MODELING LANGUAGE, SOCIAL, AND BEHAVIORAL ABSTRACTIONS FOR MICROBLOG POLITICAL DISCOURSE CLASSIFICATION

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Kristen Marie Johnson

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Dan Goldwasser, Chair

Department of Computer Science

Dr. Jennifer Neville

Departments of Computer Science and Statistics

Dr. Jean Honorio

Department of Computer Science

Dr. Xiangyu Zhang

Department of Computer Science

Approved by:

Dr. Voicu Popescu

Head of the School Graduate Program

For my great-grandparents, Edward and Leona Rodriguez, my grandfathers, Guy Edward Johnson, Sr. and Michael Angelo Biscardi, and my grandmother, Valda Dale Fontenot Biscardi.

ACKNOWLEDGMENTS

I am deeply grateful to my advisor, Dr. Dan Goldwasser, for his guidance throughout my PhD. He introduced me to the research area of Natural Language Processing and helped me to develop my understanding of this field while working towards multiple publications. His advice, knowledge, insight, and guidance have helped me tremendously during my graduate studies, while conducting research, and with developing my personal research and mentoring style. I hope to emulate the qualities that make him an amazing advisor when working with my future students, especially that of conducting meetings that are mostly about every topic except research.

My appreciation and thanks go to my committee members, Dr. Jennifer Neville, Dr. Jean Honorio, and Dr. Xiangyu Zhang, for their advice and encouragement during my qualifiers, preliminary exam, and defense. I would also like to give a special thanks to my labmates, Di Jin and I-Ta Lee, whose assistance with annotation and analysis made the works in Chapters 4 and 5 possible. I thank my Master's thesis advisor, Dr. Dongxiao Zhu, for initializing my interest in research and academia. Lastly, a heartfelt thanks to my first computer science professor and friend, Dr. Jamie Niño, for teaching me how to program, teach, and cook with style.

All of my love goes to my amazing family and friends who have encouraged me through all the sleepless deadlines and filled my days with laughter and happiness. Most importantly I thank my parents, Michelle Biscardi and Guy Johnson, Jr., who taught me that hard work and kindness will take me anywhere I want to go in life. To all of my family back home in New Orleans, especially the Newmans, Jeff Beck, Hope Redler, and Michelle Redler, thank you for your emotional support and all the late night tacos, crepes, and beignets. To my aunt, Theresa Liolios, thank you for being the best big sister ever. To my younger siblings, especially Courtney and Sam de Silva, thank you for always making me laugh.

I am fortunate to have made too many friends at Purdue to list them all here, but I would like to give special thanks to Adrian Thomas, Huda Nassar, Giselle Zeno, Ruben Varela, Victor Uceta, Radhika Bhargava, Yonghwi Kwon, Iman Alodah, Ruby Tahboub, Maurina Aranda, and Xuejiao Kang for their endless emotional and sometimes technical support, often over late night dinners or all day long in group chats. Finally, I want to thank all of my awesome friends in Atlanta, specifically Aline Pamboukian and Nancy Razzouk, for being my adventurous sisters and exploring Atlanta with me, from brunches to rooftop yoga.

I'm the luckiest woman on Earth to have you all in my life.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

Johnson, Kristen Marie Ph.D., Purdue University, August 2019. Modeling Language, Social, and Behavioral Abstractions for Microblog Political Discourse Classification. Major Professor: Dan Goldwasser.

Politicians are increasingly using social media platforms, specifically the microblog Twitter, to interact with the public and express their stances on current policy issues. Due to this nearly one-on-one communication between politician and citizen, it is imperative to develop automatic tools for analyzing how politicians express their stances and frame issues in order to understand how they influence the public. Prior to my work, researchers have focused on supervised, linguistic-based approaches for the prediction of stance or agreement of the content of tweets and classification of the frames and moral foundations used to express a single tweet. The generalizability of these approaches, however, is limited by the need for direct supervision, dependency on current language, and lack of use of social and behavioral context available on Twitter. My works are among the first to study these general political strategies specifically for politicians on Twitter. This requires techniques capable of abstracting the textual content of multiple tweets in order to generalize across politicians, specific policy issues, and time. In this dissertation, I propose breaking from traditional linguistic baselines to leverage the rich social and behavioral features present in tweets and the Twitter network as a form of weak supervision for studying political discourse strategies on microblogs. My approach designs weakly supervised models for the identification, extraction, and modeling of the relevant linguistic, social, and behavioral patterns of Twitter. These models help shed light on the interconnection of ideological stances, framing strategies, and moral viewpoints which underlie the relationship between a politician's behavior on social media and in the real world.

1 INTRODUCTION

1.1 Dissertation Statement

Beginning with the 2008 United States presidential election, Twitter has increasingly been used by all candidates to promote their agenda, interact with supporters and colleagues, and attack their opponents. These social media platforms allow politicians to quickly react to current events and gauge public interest in and support for their actions. Unlike its traditional media predecessors, Twitter requires politicians to compress their ideas, political stances, and reactions to 280 character-long tweets. Consequently, politicians must cleverly choose how to frame controversial issues, as well as how and when to react to each other. Due to this character limit, the stance, frame, or underlying moral foundation of a tweet is not independent of the social context in which it was generated. Thus, for accurate predictions, social relationships and behaviors must also be modeled.

The central role of social media platforms in today's political discourse emphasizes the importance of constructing automated tools for analyzing this content. From a technical perspective, the dynamic settings in which this content is generated raises new challenges. On the one hand, the language used to discuss new events and political agendas continuously changes, forcing automated tools to constantly adapt, i.e., the language used to discuss issues today will not be the same as the language used to discuss the same (or new) issues tomorrow. On the other hand, the rich social interactions on Twitter can be leveraged to provide a powerful alternative to direct supervision.

In order to accurately predict and classify the political strategies used in microblogs, a modeling approach is needed that is as independent of language as possible and capable of incorporating the social and behavioral relationships of politicians on

Twitter into the prediction. By not directly relying on individual words and using the relationships between politicians in the Twitter network, a model has access to more abstract, higher-level features. These abstract features act as a form of weak supervision, meaning the model will be less reliant on annotated, or manually labeled, data for learning and prediction. This naturally leads to weakly supervised models that are capable of generalizing over time, as they are no longer dependent on annotation, language, specific individuals, or specific issues.

In this dissertation, I present the effectiveness of this weakly supervised modeling approach as proven with results on three political discourse prediction tasks across a variety of policy issues frequently discussed on Twitter: (1) predicting politician stance and agreement patterns, (2) tweet framing classification, a very nuanced political discourse analysis task, and (3) classification of the moral foundations used to express ideologies in tweets. Instead of solely relying on linguistic information, I have designed natural language processing (NLP) based models to exploit the rich social and behavioral context in which the tweets were generated in order to make more accurate predictions and reduce the amount of supervision required. My works show that modeling social and behavioral features in addition to language features improves F_1 prediction scores in both supervised and unsupervised settings for all prediction tasks.

