Lexical Analysis of Ideological Polarization of Political Discourse in Online Discussion Forums Over Time

Xavier Oliva Jurgens
xavi.oliva@berkeley.edu

Max Raphael Mynter
max.mynter@berkeley.edu

Abstract

With our final project, we want to evaluate the temporal change of political opinions of users or communities and their respective idiosyncratic language on Reddit, an online discussion forum.

1 Introduction

1.1 Motivation

The proper organization of humans in society has been a topic of discussion since the advent of speech. Throughout history, there have been many different forms of government. Various forms of societal organization - from kingdoms to city-states, communalism to theocracy, dictatorships to anarcho-syndicalism - have been implemented or at least proposed and discussed. With the development of the Internet, some of this discourse has moved online. Especially the ideological polarization of online political discourse on platforms such as Twitter, Facebook, and Reddit has been in the spotlight in recent years.

We want to leverage the digital traces of these discussions on Reddit to quantify the temporal evolution of users' political opinions over time. Additionally, we want to infer if certain radicalization pipelines are more prevalent than others, if there is a typical evolution of political views, and if certain events trigger a change in political opinion for many individuals at once. While studying these online groups, we also aim to extract the idiosyncratic language used within these communities and compare them across the political spectrum.

These study of these opinion shifts, group dynamics, and sociolects are at the heart of the social sciences - *How do people talk? What do they consider proper politics?* - and political practice - *What issues should be addressed on how to gather more votes in the upcoming election?*. Additionally, online radicalization poses a threat to communities as individuals might utilize violent means to hurt people they oppose. In recent years, online platforms have been pressured to curb extremist content and misinformation, so understanding the linguistic features of users and their evolution is an important field of research.

1.2 Related Work

Recent breakthroughs in vector-space representation provide new opportunities to represent words based on their context. In (An et al., 2018), word semantics are characterized using semantic axes in word-vector spaces that can be applied across communities and domains. The authors evaluate their framework using comparative text analysis to capture nuanced linguistic representations within Reddit subcommunities. They examine the semantic differences in diverse issues, such as immigration, minorities, and guns, revealing that some words have different connotations among distinct subgroups.

Additionally, our project is closely related to previous studies on the polarization of political discourse in NLP and computational social science. Recent work has introduced a framework for finding linguistic features maximally informative about the edge topology of a social network and use it to detect polarized concepts in online discussion forums (Hofmann et al., 2021). They use graph-based architectures to include information from social networks for NLP tasks using pre-trained models like BERT (Devlin et al., 2018). For ideological framing, they project contextualized embeddings into a *moral embedding subspace* based on the five moral foundations from moral foundations theory. Their experiments show that some subreddits have experienced a pronounced shift in their ideology, turning moderate discussion forums moving towards more extreme ideological positions.

Another related work is (Maynard and Funk, 2011) where the authors map Twitter posts onto the triple, (*name*, *opinion*, *party*), using sentiment analysis techniques. To achieve this, the authors first applied common NLP preprocessing techniques such as tokenization and then employed named-entity-recognition to infer the object (party) the tweets were about. The sentiment is inferred to be either positive, neutral, or negative. If the algorithm is uncertain, it was mapped to neutral as well.
The paper (Chambers et al., 2015) also evaluates twitter data, albeit with a slightly different focus. The authors assess the metadata of tweets to infer the geospatial position of the poster and infer the sentiment towards the target country from the text corpus. In that sense, the goal is similar to the previously discussed one, with the difference that sentiments towards countries, not parties, are quantified. A significant obstacle in this analysis, however, was the identification of irrelevant data. For example, tweets about turkey which concern the bird or a New Yorker lamenting about the traffic in Little *Italy* had to be excluded to create the aggregate country sentiment quantifier.

1.3 Problem Statement

With this work, we want to answer what events cause polarization or radicalization among active online forum users. We hypothesize that while the political opinion of a majority of users remains constant, some events or developments might trigger more radicalized points of view. We hypothesize that individual users or whole subreddits experience political shifts or experience radicalization over time. To obtain these insights, we want to quantitatively infer idiosyncratic codes used by followers of specific political ideologies.

2 Methods

2.1 Dataset

To conduct our research on the shift of political opinion, we want to evaluate posts within political communities on Reddit. We hope to use the Reddit Politosphere dataset that should be made available by (Hofmann et al., 2021).[1] This dataset contains a collection of comments on over 600 subreddits manually verified to deal with real-world politics over the span of 12 years (between 2008 and 2019).

Should the mentioned dataset be unavailable within a reasonable delay, substitute data is readily available on sites like `kaggle.com` or ConvoKit [2]. Alternatively, it can be directly gathered using the Pushshift API (Baumgartner et al., 2020).

2.2 Outline

After receiving the data, we will first identify individuals actively posting over extended (multi-year) periods and reduce the data to their posts and comments and clean the data of irrelevant texts. Since we do not have access to the dataset yet and do not know how many individual users have posted a significant amount of comments over a long time range, we might consider a higher granularity, analyzing groups of users instead of individuals.
To go about our quest of quantifying a shift in political opinion over time, we first need to operationalize the political view itself. To do so, we plan to capture domain-specific word semantics using existing frameworks like SEMAXIS with political poles as dimensions (An et al., 2018) and hope to find idiosyncratic language for distinct ideological groups. Then, using this knowledge, we aim to analyze the temporal changes of the political discourse of Reddit users.

2.3 Contributions

We intend to distribute the workload of the project equally. Open tasks are formulating the research question in more detail, retrieving and preprocessing data, identifying poles to score political opinions based on anchor baselines for certain ideologies, developing evaluation metrics, and visualizing the results.

2.4 Limitations

The analysis of online data allows for a comprehensive evaluation of a large sample of individuals at a comparably low time and other resources investment. This is especially true when we consider obtaining the same amount of data through developing and deploying a questionnaire for a similar sample size. Nonetheless, the results will be limited by the self-selection bias imposed by the active participation in online forums. Additionally, it might be that not enough individual data points over a long time interval are available.

[1] We are already in contact with the authors since the dataset has not been publicly released yet.

[2] `https://convokit.cornell.edu/documentation/subreddit.html`

References

Jisun An, Haewoon Kwak, and Yong-Yeol Ahn. 2018. Semaxis: A lightweight framework to characterize domain-specific word semantics beyond sentiment. *arXiv preprint arXiv:1806.05521*.

Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, pages 830–839.

Nathanael Chambers, Victor Bowen, Ethan Genco, Xisen Tian, Eric Young, Ganesh Harihara, and Eugene Yang. 2015. Identifying political sentiment between nation states with social media. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 65–75.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Valentin Hofmann, Janet B. Pierrehumbert, and Hinrich Schütze. 2021. Modeling ideological agenda setting and framing in polarized online groups with graph neural networks and structured sparsity.

Diana Maynard and Adam Funk. 2011. Automatic detection of political opinions in tweets. In *Extended Semantic Web Conference*, pages 88–99. Springer.