Comparing Twitter Language of Congresspeople Before and During Election

Elijah Frankle, Khang Dinh, Louis Du, Luke Dillon

Department of Linguistics, University of California, Santa Barbara (UCSB), 93106, CA, United States of America

Abstract

Using a public dataset containing posts from Twitter accounts, we analyze the language of members of Congress during campaign season and office. We employ a neural network classifier that predicts differences in language use in members of Congress between the time periods of campaign season and office, informed by a set of pre-trained embeddings. We find that a difference exists between the language used on Twitter by members of Congress when they are campaigning for reelection versus when they are occupying office. We then fine-tune our predictions to explore what lexemes are influencing the predictions. The distinctions we discover have implications for the influence social media platforms can possess when utilized by politicians.

Keywords: political linguistics, natural language processing, discourse analysis, sentiment analysis

1. Introduction

Within the last few years, social media has essentially become a requisite for any serious politician hoping to garner public attention. In the age of the Internet, information can be disseminated to a global audience with hyper-speed connectivity, rendering social media websites like Twitter a tool for politicians to broadcast their messages. Featuring a free account registration model, Twitter allows its registered users to interact on the website by viewing, messaging, reposting, and following other accounts via live updates. Messages on Twitter (i.e. tweets) are characteristic for their character limit dictating how long a single tweet can be, as a consequence heightening the efficiency and productivity of information being produced and circulated on the social networking service.

The perpetual spread and permanence of information on Twitter has become an opportunity for researchers seeking to create models and analyze trends using Twitter data. With respect to research in political science, the literature has spawned quickly in recent years. Most studies look at public sentiment toward politicians [1] – [6]. Others like [7], [8] classify sentiment from the perspective of politicians' language use. This research typically develops computational methods to aid in the sentiment analysis of political social media messages; these include Naïve Bayes classifiers [2], convolutional neural networks [3], and bidirectional recurrent neural networks [8] that evidence the growing power of computational methods to comb through large amounts of data to make accurate predictions. These methods have been used to probe attitude-behavior consistency of politicians' promises [7], topic model tweets surrounding political events [4], or analyze cross-platform differences in political communities [6]. Indeed, the existing literature on politics in social media research is incredibly dense yet nascent, prompting research that assesses the quality of the literature while propounding potential models that might improve the process of conducting this kind of research [9]. We recommend two literature reviews [10], [11] beyond the scope of this paper that have summaries commenting on this kind of research, including how Twitter displays a bias for broadcasting over conversing [10], and more reflectively how the term "sentiment analysis" is inconsistent with its definition across the literature of this kind of research [11].

Our research aims to analyze trends from Twitter data, extrapolating the results to larger contexts that reveal the language use of members of the US Congress. A positive consequence of this goal is that we are able to observe the difference in language use of politicians in certain contexts; this is important in understanding whether politicians deliberately alter their language use, especially with the knowledge that their messages are being widely broadcasted. Though existing research classifies the language use of politicians in transcribed speeches [7] or news discussions on private messaging services [8], little is known about the language use of politicians on public social media with respect to temporal parameters. In a time when US politics are increasingly becoming influenced by Twitter—even insofar as defining its environment—is there a difference in the way members of the US Congress use language on Twitter when they are campaigning for reelection versus when they are occupying office?

In this paper, we shall discuss how we process our data [12], employ a neural network classifier, and train our classifier to yield telling results. Significantly, our results support the idea of utilizing "text-as-data" on the Internet as a valuable and rich resource for researchers. Our results also prove that computational methods can be an efficient and effective way of conducting research on massive amounts of data, thereby improving upon certain limitations of traditional methods of data analysis (e.g. surveys, experiments, interviews, case studies). At the end, we pose a discussion that contextualizes our results, addresses limitations of our study, and offers incentives for future related research.

