Modeling of Political Discourse on Twitter

Kristen Johnson
Department of Computer Science
Purdue University, West Lafayette, IN 47907
john1187@purdue.edu

Abstract

Our works focus on the application of NLP methods for the analysis of political discourse on Twitter. Our guiding intuition is that modeling the language used on Twitter alone is not enough for the most accurate prediction possible. Therefore, we explore how weakly supervised models can be constructed to leverage both language *and behaviors* of politicians on Twitter to identify stance and framing patterns from the discourse. By incorporating behavioral features, such as similar temporal information, stances on current and future political issues, as well as the frames used to express these issues, can be determined with higher accuracy than what is possible with language based models alone.

1 Introduction

During the 2016 United States presidential election, politicians frequently used Twitter to express their stances on current political issues. Due to the limited length of tweets and scrutiny faced by politicians for what they say on social media, politicians must carefully craft and time their tweets. The content and delivery of these tweets is therefore highly indicative of a politician's stances and frame choices. Such framing choices are an important political strategy which allows politicians to bias public perception and discussion of current issues towards their stance on the issue. For example, the debate around increasing minimum wage can be framed as a *quality of life* issue or as an *economic* issue. The first frame supports the increase because it improves workers' lives, but the second frame, by conversely emphasizing the costs involved, opposes the raise.

Prior works in stance and debate classification focus on supervised analysis of the text of individual tweets or forums (Sridhar et al., 2015; Hasan and Ng, 2014; Abu-Jbara et al., 2013; Walker et al., 2012; Abbott et al., 2011; Somasundaran and Wiebe, 2010, 2009; SemEval, 2016). Several works have also explored framing in longer text articles such as congressional speeches and news articles (Fulgoni et al., 2016; Tsur et al., 2015; Card et al., 2015; Baumer et al., 2015). Contrary to these sources, Twitter requires politicians to compress their ideas into 140 character long tweets. As a result, context is lost and frame prediction becomes a multilabel classification task since many frames may appear within one tweet.

Our works explore the intuition that an accurate stance or framing prediction requires a model which combines features of both a politician's behaviors and the language extracted from their tweets. Additionally, a supervised model can be too inflexible to accurately capture changing stances and political alliances, given the dynamic nature of Twitter. Therefore, we have designed weakly supervised holistic models which capture both the text and behavior of politicians on Twitter. These models have been applied in two settings: to predict behavior (i.e., stance) and to perform text analysis (i.e., predict frames). Our results show that progressively adding behavioral features to the language models improves results in both supervised and unsupervised experimental settings.

2 Method

To capture the stance of a politician or the frames that they use in their tweets, we designed models that incorporate different combinations of weakly supervised indicators. These indicators include language features such as the presence of unigrams, bigrams, or trigrams in the tweet. Behavioral features extracted from Twitter are also used

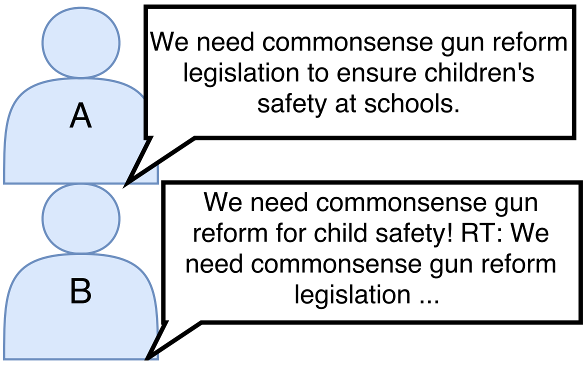

Figure 1: Simplified Example of Tweet Indicators.

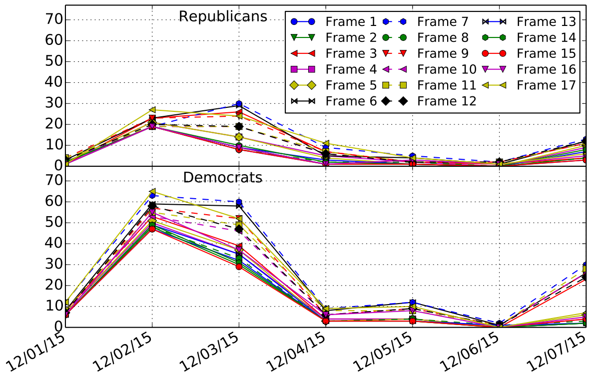

Figure 2: Frame Trends Over Time. The top and bottom panels show the predicted frames of gun related tweets from Republicans and Democrats, respectively. Frames 1-14 correspond to the frames of Boydstun et al. (2014) and Frames 15-17 are our proposed Twitter-specific frames (Johnson et al., 2017a,b).

as indicators. Such features include if the authors are in the same political party and if the tweets were posted around the same time, as well as if the authors of the tweets retweet or follow each other (or the same other people).