Given the highly dynamic nature of political discourse on Twitter, relying on traditional approaches, such as bag-of-words features, would require supervised models dependent on constant manual annotation. However, manual annotation of large quantities of tweets is too time consuming and can often result in conflicting annotations. To overcome these challenges, I design weakly supervised models for extracting meaningful patterns from the data and build the classifiers over the output of these models, rather than over the raw data directly. Depending on the prediction task, the weakly supervised models use the following types of information: (1) *directly observable information* such as political party affiliation and the issue discussed in the tweet, (2) *linguistic information* defined over relevant keywords and frequent party slogans, represented as bigrams and trigrams, and (3) *social or behavioral information* such as similar temporal activity, retweet patterns, and the follower network. These features are then declaratively compiled into a graphical model using Probabilistic Soft Logic (PSL), a recently introduced probabilistic modeling framework.[1] As described in Chapter 2, PSL specifies high level rules over a relational representation of these features. These rules are then compiled into a graphical model called a hinge-loss Markov random field (Bach et al., 2013), which is used to make the final stance, agreement, or frame prediction. Figure 1.1 shows this modeling pipeline.

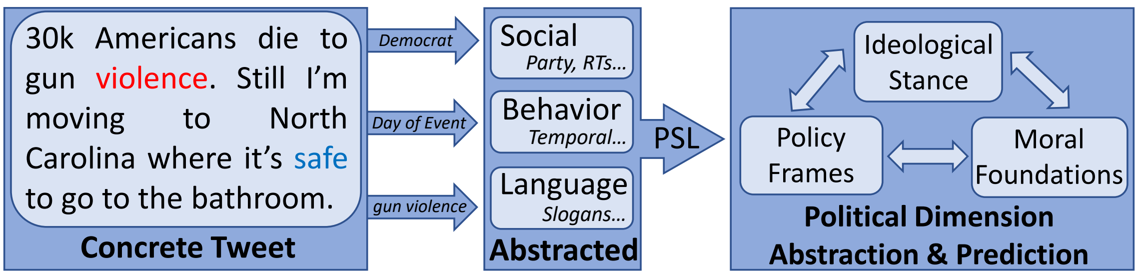


Figure 1.1. Abstraction Modeling Pipeline. Concrete words are extracted from tweets and represented as abstract features. These features are then combined in PSL models to predict higher-level abstractions of political discourse.

This approach allows us to model dependencies by connecting Twitter users who have language similarities, social connections, or behavioral similarities. Language similarities indicate that the same issue has been discussed or the presence of political slogans in tweets, which may further indicate shared (or differing) underlying political ideologies. Social connections are directed dependencies that represent the followers of each user as well as retweeting behavior (i.e., user A retweets user B's content). Interestingly, such social connections capture the flow of influence within political parties; however, the number of connections that cross party lines is extremely low. Instead, we rely on capturing behavioral similarity between users to provide this information. For example, users whose Twitter activity peaks at similar times tend

[1]PSL is available at: `http://psl.cs.umd.edu`.

to discuss issues in similar ways, providing indicators of their similar stance or frame usage for those issues.

1.2 Dissertation Contributions

First, I present natural language processing techniques for exploiting the social context of tweets. Traditional approaches assume full supervision is available, i.e., all tweets are labeled with the prediction target. However, acquiring millions of labels is not always possible or feasible. Twitter provides access to additional information: social context. Using the available social information reduces the burden to obtain annotations, which will allow the model to handle the dynamic, constantly changing language of Twitter. This further allows the model to adapt over time to new politicians and political issues without additional supervision.

Second, I propose viewing the political strategies of policy framing and moral foundations as a new type of abstraction over textual content. Different from previous works that have explored specific instances of framing, I adapt guidelines from political science to study general framing in a social media setting. Furthermore, moral foundations are typically classified through self-reporting questionnaires for ideological studies, but in this dissertation I adapt the theory guidelines for the classification of the moral foundations used to express ideological stances in tweets. Framing and moral foundations provide a more abstract view of language and behavior and thus help to reduce the need for annotation. Furthermore, when used as model features, these two political aspects act as forms of weak supervision for the prediction of stance as well as each other.

Finally, I provide models that capture the raw data and present it in a form that is conducive to learning. I show that modeling the social and behavioral information present in Twitter results in more accurate predictions for the tasks of stance and agreement prediction and frames and moral foundations classification than approaches relying on linguistic features alone.

1.2.1 Stance and Agreement Prediction

Stance and agreement prediction is a popular task in natural language processing, specifically for studying online debate forums. Stance is the position a person takes on an issue, usually predicted in the form of supporting or not supporting an issue. When choosing which politician to vote for, knowing their stance on an issue can help the public make the best decision. However, voter assisting websites do not always have a clear-cut yes or no, support or does not support, stance for every politician on every issue. Therefore, tools which can automatically determine a politician's stance on an issue, based off of what they say online, would be beneficial for citizens.

During the 2016 United States presidential election, politicians increasingly began to use Twitter to express their beliefs, stances on current political issues, and reactions concerning national and international events. Given the limited length of tweets and the scrutiny politicians face for what they choose or neglect to say, they must craft and time their tweets carefully. The content and delivery of these tweets is therefore highly indicative of a politician's stances. In this chapter, I will present a weakly supervised method for extracting social and behavioral information, such as how issues are framed and the temporal activity patterns on Twitter for politicians, to determine their stances on popular issues of the 2016 election. These behavioral components are combined into a global model which collectively infers the most likely stance and agreement patterns among politicians, with respective accuracies of 86.44% and 84.6% on average.

1.2.2 Political Discourse Frame Classification

Framing is a political strategy in which politicians carefully word their statements in order to control public perception and discussion of policy issues. This technique allows politicians to *spin* an issue in such a way as to influence how the public or traditional forms of media will view and discuss the issue. Modeling and understanding how politicians use language and behavior to frame their stance on an issue would help inform the public when political bias is present.

Previous works exploring political framing typically analyze frame usage in longer texts, such as congressional speeches. In this chapter, I will present a collection of weakly supervised models which harness collective classification to predict the frames used in political discourse on Twitter. My global probabilistic models show that by combining a variety of lexical features extracted from tweets with network-based behavioral features of Twitter, the average, unsupervised F_1 score increases by 21.52 points over a lexical baseline alone.

1.2.3 Moral Foundations Classification

Previous works in computer science, as well as political and social science, have shown correlation in text between political ideologies, framing strategies, and the moral foundations (Graham et al., 2009) expressed within that text. Based on these associations, this chapter first presents models exploiting both the language and how politicians frame issues on Twitter, in order to predict the moral foundations that are used by politicians to express their stances on issues. However, given the length restriction of tweets, politicians must carefully word their statements to ensure their message is understood by their intended audience. This constraint often eliminates the context of the tweet. To overcome this lack of information, I have designed relational models which combine high-level abstractions of political language and politician behavior to reveal the moral foundations underlying the discourse of United States politicians online, *across* differing governing administrations, revealing how party talking points remain cohesive or change over time.

1.3 Dissertation Organization

This dissertation will present the evolution of the works presented in each chapter. I will highlight the usefulness of social and behavioral features for the prediction of three types of political behavior. The overall organization of this dissertation is as follows:

- Chapter 1 has introduced the benefits and unique setting of political discourse analysis on Twitter. I have presented a high-level overview of my contributions to this area.

- Chapter 2 presents relevant previous research efforts as well as the technical details of the Probabilistic Soft Logic (PSL) graphical model used in my works.

