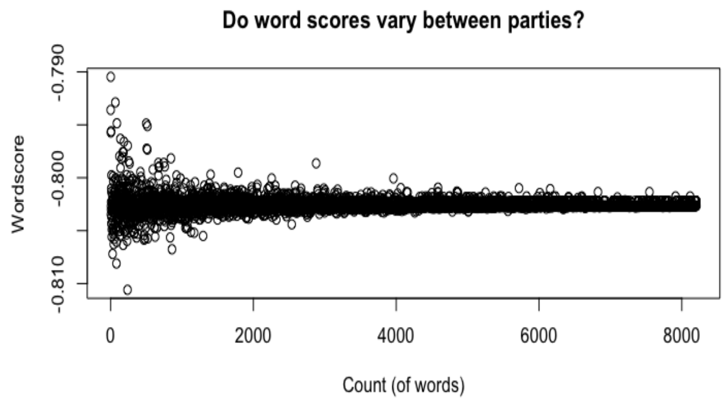
Finally, we engineer a binary feature, “viral”, to classify if a Tweet is viral or not. To form this feature, we first calculate the mean retweets for each congressperson. If a certain Tweet is retweeted greater than twice the mean number of retweets for that congressperson, we classify it as viral.

3 Assessing Tweet Virality

We first seek to learn a relationship between words in a Tweet and how popular that Tweet becomes (i.e. virality). We train three supervised models to this end: Naive Bayes classifier, support-vector machine (SVM), and logistic regression. We vary training size for each algorithm and find little difference in performance. The following steps are taken for each classifier: setting the training size, creating an index of the number of rows in the data set, removing stop words and punctuation, and splitting a Document Feature Matrix (DFM) into train and test sets. A model is then trained on these DFMs and evaluated.

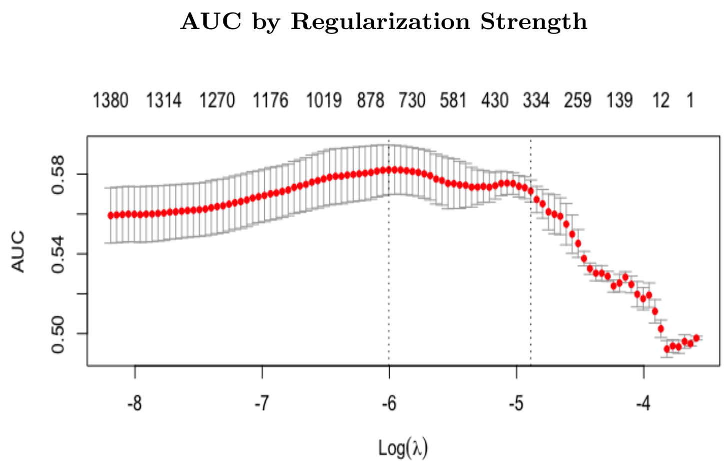
The Naive Bayes classifier learns posterior probabilities by calculating the product of the prior probabilities of a document occurring in a class and the likelihood that class is correct. The naive assumptions assume conditional and positional independence. For example, the probability of observing “Biden” is independent of the probability of observing “president” given the Tweets are from a Democratic congressperson. Once evaluated, we find that the classifier achieves poor out-of-sample performance (Accuracy: 0.80, Recall: 0.23, Precision: 0.16, F1-Score: 0.19). Accuracy is not a useful metric in our imbalanced data set with a base rate of 10% of Tweets being viral. The low precision and recall suggests predicting virality using a Naive Bayes approach is ineffective.

We employ wordscores (Laver et al., 2003), utilizing our training set as a reference to generate relative word scores. After generating these scores, we find that the importance of words in our Tweets with respect to virality is relatively uniform, as seen below. However, we do find some words hold higher weight such as “putin”, “energy”, “oil”, “china”, and “prices”. This suggests that tweets with more inflammatory words spark stronger sentiment and cause users to retweet more often.



For the SVM classifier, we follow an identical workflow. SVM learns by finding a separating hyperplane to classify Tweets into viral or not. In this setup, we employ a 5-fold cross validation to tune our parameters and avoid overfitting. We find familiar results (Accuracy: 0.90, Recall: 0.17, Precision: 0.14, F1-Score: 0.16). It is unsurprising that this new classifier was unable to find a relationship in the data as it is does not appear to be linearly separable.