2. Data

The data used in this study came from the "Congressional Tweet Archive", gathered by Bryan Gervais in 2019, accessed through the Harvard Dataverse. It contained eight files, one for each congressional chamber and each political party for the 112th and 113th Congressional Sessions (2011-2015). Each file contained columns corresponding to congresspeople, the rows below populated with the body of tweets and their dates.

To clean this data, we first removed the files relating to Senators as our focus was on members of the US House of Representatives. We then moved on to using a Python script to algorithmically clean and prepare the data for the neural network. We removed all congress people that were either not running for re-election or did not complete their full term in office. We removed all hyperlinks from the body of the tweets as these links were not readable and therefore did not contain any valuable linguistic information. We removed all tweets that did not have a date associated with them as they would not be able to be categorized without their date. We moved all "@" symbols so "@NIH" and "NIH" would count as the same token as they carry similar semantic information.

Then, for each eligible congressperson in the dataset, we counted the number of tweets they had before a certain date threshold (see Methods for more detail on the date threshold), corresponding to the "in office" category, and the number of tweets they had after the threshold, corresponding to the "campaigning" category. If they had more than 300 tweets in each category, we took exactly 300 tweets from each category. If they had less than 100 tweets in either category they were deemed to have too little Twitter usage and were thus removed. If both categories had at least 100 tweets but at least one had below 300, we took that lower number of tweets from each category. From the tweets we took, 75% were put in training data and 25% were put in testing data, assigned randomly. The training and testing data were then exported as two new csv files. When the body of the tweets were tokenized, we checked for co-occurrence between the vocabulary in each classification. Any tokens that did not appear at least once in both classifications were removed. This was to avoid time-specific current events from influencing our learning of broader linguistic trends.

3. Methods

We train a three-layer feed-forward mean-bag-of-embeddings neural network classifier with two ReLU hidden layers of 100 units. The output is a simple sigmoid function, which yields a value between 0 and 1. This represents the probability of deviating to the alternative class (campaign). Due to our limited computational resources and time we wanted to use a reasonably simple model as it would cut down the time we needed to train. The neural network is initialized with unfrozen pre-trained GloVe embeddings with 50 dimensions, which were pre-trained on Twitter data. Additionally, the GloVe embeddings would help us cut down on training time, while also giving us very relevant, Twitter-specific embeddings to start out with.

We train a few models on different training data. Each model is based on a different date cutoff (Fig. 1) to separate the period of in office versus campaigning. The different cutoffs we tested were April, May, June, July, and August. We choose the date cutoff that yields the best validation accuracy and loss. These different models perform relatively similarly, so we simply take the best performing model which corresponds to the July cutoff. That is to say, we consider "near a campaign" as anything after and including the first day of July in the second year of a Congressperson's term.

We then use this date cutoff in conjunction with taking a minimum of 100 tweets and a maximum of 300 tweets from each representative during each time period (campaign and in-office) in an attempt to get more data to work with (before we only took 400 tweets from each individual that had at least 400 tweets, which removed too many accounts that didn't use Twitter that frequently). The model we use is trained for 50 epochs and yields 0.662 validation accuracy and 0.666 validation loss.

A mean-bag-of-embeddings model simply pools together all of the embeddings—GloVe in this case—and gets the mean. We can treat this mean as a sentence, or tweet, embedding. Since we are using a mean-bag-of-embeddings model, we cannot analyze word order in context as this type of model does not take context into account. We explore the patterns of the model by creating our own prompts and feeding them into our trained model.

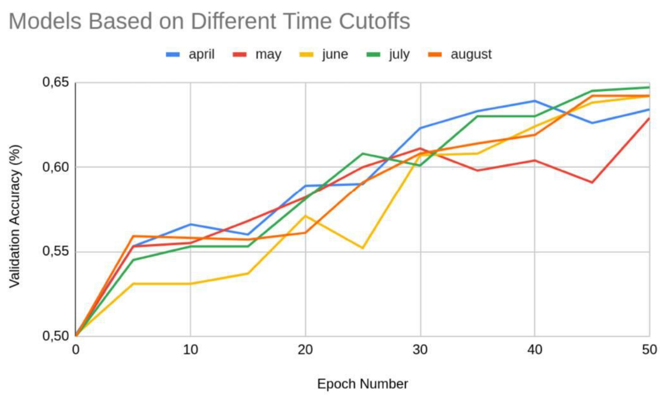

Figure 1: The validation accuracy is on the vertical axis and epoch number is on the horizontal axis. Cutting the threshold too close to election season yields a lower accuracy. July performs best in training and demonstrates a steady growth.