- Chapter 3 focuses on my work on stance and agreement prediction. This chapter introduces the weakly supervised modeling approach used throughout this dissertation, which combines linguistic content and behavioral features for analyzing and understanding the political Twitter domain. A first step towards frame classification and its potential for real-world behavior analysis is discussed. The results of this chapter show the effectiveness of this modeling approach for the prediction of politician stances and agreement patterns.

- Chapter 4 presents the combined results of my work towards solving the problem of frame classification. This chapter presents the problem of frame classification, provides guidelines for frame annotation, as well as an annotated dataset, and builds a strong linguistic baseline. The results of this chapter support the motivating idea behind my research – that social and behavioral information combined with linguistic information results in more stable models and better predictions. Finally, I present interesting real world applications of frame classification of events over time and aisle-crossing voting behavior.

- Chapter 5 presents the application of accumulated findings from previous works for the classification of the moral foundations used to express political stances on issues discussed in tweets. The contributions of this section include a dataset annotated with the moral foundations, annotation guidelines, and probabilistic graphical models which show the usefulness of jointly modeling political slogans, as opposed to the unigrams of previous works, with social and behavioral features for the prediction of the morality underlying political tweets.

- Chapter 6 concludes this dissertation. This chapter summarizes the key findings of my publications and emphasizes the interconnection of the three facets of political discourse studied in this dissertation: ideological stance, policy framing strategies, and the moral foundations underlying political discourse.

The work presented in Chapter 3 was done in collaboration with my advisor, Dr. Dan Goldwasser. Chapter 4 presents collaborative work with my advisor and labmate, Di Jin. Finally, the works of Chapter 5 were a collaborative effort with my advisor and labmate, I-Ta Lee.

2 RELEVANT BACKGROUND

2.1 Related Works

Over recent years there has been a growing interest in analyzing political discourse on social media microblogs. In this dissertation, I explore how political ideology, language, framing, and morality are expressed and interact with each other in political discourse on Twitter. Modeling and understanding these aspects of political discourse draws on previous research in text, opinion, and network analysis, as well as human behavior analysis, from a multitude of areas including: computer science, political science, psychology, sociology, and communications.

2.1.1 Understanding Stance and Opinions Through Text

Chapter 1 presents the first work predicting the stances of politicians using Twitter data based on content, frames, and temporal activity (Johnson and Goldwasser, 2016). This task is related to the traditional natural language processing tasks of opinion mining and stance prediction but is applied to Twitter analysis and specifically studies the language of politicians. Several previous works have studied mining opinions and predicting stances in online debate forum data by exploiting argument and threaded conversation structures, both of which are not always present in short Twitter data.

Somasundaran and Wiebe (2009) first used associations between people and discourse indicators for opinion analysis, specifically to find an individual's stance. Later, the authors extended this work to show that more abstract features of textual information, such as indicators of opinions and sentiment, could be used to predict stance (Somasundaran and Wiebe, 2010). Abbott et al. (2011) showed that using metadata, as well as contextual and dependency features of forums, increased stance disagreement prediction accuracy over a unigram baseline. By using opinion mining techniques, Abu-Jbara et al. (2013) were able to show how debate participants split into subgroups with differing opinions. Hasan and Ng (2014) take stance and opinion identification one step further by designing models capable of determining the reasoning behind a particular stance. Sridhar et al. (2015) determined that the best modeling approach for online dialogues is a collective, probabilistic framework due to the interrelated and dependent nature of stances online. Probabilistic modeling supports reasoning about agreement and disagreement, collective modeling of dialog structure, and stance classification at both the author and post level for forums.

In addition to identifying the stance of the author of a post in an online forum, there has also been work towards identifying the stance of the post itself. Walker et al. (2012) showed that dialog structure could be represented by collectively modeling the agreement relationships among post authors in order to classify the positive or negative stance of online debate posts. Two recent works (SemEval, 2016; Ebrahimi et al., 2016) aimed to detect the stance of *individual tweets*. In contrast to this task, as well as the aforementioned related work on debate stance prediction, the models presented in Chapter 3 *do not assume that each tweet expresses a stance*. Instead, these models are used to investigate how a politician's overall Twitter behavior, as represented by combined content and temporal indicators, is indicative of a stance, e.g., by also capturing when politicians *fail to tweet about a topic*.

2.1.2 Using Text to Reveal Morality and Framing Strategies

Compared to stance prediction, frame classification is a more difficult, finer granularity task and describes how someone expresses their view on an issue, not whether they support the issue. Issue framing is related to the broader natural language processing challenges of analyzing biased language and subjectivity, i.e., the aspects of language that are used to express opinions. Recasens et al. (2013) identified framing bias as "perspective-specific words", collected common linguistic cues indicating this bias, and developed models to identify the bias-inducing word. Tan et al. (2014) analyzed how wording affects message propagation on Twitter, an important aspect for determining how issues are framed on social networks. Other works have focused on specific issues, e.g., the debate over genetically modified organisms (GMOs), to determine if opposing sides in the debate adopted more or less scientific tones when discussing the issues (Choi et al., 2012). Rather than look at individual words or phrases, Greene and Resnik (2009) proposed using how ideas are structured, or packaged, together for implicit sentiment, or perspective, classification. Wiebe et al. (2004) identified indicators of subjectivity and showed that the more frequently these indicators appeared in the surrounding context of a word, the more likely that word was subjective, and thus useful for predicting opinions.

Several previous natural language processing works have explored framing in public statements or congressional speeches (Tsur et al., 2015) and news articles (Card et al., 2015; Baumer et al., 2015; Fulgoni et al., 2016). Tsur et al. (2015) used probabilistic topical models to analyze framing and agenda setting in public statements from members of the United States Congress to understand the effects of framing techniques on party divergence across four topics. While Card et al. (2015) and Fulgoni et al. (2016) studied framing trends in news articles, Baumer et al. (2015) explored different word and sentence features in order to determine the most useful indicators for automatically identifying framing language. Other works focus on identifying and measuring political policies by jointly modeling framing language during legislative debates and voting patterns (Nguyen et al., 2015) or by using probabilistic topic models of legislative bills to predict stances and voting behavior (Gerrish and Blei, 2012).

The modeling and annotation approach presented in Chapter 4 (Johnson et al., 2017a,b) builds upon the previous work on frame analysis by Boydstun et al. (2014), by adapting and applying their Policy Frames Codebook for Twitter. This codebook is a set of annotation guidelines for the labeling of general, issue-independent frames of longer texts. These guidelines and frames were extended for use on Twitter political discourse and studied in a computational setting as described in Chapter 4.

Several works from political and social science research have studied the role of Twitter and framing in molding public opinion of certain events, e.g., the Vancouver riots (Burch et al., 2015) and the Egyptian protests (Harlow and Johnson, 2011; Meraz and Papacharissi, 2013). Others have covered framing and sentiment analysis of opponents (Groshek and Al-Rawi, 2013) and network agenda modeling (Vargo et al., 2014) in the 2012 U.S. presidential election. Jang and Hart (2015) studied frames used by the general population specific to global warming. In contrast to these works, this dissertation presents models which predict the *issue-independent* general frames of tweets, by U.S. politicians, which discuss six different policy issues (Johnson et al., 2017b,a).