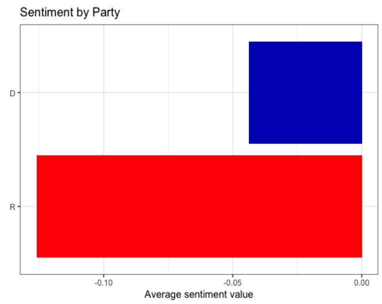
Finally, we train a logistic regression model as this classifier is interpretable and often performs well on smaller data sets. Yet again, we find unsatisfying out-of-sample results (Accuracy: 0.90, Recall: 0.08, Precision: 0.17, F1-Score: 0.11). However, when we train using a penalized maximum likelihood, we find more intuitive results, as seen below. The figure shows the regularization path of area under the receiver operating characteristic curve (AUC) computed for a ridge regression with the regularization parameter lambda on the x-axis (Friedman et al., 2010). If we condition on an optimal lambda, AUC approaches 0.58 which as we find before is not a very predictive classifier, but still better than randomly guessing (AUC = .50) which suggests some learning does occur. Since AUC is invariant to the base rate, this is a good metric to analyze. When using lasso regression, we achieve an AUC of 0.64. While higher, this is still low and we conclude that tweeted words about the Russo-Ukraine War from congresspeople is not very predictive of virality.



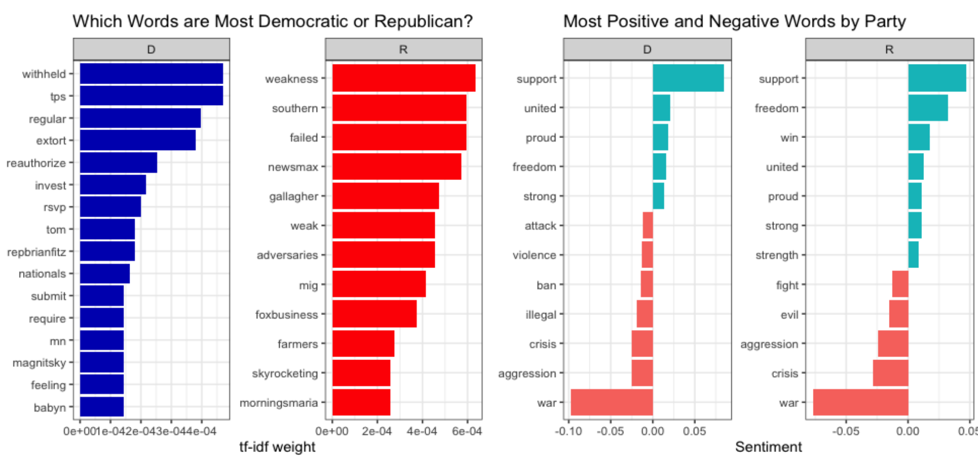
Similar to the wordscores, we attempted to find topics that are most predictive of virality using logistic regression coefficients as well as feature importance with random forest. We trained a 40-topic Structural Topic Model (STM) and assigned each document to the topic with the highest respective gamma value (probability of topic given document). While feature importance produced a ranking of top features, there is not enough signal in topic assignment to predict virality and in each case the model simply predicts the majority class per topic (which is always non-viral). This further points to the difficulty in predicting virality.

4 Sentiment Analysis

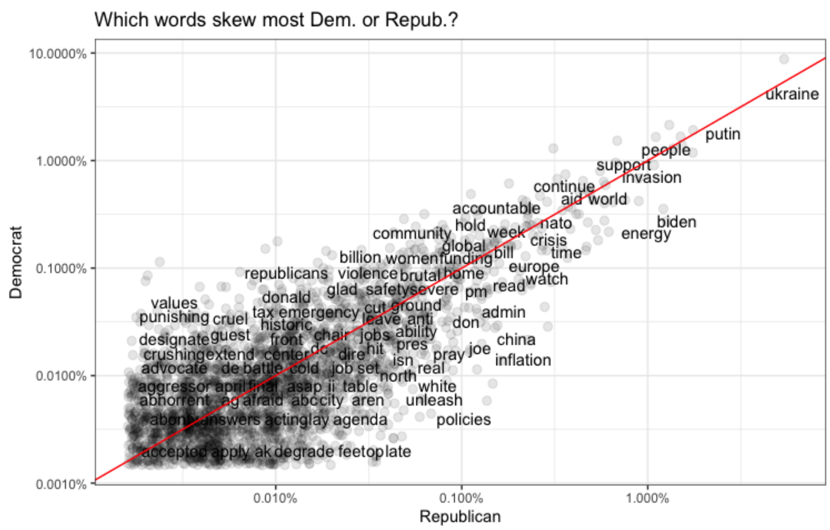
Transitioning to an unsupervised setting, we compare and contrast how Democratic and Republican congresspersons tweet about the war through sentiment analysis. We use a number of different sentiment analysis approaches throughout this section, and much of this analysis is attributed to similar analysis and code found in Silge and Robinson, 2017. We first use a dictionary approach. We apply the AFINN dictionary (Nielsen, 2011) provided in the tidytext package to each Tweet and take the average sentiment score of each Tweet, displayed below. Via this approach, we see that, on average, Republicans tweet with a slightly more negative sentiment about Ukraine than Democrats.