4. Results

Besides the model performance, we are also able to prove the presence of a change in language use by visualizing the probability values of the two classes in a box plot (Fig. 2) Though both classes have outliers, it is obvious that the mean and median probabilities of the class "campaign" are significantly higher than "office." This means that the candidates do

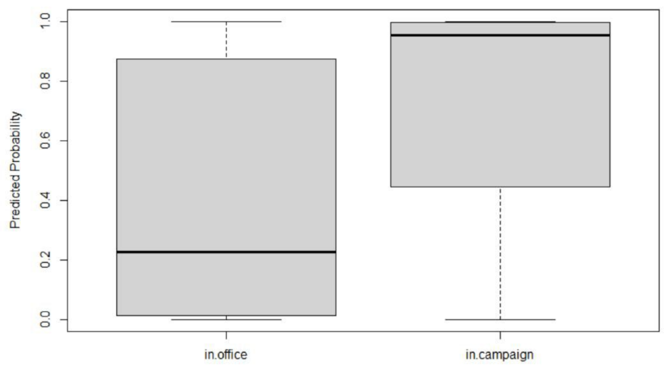

Figure 2: The predicted probabilities of "in office" and "in campaign" classifiers from the feed-forward mean-bag-of-embeddings neural network classifier. There is a noticeable difference in the mean and interquartile range between the two classes.

tweet differently before and after the election, and that our model is effectively capturing it.

Seeing that the model is indeed differentiating between the two classes, we then turn to explore what these differences are. We accomplish this by making predictions on modified test cases. We first inspect our test results, manually picking out test cases that have high or low probabilities. We then hypothesize that certain features in these sentences are causing the differences. To verify that, we perform predictions on the individual words and can see their relation with the two classes. With this in mind, we pick out five sentences from each class, with their predicted probabilities (Fig. 3-4)

1	My speech on the house floor today againt #sopa	0.003387954
2	Rick speaks about fighting the flood waters on Garrison Dam...	0.004512765
3	I'm pleased to announce the 2012 Academy Days schedule of events for the Third District. Hope to see you there! | #AR3	0.006725665
4	I just voted to disapprove the President's request to raise the debt limit, we need to cut spending and stop the insanity	0.007613
5	Just completed interview w/ Channel 9 KMBC News re: the budget negotiations & a possible govn. shutdown.	0.007970719

Figure 3: Sentence predictions from the test results of office classifications represented by values close to 0. A declaratory and promising tone characterizes these sentences with optimistic phrases such as, "Just completed", "I just voted", and "I'm pleased to announce".

For the tweets made during the politicians' terms, we have the intuition that they feel like simple announcements of what they are doing; they seem like daily reports in this sense. It is possible that this intuition comes from their choice of verbs. This is reflected in the words "speech", "speak", "announce", and "complete." In contrast, the tweets made during the period where they are campaigning for reelection have more emotion in the choice of verbs where they are much more eager to express their opinions, such as in sentence 9 through the word "condemn." Topics like natural disasters are much more common as well. To test our intuition, we made predictions on inputted texts (Fig. 5) based on their perceived test influence.