The connection between morality and political ideology has been studied in the fields of psychology and sociology (Graham et al., 2009, 2012). The Moral Foundations Theory has been classified via applications of the Moral Foundations Dictionary (MFD) to identify moral foundations in partisan news sources (Fulgoni et al., 2016) and to construct features for other downstream tasks (Volkova et al., 2017). Several recent works have explored using data-driven methods that go beyond the Moral Foundations Dictionary to study tweets related to specific events, such as natural disasters like Hurricane Sandy (Garten et al., 2016; Lin et al., 2017). Furthermore, Fulgoni et al. (2016) expound upon the underlying moral foundations that drive the framing strategies used by different political parties to discuss policy issues.

2.1.3 Social Media Microblog Analysis

Unsupervised and weakly supervised models of Twitter data for several various tasks have been suggested, including: profile (Li et al., 2014b) and life event extraction (Li et al., 2014a), conversation modeling (Ritter et al., 2010), and methods for dealing with the unique language used in microblogs (Eisenstein, 2013).

Predicting political affiliation and other characteristics of Twitter users has been explored (Volkova et al., 2015, 2014; Conover et al., 2011). Other works have focused on political sentiment analysis (Pla and Hurtado, 2014; Bakliwal et al., 2013), predicting ideology (Iyyer et al., 2014; Bamman and Smith, 2015; Sim et al., 2013; Djemili et al., 2014), analyzing types of tweets and Twitter network effects around political events (Maireder and Ausserhofer, 2013), automatic polls based on Twitter sentiment and political forecasting using Twitter (Bermingham and Smeaton, 2011; O'Connor et al., 2010; Tumasjan et al., 2010), and distant supervision applications (Marchetti-Bowick and Chambers, 2012).

2.1.4 Social Network Modeling

Works focusing on inferring signed social networks (West et al., 2014), social group modeling (Huang et al., 2012), and PSL collective classification (Bach et al., 2015) are most relevant to the modeling approach presented in this dissertation. However, these approaches typically operate in supervised settings. In this dissertation, I propose new models that work in weakly supervised and unsupervised settings in order to overcome the annotation cost of continually labeling massive quantities of tweets.

2.2 Global Modeling Using Probabilistic Soft Logic (PSL)

PSL is a declarative modeling language which can be used to specify weighted, first-order logic rules. These rules are then combined into predicates to form PSL models. In PSL notation, P_1, P_2, P_3, and P_4 represent predicates (e.g., political party, issue, presence of n-grams, and frame) and x, y are variables. Each rule is composed of observed predicates added together on the left hand side and a target predicate on the right hand side of the rule. Each rule also has a weight λ which reflects the importance of that rule and is learned using the Expectation-Maximization algorithm in our unsupervised experiments. Using concrete constants *a, b* (e.g., tweets and words) which instantiate the variables x, y, model atoms are mapped to continuous

[0,1] assignments. More important rules (i.e., those with larger weights) are given preference by the model.

These rules are compiled into a hinge-loss Markov random field which defines a probability distribution over possible continuous value assignments to the random variables of the model (Bach et al., 2015). These continuous value assignments differ from other probabilistic logical models, e.g. MLNs, in which the random variables of the model are strictly true or false. This probability density function is represented as:

$$P(\mathbf{Y} \mid \mathbf{X}) = \frac{1}{Z} \exp\left(- \sum_{r=1}^{M} \lambda_r \phi_r(\mathbf{Y} \ , \ \mathbf{X}) \right)$$

where Z is a normalization constant, λ is the weight vector, and

$$\phi_r(\mathbf{Y}, \mathbf{X}) = (\max\{l_r(\mathbf{Y}, \ \mathbf{X}), 0\})^{\rho_r}$$

is the hinge-loss potential specified by a linear function l_r. The exponent $\rho_r \in 1, 2$ is optional. Each potential represents the instantiation of a rule, which takes the following form using the predicates, variables, and weights as described previously:

$$\lambda_1 : P_1(x) \wedge P_2(x, y) \rightarrow P_3(y)$$
$$\lambda_2 : P_1(x) \wedge P_4(x, y) \rightarrow \neg P_3(y)$$

2.3 Evaluation Metrics

Since each tweet can have more than one frame or moral foundation label, the prediction task is viewed as a multilabel classification task throughout this dissertation. The precision of a multilabel model is the ratio of how many predicted labels are correct:

$$Precision = \frac{1}{T} \sum_{t=1}^{T} \frac{|Y_t \cap h(x_t)|}{|h(x_t)|} \tag{2.1}$$

The recall of this model is the ratio of how many of the actual labels were predicted:

$$Recall = \frac{1}{T} \sum_{t=1}^{T} \frac{|Y_t \cap h(x_t)|}{|Y_t|} \tag{2.2}$$

In both formulas, T is the number of tweets, Y_t is the true label for tweet t, x_t is a tweet example, and $h(x_t)$ are the predicted labels for that tweet. The F_1 score is computed as the harmonic mean of the precision and recall.

3 STANCE AND AGREEMENT PREDICTION

Converse to previous works which predict stance per individual tweet (SemEval, 2016), this chapter presents a novel approach better suited to model the dynamic political arena of Twitter, by using the *overall* Twitter behavior per politician to predict a *politician's* stance on an issue. This section explores two aspects of the problem, *stance* prediction over a wide array of issues, as well as stance *agreement and disagreement* patterns between politicians over these issues. While the two aspects are related, they capture different information, as identifying agreement patterns reveals alliances and rivalries between candidates, across and within their party. In an extreme case, even the lack of Twitter activity on certain issues can be indicative of a stance.

For example, consider the three tweets on the issue of gun control shown in Figure 3.1. To identify the stance taken by each politician, the PSL model combines both content and behavioral features, accumulated from all of a politician's tweets on that issue. First, the tweet's relevance to an issue can be identified using *issue* indicators (highlighted in green). Second, the similarity between the stances taken by two of the politicians (agreement) can be identified by observing differences in how the issue is *framed* (shown in yellow), a tool often used by politicians to create bias toward a stance and contextualize the discussion (Tsur et al., 2015; Card et al., 2015). Tweets (1) and (3) frame the issue as a matter of safety, while tweet (2) frames it as pertaining to personal freedom, thus revealing the agreement and disagreement patterns between the politicians. Third, we can consider the timing of these tweets, i.e., whether these tweets are posted continually or just around events concerning gun violence. Finally, we can also use sentiment indicators (e.g., the negative sentiment of tweet (1)). Notice that each feature individually might not contain sufficient information for correct classification, but combining all aspects, by propa-

(1) **Hillary Clinton (@HillaryClinton):** *We need to keep guns out of the hands of domestic abusers and convicted stalkers.*

(2) **Donald Trump (@realDonaldTrump):** *Politicians are trying to chip away at the 2nd Amendment. I won't let them take away our guns!*

(3) **Bernie Sanders (@SenSanders):** *We need sensible gun-control legislation which prevents guns from being used by people who should not have them.*

Figure 3.1. Tweets Discussing the Issue of Gun Control. Issue indicators (e.g., guns and gun-control) are highlighted in green and different frame indicators (e.g., domestic abusers or 2nd Amendment) are highlighted in yellow.

gating stance bias (e.g., from sentiment) to politicians who hold similar or opposing views (as determined from frame analysis), leads to a more reliable prediction.