We next analyze which specific words are most Democratic or Republican, and which words are most positive or negative within each party. Using term frequency inverse document frequency (TF-IDF) of each word in each Tweet, we select the words with the highest TF-IDF weight. Intuitively, these words are most unique to one party over the other. Interestingly, top Republican words include "weakness", "failed", and "weak" – words often tweeted by former President Trump. Democratic words are largely more policy-oriented (e.g., "withheld", "reauthorize", and "invest"). The AFINN dictionary includes a continuous scale for sentiment scores (i.e., "killed" is more negative than "war"), so we can calculate the most negative words for each party as well. These do not drastically differ between parties, but it is worth noting that Democrats tweet more about "unity" while Republicans tweet more about "freedom", both of which are consistent with their respective party's dogmas. In the below plot, Democrats are the first and third panels and Republicans are the second and fourth.



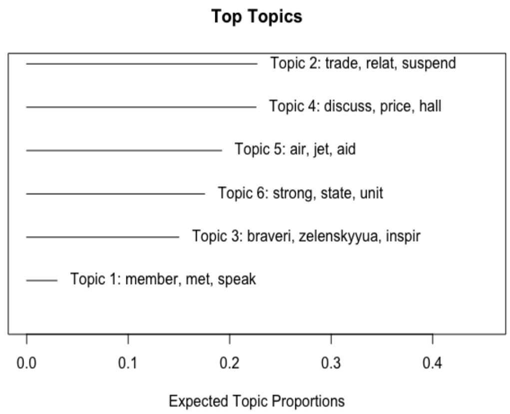
As a final analysis, we take a slightly different approach than the TD-IDF vectorization to analyze which words skew more left or right. Using a count-based approach, we count word occurrences within each party then scale this value by the total number of words used by each party to create a word frequency for each word for each party. We select only words that occur in tweets by both parties, then plot the word with the Republican frequency of the word as the x-axis and the Democratic frequency of the word as the y-axis. This produces some interesting conclusions. In the context of Ukraine, Democrats are more likely to tweet about Republicans, and Republicans are more likely to tweet about President Biden. Republicans are also more likely to tweet about what they view as shortcomings of the Democratic presidency ("inflation" and "energy"), while, like the previous analysis, Democrats are more focused on "community" and "values".



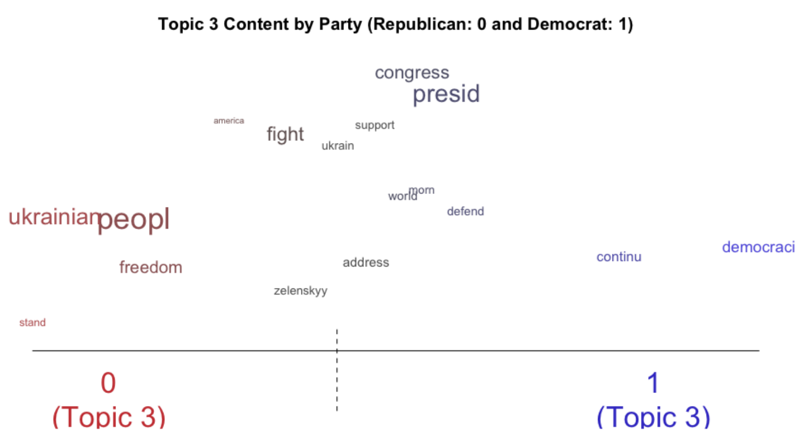
5 Topic Modeling

In addition to analyzing how sentiments differ between the right and left in the context of Ukraine, we analyzed how congresspersons' tweets form topics. To do this, we build a STM using semi-collapsed variational expectation–maximization (EM). STMs vary from Latent Dirichlet Allocation (LDA) models in that prevalence and content are derived from metadata instead of a Dirichlet distribution. STMs drop the LDA assumption that, given a topic, word distribution is stationary. STMs return a probability of topic assignment for each document. Further mechanics of the STM model are outlined in the stm package documentation (Roberts et al., 2019). For our model, the topic content varies according to party assignment, and the prevalence varies according to party assignment and the spline of the date that the Tweet was sent.