6	Governor McDonnell Declares State of Emergency in Preparation for Hurricane Sandy:	0.9931639
7	It's no secret that severe drought across the state this year has hurt a number of Arkansas... #AR3 USDA #Arkansas	0.991412
8	Press conference on why we need a bipartisan solution to fiscal cliff starting shortly	0.9914266
9	I absolutely condemn the senseless attacks that killed Ambassador Stevens & 3 other American diplomats in #Libya:	0.9900165
10	I'll be discussing the fiscal cliff on tamronhall in a few minutes. Tune in via msnbc!	0.9864889

Figure 4: Sentence predictions from the test results of campaign classifications represented by values close to 1. An urgent and demanding tone characterizes these sentences discussing violent or severe world events using phrases such as, "It's no secret", "why we need a bipartisan solution", and "I absolutely condemn".

	Input text	Classifier probability	Output label
0	speech	0.000173	office
1	speak	0.205794	office
2	announce	0.007564	office
3	just voted	0.292617	office
4	just completed	0.002853	office
5	hurricane sandy	1.000000	campaign
6	discussing	0.999080	campaign
7	condemn	0.999442	campaign
8	flood	0.000007	office
9	drought	1.000000	campaign
10	disapprove	0.000001	office
11	approve	0.978375	campaign

Figure 5: Predicted results from the classifier on specific input texts. Declaratory words are accurately classified as office, while critical or declamatory words are accurately classified as campaign. The neural network classifier model is able to reasonably make accurate distinctions between the language style members of Congress use in different time periods.

Most of these predictions line up with our intuition except for the word "flood". Ideally, "flood" and "drought" should be considered as similar features, but in fact they appear as two opposite extremes. The reason behind this hides deep within our model and training data which we do not have the time to explore within the constraints of this paper. Still, it can be reasonably inferred that they can lead to incorrect prediction results, and the outliers in the prediction of each class might be caused by similar features like this.

5. Discussion

Based on the model performance and the difference in predicted probabilities between the two classes, we can clearly confirm our hypothesis that politicians use different styles of language when they are running for reelection versus when they are in office. Finding out the exact cause of these differences on a linguistic level is a more difficult task. We were able to identify notably different features and verify their contribution to the prediction probability. However, to find out exactly how these features affect the classification, we would need more quantitative methods. In the future, we could use different models to map out the contributions of different features and how much weight they apply to the result probability. We could further optimize our findings by annotating the potentially influential variables that might affect politicians' language use in regards to certain events or time periods in the political season [7]. We could analyze the language on a morphological or syntactic level by using subword-based embeddings like in [3] rather than vector representations based on words. Our test accuracy could have been improved by combining our network-based feature predictions with lexical features, which has incidentally been proven to yield more accurate test results [9]. Finally, our model could have performed more accurate generalizations of the language use of politicians had we included cross-platform modeling [6] in our training data (i.e. our model uses Twitter data, but other social media services exist in which US members of Congress also frequently use).

We can infer from the results of the manually inputted lexemes that politicians adopt a critical and initiatory tone when campaigning. Conversely, politicians appear to make more clinical and optimistic statements when occupying office. Politicians broach more serious and emotionally affective topics when campaigning. Messages from politicians are more concerned with government when they are in office. From these incipient results, we can infer that members of Congress might be altering their language to pander to issues that voters care about when they are campaigning. When members of Congress are in office, however, they appear to abandon this rhetorical approach to their language use. Regardless, our results on the specific features influencing the "campaign" and "office" predictions are not statistically significant, so we cannot make a proven claim about the cause or decisions by members of Congress to alter their language across different date thresholds. However, our early results show promise.

That said, our main research is novel with respect to treating classification in the context of observing temporal differences for the language use of US members of Congress. We find significant differences in Twitter language use when members of the US House of Representatives are campaigning versus when they occupy incumbency. This research has implications for discourse analysis and the role of social media in influencing the language use of politicians. Future research can explore what these differences are on a linguistic level, the broader impetus for why this kind of "style-switching" occurs in politicians' language use, and how the intrinsic qualities of social media can influence the tide of politics—which in turn can influence a nation's outcomes—when paired with the motives or agenda of politicians.