Given the dynamic nature of political discourse on Twitter, we design the modeling approach to use minimal supervision and naturally adapt to new issues. First, several weakly supervised local learners are built, whose only supervision is a small seed set of issue and frame indicators which characterize the stance of tweets (based on lexical heuristics (O'Connor et al., 2010) and framing dimensions (Card et al., 2015)), and Twitter activity statistics which capture temporally similar patterns between politicians. The final model represents agreement and stance bias by combining these weak models into a weakly supervised joint model through Probabilistic Soft Logic (PSL), a recent probabilistic modeling framework (Bach et al., 2013). The information gained from the weakly supervised local learners is the only supervision used by PSL; the rest of the prediction is completely unsupervised. PSL combines these aspects declaratively by specifying high level rules over a relational representation of the politicians' activities (exemplified in Figure 3.2), which is further compiled into a graphical model called a hinge-loss Markov random field (Bach et al., 2013), and used to make predictions about stance and agreement between politicians.

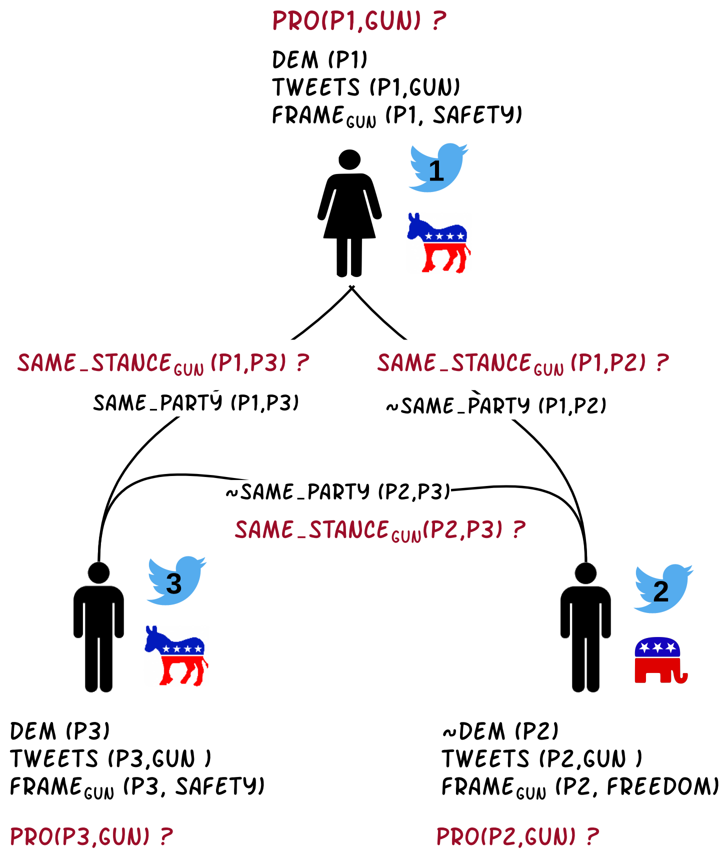


Figure 3.2. Relational Representation Example of Twitter Activity. P1, P2, and P3 represent three different politicians. Prediction target predicates (PRO and SAMESTANCE) are shown in red. Indicators of Twitter content and behavior include: DEM, TWEETS, FRAME$_{\text{GUN}}$, SAMEPARTY. GUN refers to the issue of gun control; SAFETY and FREEDOM refer to different frames for the issue.

The models are used to analyze the Twitter activity of 32 prominent U.S. politicians, including those who were U.S. 2016 presidential candidates, on 16 different issues. The experiments demonstrate the effectiveness of this global modeling approach, which outperforms both the weak learners that provide the initial supervision and a supervised text based baseline. Furthermore, the results verify that understanding political discourse on Twitter requires modeling not only the word content of tweets but the social behavior associated with those tweets as well.

3.1 Data Collection

Using the Twitter API, we collected tweets for 32 politicians: the 16 Republicans (all 2016 presidential candidates) and 16 Democrats (5 of which were candidates) listed in Table 3.1. The initial goal was to compare politicians participating in the 2016 U.S. presidential election. To increase representation of Democrats, tweets were collected from Democrats who hold leadership roles within their party, because more well known politicians tend to focus their tweets on national rather than local (district/state) events. For all 32 politicians there is a total of 99,161 tweets: 39,353 Democrat and 59,808 Republican[1].

Based on tweet availability and politician coverage[2], we chose 16 issues (shown in Table 3.2) derived from the 58 questions used by `ISideWith.com`, which match a user to politicians based on their responses, as our stance prediction goals. These issues range over common policies including domestic and foreign policy, economy, education, environment, health care, immigration, and social issues.

[1]Our Twitter dataset, keywords, and PSL scripts are available at: `purduenlp.cs.purdue.edu/projects/politicaltwitter`.

[2]For each of these 16 issues, at least 15 (with an average of 26) of the 32 politicians have tweeted on that issue; for the remaining issues, we found fewer than half (or none) of the politicians tweeted about that issue.

Table 3.1.
Politicians Tracked in This Chapter. All Republicans ran as 2016 presidential candidates. Democrats are divided based on whether or not they ran as a candidate.

Republicans	Jeb Bush, Ben Carson, Chris Christie, Ted Cruz, Carly Fiorina, Lindsey Graham, Mike Huckabee, Bobby Jindal, John Kasich, George Pataki, Rand Paul, Rick Perry, Marco Rubio, Rick Santorum, Donald Trump, Scott Walker	
Democrats	Candidates:	Lincoln Chafee, Hillary Clinton, Martin O'Malley, Bernie Sanders, Jim Webb
	Non-candidates:	Joe Biden, Kirsten Gillibrand, John Kerry, Ben Lujan, Ed Markey, Nancy Pelosi, Harry Reid, Chuck Schumer, Jon Tester, Mark Warner, Elizabeth Warren

3.1.1 Data Preprocessing

Using all tweets, we compiled a set of frequent keywords (with an average of seven) for each issue. This set is small to avoid overselection, i.e., avoiding tweets about praying for a friend's *health* but keeping tweets discussing *health care*. Via Python scripts, these keywords are used to retain tweets related to the 16 issues shown in Table 3.2, while eliminating all irrelevant tweets (e.g., those about personal issues, campaigning, duplicates, and non-English tweets).

`ISideWith.com` uses a range of yes/no answers to their questions and provides proof (through quotes or voting records) of a politician's stance on that issue, *if available*. When unavailable, the site assigns an answer based on party lines or often provides no answer. Also, less popular politicians are not featured on the site. For these cases, we manually annotated the stance using online searches of newspapers or voting records. These stances are only used for evaluation of our predictions. Our weakly

Table 3.2.
Sixteen Political Issues Used in This Chapter. Issues and their corresponding Yes/No questions were taken from `ISideWith.com`.