Setting the number of topics to six, the model returns the following topics with the associated expected topic proportions and top keywords. Some topics are noteworthy. For example, the second topic considers the economic impacts of trade due to the war, the third topic speaks to the bravery and inspiration exhibited by Ukrainian President Volodymyr Zelenskyy, and the fifth topic relates to the aid provided by outside countries via air and jets.



Diving deeper into the third topic, we can observe which words are most likely to come from Democrats and which are most likely to come from Republicans. Following the trends of the other analyses we have discussed above, in the context of Tweets about President Zelenskyy, Democrats tweet more about unity (e.g., “democraci”, continu“) while Republicans are more focused on individual virtues (e.g., ”fight“, ”freedom“).



6 Conclusion

In this paper, we sought to answer three questions using tweets by American members of Congress about the Russo-Ukrainian War: (1) can we predict if a Tweet will go viral based solely on its text, (2) do the right and left differ in sentiment when tweeting about this war, and (3) what topics can we derive from these Tweets to better understand overarching themes and how do these differ by party lines?

The first question presents a supervised learning problem, while the second and third questions are unsupervised in nature. We find that predicting Tweet virality is a difficult task that likely requires more input than just the words of a Tweet. In the unsupervised portion, we conclude that Republicans and Democrats do differ by sentiment and topic. Republican messaging is slightly more negative and individual-focused, while Democrats tweet to more policy-oriented and community-minded ideals.

Future work may entail using more Tweets as the war progresses to increase the data size. A classifier was unable to properly learn a relationship between words and virality in our setting; however, with more data a stronger signal may emerge. Alternatively, it may entirely be the case that Democrats and Republicans are trying to temper the aggressiveness of their tweets and so they don't invoke much outrage (or retweets). As we saw in Section 3, there were some inflammatory words (e.g., "oil“ and "china“) that were used sparingly by congresspeople.

There are other interesting questions to explore, as well. For example, there is a well documented history of disinformation used as propaganda in Russia (Paul and Matthews, 2016). Through topic modeling and sentiment analysis of Russian citizens, we could study how propaganda in Russia has changed over time and if it is becoming more or less effective.

References

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, *97*(2), 311–331. https://doi.org/10.1017/S0003055403000698

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the u.s. congress. *Journal of the American Society for Information Science and Technology*, *61*(8), 1612–1621. https://doi.org/https://doi.org/10.1002/asi.21344

Nielsen, F. Å. (2011). Afinn. http://www2.compute.dtu.dk/pubdb/pubs/6010-full.html

Paul, C., & Matthews, M. (2016). *The russian firehose of falsehood propaganda model: Why it might work and options to counter it*. RAND Corporation. https://doi.org/10.7249/PE198

Silge, J., & Robinson, D. (2017). *Text mining with r: A tidy approach* (1st). O'Reilly Media, Inc.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1–40. https://doi.org/10.18637/jss.v091.i02

Roesslein, J. (2020). Tweepy: Twitter for python! *URL: https://github.com/tweepy/tweepy*.

Maldonado-Sifuentes, C. E., Angel, J., Sidorov, G., Kolesnikova, O., & Gelbukh, A. (2021). Virality prediction for news tweets using roberta. In I. Batyrshin, A. Gelbukh, & G. Sidorov (Eds.), *Advances in soft computing* (pp. 81–95). Springer International Publishing.

Chen, E., & Ferrara, E. (2022). Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between ukraine and russia. https://doi.org/10.48550/ARXIV.2203.07488

Doherty, E. (2022). Un: Over 5 million refugees have fled ukraine since russian invasion began. https://www.axios.com/2022/02/26/ukraine-refugees-russian-invasion

Jeong, A., Demirjian, K., Bella, T. B., & Stein, J. (2022). Biden authorizes $350 million more in defense aid for ukraine in response to russian invasion. *The Washington Post*.

Viala-Gaudefroy, J. (2022). Ukraine war: While most americans express outrage, putin's spell continues to hang over republicans. *The Conversation*.