References

[1] R. Ali, G. Pinto, E. Lawrie, and E. Linstead, "A large-scale sentiment analysis of tweets pertaining to the 2020 US presidential election," *Journal of Big Data*, vol. 9, no. 79, pp. 1-12, June 2022. Accessed: May 4, 2023. doi: 10.1186/s40537-022-00633-z. [Online]. Available: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00633-z

[2] J. Macrohon, C. Villavicencio, X. Inbaraj, and J. Jeng, "A semi-supervised approach to sentiment analysis of tweets during the 2022 Philippine presidential election," *Information*, vol. 13, no. 10, p. 484, Oct. 2022. Accessed: May 4, 2023. doi: 10.3390/info13100484. [Online]. Available: https://www.mdpi.com/2078-2489/13/10/484

[3] M. Pota, M. Esposito, M. Palomino, and G. Masala, "A subword-based deep learning approach for sentiment analysis of political tweets," in *32nd Int. Conf. Adv. Information Networking and Applications Workshops (WAINA)*, Krakow, Poland, 2018, pp. 651-656, doi: 10.1109/WAINA.2018.00162. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8418147/authors#authors

[4] Y. Theocharis, P. Barberá, Z. Fazekas, and S. Popa, "The dynamics of political incivility on Twitter," *SAGE Open*, vol. 10, no. 2, May 2020. Accessed: May 11, 2023. doi: 10.1177/2158244020919447. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/2158244020919447

[5] P. Rita, N. António, and A. Afonso, "Social media discourse and voting decisions influence: sentiment analysis in tweets during an electoral period," *SNAM*, vol. 13, no. 46, Mar. 2023. Accessed: June 6, 2023. doi: 10.1007/s13278-023-01048-1. [Online]. Available: https://link.springer.com/article/10.1007/s13278-023-01048-1

[6] A. Shevtsov, M. Oikonomidou, D. Antonakaki, P. Pratikakis, and S. Ioannidis, "What Tweets and YouTube comments have in common? Sentiment and graph analysis on data related to US elections 2020," *PLOS ONE*, vol. 18, no. 1, Jan. 2023. Accessed: June 6, 2023. doi: 10.1371/journal.pone.0270542. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0270542

[7] J. Choi, R. Park, Y. Baek, and J. Lee, "Unraveling the thought probes of US legislators on trade negotiations: sentiment analysis of the 108th and 112th US congressional speeches," *Inha Journal of International Studies*, vol. 38, no. 1, pp. 83-104, Mar. 2023. Accessed: May 4, 2023. doi: 10.1111/pafo.12220. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/pafo.12220

[8] K. Svetlov and K. Platonov, "Sentiment analysis of posts and comments in the accounts of Russian politicians on the social network," in *25th Conf. Open Innovations Association (FRUCT)*, Helsinki, Finland, 2019, pp. 299-305, doi: 10.23919/FRUCT48121.2019.8981501. [Online]. Available: https://ieeexplore.ieee.org/document/8981501

[9] A. Alshahrani, "Linguistic implications of political tweets," *IJEL*, vol. 10, no. 4, June 2020. Accessed: June 6, 2023. doi: 10.5539/ijel.v10n4p203. [Online]. Available: https://www.ccsenet.org/journal/index.php/ijel/article/view/0/43093

[10] A. Jungherr, "Twitter use in election campaigns: a systematic literature review," *J. Inf. Technol. Politics*, vol. 13, no. 1, pp. 72-91, Mar. 2016. Accessed: June 6, 2023. doi: 10.1080/19331681.2015.1132401. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/19331681.2015.1132401

[11] G. Abercrombie and R. Batista-Navarro, “Sentiment and position-taking analysis of parliamentary debates: a systematic literature review,” *JCSS*, vol. 3, no. 1, pp. 245-270, Jan. 2020. Accessed: June 6, 2023. doi: 10.1007/s42001-019-00060-w. [Online]. Available: https://link.springer.com/article/10.1007/s42001-019-00060-w

[12] B. Gervais, 2019, “Congressional Tweet Archive”, Harvard Dataverse. [Online]. Available: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BOK1CF