Issue	Question
Abortion	*Do you support abortion?*
ACA	*Do you support the Patient Protection and Affordable Care Act (Obamacare)?*
Confederate	*Should the federal government allow states to fly the confederate flag?*
Drugs	*Do you support the legalization of Marijuana?*
Environment	*Should the federal government continue to give tax credits and subsidies to the wind power industry?*
Guns	*Do you support increased gun control?*
Immigration	*Do you support stronger measures to increase our border security?*
Iran	*Should the U.S. conduct targeted airstrikes on Iran's nuclear weapons facilities?*
ISIS	*Should the U.S. formally declare war on ISIS?*
Marriage	*Do you support the legalization of same sex marriage?*
NSA	*Do you support the Patriot Act?*
Pay	*Should employers be required to pay men and women, who perform the same work, the same salary?*
Religion	*Should a business, based on religious beliefs, be able to deny service to a customer?*
Social Security	*Should the government raise the retirement age for Social Security?*
Student	*Would you support increasing taxes on the rich in order to reduce interest rates for student loans?*
TPP	*Do you support the Trans-Pacific Partnership?*

supervised approach requires *no* prior knowledge of the politician's stance, allowing it to generalize to situations such as these, where stance information is unavailable.

3.1.2 Prediction Goals

The collected stances represent the ground truth of whether a politician is for or against an issue. Based on these we define two target predicates using PSL notation to capture the desired output as soft truth assignments to these predicates. The first predicate, PRO(P1, ISSUE), captures a positive stance by politician P1, on an ISSUE. A negative stance would be captured by its negation: ¬PRO(P1, ISSUE). The second target predicate, SAMESTANCE_I(P1, P2), classifies if two politicians share a stance for a given issue, i.e., if both are for or against an issue, where I represents one of the sixteen issues of interest. Although the two predicates are clearly inter-dependent, we chose to model them as separate predicates since they can depend on different Twitter behavioral and content cues. Given the short and context-free style of Twitter we can often find indicators of politicians holding similar stances, *without* clear specification for which stance they actually hold.

3.2 Local Models of Twitter Activity

The only supervision required by our method consists of the keywords describing issues and frames, Twitter behavior patterns, and party affiliation, all of which is easily attainable and adaptable for new domains (e.g., different keywords can be used to capture issues of another country). The weakly supervised local models described in this section capture similarities between tweet content and temporal activity patterns of users' timelines, as well as stance bias, and are used to provide the initial bias when learning the parameters of the otherwise unsupervised global PSL model.

3.2.1 Issue of Tweets

To capture which issues politicians are tweeting about, we used a keyword based heuristic, similar to the approach described in O'Connor et al. (2010), where each issue is associated with a small set of pre-selected keywords (as described previously). The keywords appearing in a given tweet may be mutually exclusive (e.g., *fracking* for Environment will not appear in tweets discussing other issues); however, some may fall under multiple issues at once (e.g., *religion* may indicate the tweet refers to ISIS, Religion, or Marriage). Tweets are classified as relating to a certain issue based on the majority of matching keywords, with rare cases of ties manually resolved. The output of this classifier is all of the issue-related tweets of a politician, which are used as input for the PSL predicate TWEETS(P1, ISSUE), a binary predicate which indicates if that politician has tweeted about the issue or not.

3.2.2 Sentiment of Tweets

The sentiment of a tweet can indicate a politician's stance on a certain issue. OpinionFinder 2.0 (Wilson et al., 2005) is used to label each politician's issue-related tweets as positive, negative, or neutral. We observed, however, that for all politicians, a majority of tweets will be labeled as neutral. This may be caused by the difficulty of labeling sentiment for Twitter data. When this results with a politician having no positive or negative tweets, they are assigned their party's majority sentiment for that issue. The majority sentiment of a party is calculated by running all party politicians' tweets through OpinionFinder and using whichever sentiment (positive or negative) is assigned the most per party. This output is used as input to the PSL predicates TWEETPOS(P1, ISSUE) and TWEETNEG(P1, ISSUE).

3.2.3 Content Agreement and Disagreement Patterns

We expect politicians that have a similar stance on an issue to use similar words in their tweets. To determine how well tweet content similarity captures agreement between politicians, we computed the pair-wise cosine similarity between all of the politicians' words used in tweets per issue. However, the use of similar words per issue resulted in most politicians being grouped together, even across different parties. To overcome this, we calculated the *frequency* of similar words within tweets (per issue). For each issue, all of a politician's words from tweets are aggregated and the frequency of each word is compared to all other politicians' word frequencies. Politicians, P1 and P2, are considered to have a similar $\textsc{LocalSameStance}_I(\mathrm{P1},\ \mathrm{P2})$ if their frequency counts per shared word of an issue I are within the same range. For this study, we used a window of 10 (i.e., if the frequency count of a word is 30, then a count of 20 to 40 would be considered similar) to ensure that politicians who briefly mention an issue are not considered equivalent to those who discuss it more frequently.

3.2.4 Temporal Activity Patterns

We observed from reading Twitter feeds that most politicians tweet about an event the day it happens. However, for general issues, politicians will comment as frequently as desired to express their support or lack thereof for that particular issue. For example, Rand Paul tweeted daily in opposition of the NSA during his filibuster of the Patriot Act renewal. Conversely, Hillary Clinton has no tweets concerning the NSA or Patriot Act. To capture agreement patterns between politicians, we align their timelines based on days where they have tweeted about an issue. When two or more politicians tweet about the same issue on the same day, they are considered to have similar temporal activity, which may indicate stance agreement. This information is used as input to the predicate $\textsc{SameTemporalActivity}_I(\mathrm{P1},\ \mathrm{P2})$.

3.2.5 Political Framing

Framing is a political strategy that describes the concept of how politicians word their statements in order to control the way the public views and discusses current issues. To investigate the intuition that the way politicians contextualize their tweets is strongly indicative of their stance on an issue, we compiled a list of unique keywords for each political framing dimension as described in Boydstun et al. (2014) and Card et al. (2015). We again use the keyword matching approach described in Section 3.2.1 to classify all tweets with a political frame. As noted in Card et al. (2015), some tweets may fall into multiple frames. After all tweets are classified, we sum over the total number of each frame type and use the frames with the maximum and second largest counts as that politician's frames for that issue. The top two frames are used because for most politicians a majority of their issue-related tweets will fall into two frames. In the event of a tie we assign the frame that appears most frequently within that politician's party. These frames are used as input to the PSL predicate FRAME(P1, ISSUE).

3.2.6 Temporal Framing Patterns

While we expect politicians within a party to use similar frames per issue (as captured by the PSL predicate FRAME), it is also possible that politicians will use certain frames around events. For example, when a mass shooting occurs, we observe that Democrats will tweet about enacting gun legislation and typically frame these tweets as a matter of a needed preemptive action for public safety (which falls under the *Health and Safety* frame). In reaction to this, Republicans will tweet about protecting American citizens' rights to gun ownership, which falls under the *Constitutionality* frame. Therefore, we expect similarities and differences in framing usage around events to indicate similarities and differences in stances and agreement patterns. To capture this idea, we combine the approaches of Sections 3.2.4 and 3.2.5: we align the politicians' timelines per issue and compare the frames used to discuss the issue-related events. When two or more politicians use the same frame for an issue on the same day, we consider them to have similar temporal framing patterns. This is used as input to the PSL predicate $\textsc{SameTemporalFrame}_I$(P1, P2).