Figure 1 shows an example of tweets and the indicators that can be extracted from them. In this example, two Democrats are tweeting about the need for gun legislation. Politician A frames the need for gun control as a matter of safety (e.g., "to ensure child safety"). Politician B uses the same frame and retweets A. Finally, the trigram "commonsense gun reform", a popular phrase used by Democrats, is also present.

By combining these indicators into a single, joint model, we are able to make a stronger prediction than what is possible with individual indicators alone. For example, both tweets contain the unigram "gun", but this gives little indication of the politician's stance on gun control or how they will frame this issue. However, by combining the presence of unigram "gun" with trigram "commonsense gun reform", as well as shared party and similar frame indicators, the joint model is able to determine stance (e.g., this politician supports gun control) or frames (e.g., this politician frames gun control as a matter of safety).

3 Results

In our experiments, we used a supervised, lexical model as the baseline. Since our predictions are text-based decisions, similar to sentiment classification, we expect this baseline to perform well. Additionally, if our intuition that language alone does not produce the best prediction is correct, our models should outperform the baseline.

We combine the weakly supervised indicators into different holistic models designed to capture the aspects of tweets that we believe are relevant for modeling the decision (either stance or frame). Each model builds upon the previous model by adding more powerful language or behavioral indicators to the decision. For example, our first model uses presence of unigrams to determine the frame of a tweet. The next model uses unigrams and bigrams. The final model uses both language features (e.g., unigrams, bigrams, trigrams) as well as behavioral features (e.g., same time activity, retweet patterns, and the follower network) to achieve the highest prediction possible.

We analyzed our models under both supervised and unsupervised settings. For the task of stance prediction a supervised baseline achieves 47.6% average accuracy, while our language and behavior model achieves 86.44% (Johnson and Goldwasser, 2016). Similarly, for frame prediction the language baseline achieves an F_1 score of 55.21, while our joint model achieves an F_1 score of 75.95 (Johnson et al., 2017a,b). Given the high performance of our models under supervised settings, we are able to apply them to *unlabeled* tweets to observe changes in political discourse on Twitter over time. Figure 2 shows an example of frame trends over time around the San Bernadino, CA shooting. For brevity, the most interesting observation from this figure is that Democrats use the Safety Frame to push for gun control legislation, while Republicans use the same frame to redirect the discussion to focus on terrorist threats instead.

4 Conclusion

Our works present weakly supervised indicators and models for political discourse analysis for the tasks of stance and frame prediction. By incorporating Twitter behaviors, such as similar activity times, we are able to increase prediction accuracy over language-based models alone.

References

Rob Abbott, Marilyn Walker, Pranav Anand, Jean E. Fox Tree, Robeson Bowmani, and Joseph King. 2011. How can you say such things?!?: Recognizing disagreement in informal political argument. In *Proc. of the Workshop on Language in Social Media*.

Amjad Abu-Jbara, Ben King, Mona Diab, and Dragomir Radev. 2013. Identifying opinion subgroups in arabic online discussions. In *Proc. of ACL*.

Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta, and Geri Gay. 2015. Testing and comparing computational approaches for identifying the language of framing in political news. In *In Proc. of NAACL*.

Amber Boydstun, Dallas Card, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues.

Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proc. of ACL*.

Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preotiuc-Pietro. 2016. An empirical exploration of moral foundations theory in partisan news sources. In *Proc. of LREC*.

Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proc. of EMNLP*.

Kristen Johnson and Dan Goldwasser. 2016. All i know about politics is what i read in twitter: Weakly supervised models for extracting politicians' stances from twitter. In *Proc. of COLING*.

Kristen Johnson, Di Jin, and Dan Goldwasser. 2017a. Leveraging behavioral and social information for weakly supervised collective classification of political discourse on twitter. In *Proc. of ACL*.

Kristen Johnson, Di Jin, and Dan Goldwasser. 2017b. Modeling of political discourse framing on twitter. In *Proc. of ICWSM*.

SemEval. 2016. Task 6. http://alt.qcri.org/semeval2016/task6/.

Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proc. of ACL*.

Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological on-line debates. In *Proc. of NAACL Workshops*.

Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. 2015. Joint models of disagreement and stance in online debate. In *Proc. of ACL*.

Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proc. of ACL*.

Marilyn A. Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proc. of NAACL*.