3.3 Models

Information obtained from the local models alone is not strong enough to quantify stance or agreement for politicians, as shown by our baseline measurements in Section 4.4. Therefore, we use PSL to build global connections among the output of the local models (which acts as weak supervision), resulting in global PSL models which successively build upon the previous model in order to obtain the highest accuracy possible. In addition to the PSL predicates representing the target output (Pro and $\textsc{SameStance}_I$) [3] and local models (as defined in Section 3.2), we also use directly observed information: party affiliation, denoted Dem(P1) for Democrat and ¬Dem(P1) for Republican, and SameParty(P1, P2) to denote if two politicians belong to the same political party.

3.3.1 Baseline: Using Local Classifiers Directly

To show that the local models do not provide enough information individually to make an accurate prediction, we implement a local baseline (LB) PSL model which does not take advantage of the global modeling framework. It instead learns weights over rules (shown in Table 3.3), which directly map the output of the local noisy classifiers described in Section 3.2 to PSL target predicates.

[3] In a supervised setting, jointly modeling the two target predicates can improve performance. Experiments using this approach yielded improvement in performance *and* a more complex model containing more parameters, resulting in slower inference.

Table 3.3.
Subset of PSL Rules Used in the Local Baseline Model.

PSL Rules: Local Baseline Model (LB)
LocalSameStance$_I$(P1, P2) $\rightarrow$ SameStance$_I$(P1, P2)
$\neg$LocalSameStance$_I$(P1, P2) $\rightarrow$ $\neg$SameStance$_I$(P1, P2)
Tweets(P1,Issue) $\wedge$ TweetPos(P1,Issue) $\rightarrow$ Pro(P1, Issue)
Tweets(P1,Issue) $\wedge$ TweetNeg(P1,Issue) $\rightarrow$ $\neg$Pro(P1, Issue)

3.3.2 Model 1: Agreement with Party Lines

The observation that politicians tend to vote with their political party on most issues is the basis of our initial assumptions in Model 1. The PSL rules listed in Table 3.4 are designed to capture this party based agreement. For some issues we initially assume Democrats (Dem) are for an issue, while Republicans ($\neg$Dem) are against that issue, (e.g., $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P1, Issue)), or vice versa. In the latter case, the rules of the model would change accordingly, e.g., the second rule would become $\neg$Dem(P1) $\rightarrow$ Pro(P1, Issue), and likewise for all other rules. Similarly, if two politicians are in the same party, we expect them to have the SameStance, or agree, on an issue. Though this is a strong initial assumption, the model can incorporate other indicators to overcome this bias when necessary. For all PSL rules, the reverse also holds, e.g., if two politicians are not in the same party, we expect them to have different stances.

3.3.3 Model 2: Politicians' Twitter Activity

Model 2 builds upon the initial party line bias of Model 1. In addition to political party based information, we also include representations of the politician's Twitter activity, as shown in Table 3.5. This includes whether or not a politician tweets about an issue (Tweets) as well as the sentiment of the tweets as determined

Table 3.4.
Subset of PSL Rules Used in Model 1.

PSL Rules: Model 1 (M1)
SameParty(P1, P2) $\rightarrow$ SameStance$_I$(P1, P2)
Dem(P1) $\rightarrow$ Pro(P1, Issue)
$\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P1, Issue)
SameParty(P1, P2) $\wedge$ Dem(P1) $\rightarrow$ Pro(P2, Issue)
SameParty(P1, P2) $\wedge$ $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P2, Issue)
SameParty(P1,P2) $\wedge$ Pro(P1, Issue) $\wedge$ Dem(P1) $\rightarrow$ Pro(P2, Issue)
SameParty(P1, P2) $\wedge$ $\neg$Pro(P1, Issue) $\wedge$ $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P2, Issue)

in Section 3.2.2. The predicate TweetPos models if a politician tweets positively on the issue, whereas TweetNeg models negative sentiment. Two sentiment predicates are used instead of the negation of TweetPos, which would cause all politicians for which there are no tweets, and hence no sentiment, on that issue to also be considered.

Table 3.5.
Subset of PSL Rules Used in Model 2. SameStance$_I$ is also abbreviated as SS$_I$.

PSL Rules: Model 2 (M2)
Tweets(P1, Issue) $\wedge$ Dem(P1) $\rightarrow$ Pro(P1, Issue)
Tweets(P1, Issue) $\wedge$ $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P1, Issue)
Tweets(P1, Issue) $\wedge$ Tweets(P2, Issue) $\wedge$ SameParty(P1, P2) $\rightarrow$ SS$_I$(P1, P2)
Tweets(P1, Issue) $\wedge$ Tweets(P2, Issue) $\wedge$ Dem(P1) $\rightarrow$ Pro(P2, Issue)
Tweets(P1, Issue) $\wedge$ Tweets(P2, Issue) $\wedge$ $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P2, Issue)
TweetPos(P1, Issue) $\wedge$ TweetPos(P2, Issue) $\rightarrow$ SameStance$_I$(P1, P2)
TweetPos(P1, Issue) $\wedge$ TweetNeg(P2, Issue) $\rightarrow$ $\neg$SameStance$_I$(P1, P2)
TweetPos(P1, Issue) $\wedge$ TweetPos(P2, Issue) $\wedge$ Dem(P1) $\rightarrow$ Pro(P2, Issue)
TweetNeg(P1, Issue) $\wedge$ TweetNeg(P2, Issue) $\wedge$ $\neg$Dem(P1) $\rightarrow$ $\neg$Pro(P2, Issue)
TweetPos(P1, Issue) $\wedge$ TweetPos(P2, Issue) $\wedge$ SameParty(P1, P2) $\rightarrow$ SS$_I$(P1, P2)
TweetPos(P1, Issue) $\wedge$ TweetNeg(P2, Issue) $\wedge$ $\neg$SameParty(P1, P2) $\rightarrow$ $\neg$SS$_I$(P1, P2)

3.3.4 Model 3: Politicians' Agreement Patterns

Table 3.6 presents a subset of the rules used in Model 3 to incorporate higher level Twitter information into the model. The incorporation of this information allows Model 3 to overcome Model 2 inconsistencies between stance and sentiment (e.g., when someone is attacking their opposition). Our intuition that politicians who have similar tweets would also have similar stances on issues is represented with the predicate LocalSameStance_I. SameTemporalActivity represents the idea that if politicians tweet on an issue around the same time range then they also share a stance for that issue. Frame indicates the frame used by that politician for different issues. Finally, $\text{SameTemporalFrame}_I$ conveys that two politicians use the same frames for an issue at the same time. More details on these predicates are in Sections 3.2.3, 3.2.4, 3.2.5, and 3.2.6 respectively.

Table 3.6.
Subset of PSL Rules Used in Model 3. SameStance_I is also abbreviated as SS_I.

PSL Rules: Model 3 (M3)
$\text{LocalSameStance}_I(\text{P1, P2}) \wedge \text{Pro}(\text{P1, Issue}) \rightarrow \text{Pro}(\text{P2, Issue})$
$\text{SameTemporalActivity}_I(\text{P1, P2}) \wedge \text{SameParty}(\text{P1, P2}) \rightarrow \text{SameStance}_I(\text{P1, P2})$
$\text{SameParty}_I(\text{P1, P2}) \wedge \text{Frame}(\text{P1, Issue}) \wedge \text{Frame}(\text{P2, Issue}) \rightarrow \text{SS}_I(\text{P1, P2})$
$\text{Frame}(\text{P1, Issue}) \wedge \text{Frame}(\text{P2, Issue}) \rightarrow \text{SameStance}_I(\text{P1, P2})$
$\text{Frame}(\text{P1, Issue}) \wedge \text{Frame}(\text{P2, Issue}) \wedge \text{SameParty}(\text{P1, P2}) \rightarrow \text{SS}_I(\text{P1, P2})$
$\text{Frame}(\text{P1, Issue}) \wedge \text{Dem}(\text{P1}) \rightarrow \text{Pro}(\text{P1, Issue})$
$\text{Frame}(\text{P1, Issue}) \wedge \neg\text{Dem}(\text{P1}) \rightarrow \neg\text{Pro}(\text{P1, Issue})$
$\text{SameTemporalFrame}_I(\text{P1, P2}) \wedge \text{SameParty}(\text{P1, P2}) \rightarrow \text{SameStance}_I(\text{P1, P2})$
$\text{SameTemporalFrame}_I(\text{P1, P2}) \wedge \text{Pro}(\text{P1, Issue}) \rightarrow \text{Pro}(\text{P2, Issue})$

Table 3.7.
Stance and Agreement Accuracy by Issue. The SVM columns show the results of the tweet-based, supervised baseline. LB columns show the results when using only the weak local models. M1 columns are the results based on party line agreement, M2 columns are the results when adding Twitter activity to M1, and M3 columns are the results when adding higher level Twitter behaviors to M1 and M2.

Issue	Stance (Results of Pro Prediction)					Agreement (SameStance Prediction)				
	SVM	LB	M 1	M 2	M 3	SVM	LB	M 1	M 2	M 3
Abortion	61.25	81.25	96.88	96.88	96.88	44.34	49.31	93.75	93.75	95.36
ACA	87.5	96.88	100	100	100	79.7	51.61	100	100	100
Confederate	16.56	34.38	78.12	84.38	87.5	0	51.31	69.6	77.7	80.18
Drugs	48.13	87.5	78.12	88.88	96.88	44.34	50.42	63.6	84.07	84.07
Environment	69.06	53.12	78.12	78.13	81.08	65.86	45.16	65.59	68.75	71.37
Guns	84.38	93.75	93.75	93.75	93.75	57.33	48.59	68.54	99.5	99.59
Immigration	73.44	37.5	81.25	81.25	86.36	51.82	53.62	68.55	69.06	69.56
Iran	74.56	84.38	65.62	65.63	84.38	69.25	35.57	79.73	100	100
ISIS	80.0	40.32	76.28	93.75	94.44	74.19	59.68	76.28	76.28	90.04
Marriage	33.44	62.5	90.62	90.62	90.9	12.5	50.57	87.12	87.13	87.43
NSA	21.25	37.5	53.12	53.12	61.54	2.61	34.15	49.2	56.66	60.08
Pay	34.38	84.38	84.38	89.47	90.62	29.59	64.30	72.92	74.31	80.31
Religion	42.81	75	68.75	81.25	81.25	56.89	47.62	76.24	76.46	79.44
Social Security	35.31	28.12	78.12	78.13	78.13	0.91	53.76	73.25	90.03	90.88
Student	0	93.75	96.88	96.88	96.88	0	51.61	100	100	100
TPP	0	62.5	62.5	62.5	62.5	0	45.43	48.39	54.64	65.32

3.4 Experimental Results

3.4.1 Experimental Settings

Supervised Baseline. Previous works exploring stance classification typically predict stance based on an *individual item of text* (e.g., forum post or single tweet) in a *supervised* setting, making it difficult to directly compare to our approach. To facilitate comparison, we implemented a tweet-based supervised baseline, per issue. We labeled each tweet with the politician's stance (either for or against) on that tweet's issue. We trained an SVM on 80% of the politicians' tweets and tested on the remaining 20%, using 5-fold cross-validation. Because we aim to predict each politician's stance and *not* the stance of each tweet, we aggregated the SVM predictions by politician, i.e., the SVM predicts a stance for each tweet and the majority prediction among a politician's tweets is used as his or her stance. For agreement prediction, we compared this stance prediction across politicians to determine if the predicted stances agreed and whether or not this agreement was correct.

PSL Models. As described in Section 3.2, the data generated from the local models is used as weak supervision to initialize the PSL models. The Local Baseline model (LB) is initialized with only the information from the weak local models. We initialize Model 1 (M1), as described in Section 3.3.2, using knowledge of the politician's party affiliation. Model 2 (M2) builds upon (M1) by incorporating the results of the issue and sentiment analysis local models, as described in Sections 3.2.1 and 3.2.2 respectively. Model 3 (M3) combines all previous models with higher level knowledge of Twitter activity: tweet agreement (Section 3.2.3), temporal activity (Section 3.2.4), frames (Section 3.2.5), and temporal framing patterns (Section 3.2.6). We implement our PSL models to have an initial bias that candidates do not share a stance and are against an issue. Stances collected in Section 3.1.1 are used as the ground truth for evaluation of the predictions of the PSL models only, not for any form of supervision.

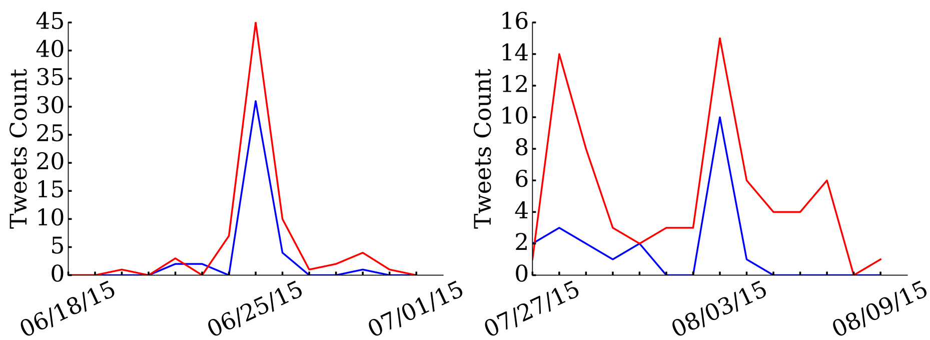
3.4.2 Quantitative Results

Results Per Issue. Table 3.7 presents the results of using the supervised baseline and our three proposed PSL models. While the supervised baseline results (SVM) are not directly comparable to our weakly supervised model, since the supervised model uses a different data split and approach, it does show that direct supervision does not lead to immediate prediction improvement and can result in weaker prediction scores. LB refers to using only the weak local models for prediction with no additional information about party affiliation. We observe that for prediction of stance (PRO) LB performs better than random chance in 11 of 16 issues; for prediction of agreement ($\textsc{SameStance}_I$), LB performs slightly lower overall, with only 9 of 16 issues predicted above chance. Using M1, we improve stance prediction accuracy for 10 of the issues and agreement accuracy for all issues. M2 further improves the stance and agreement predictions for an additional 8 and 12 issues, respectively. M3, the combination of the previous models with Twitter behavioral features, increases the stance prediction accuracy of M2 for 9 issues and the agreement accuracy for 12 issues.

The final agreement predictions of M3 are notably improved over the initial LB for all issues, indicating that similarities and differences in Twitter behaviors help capture agreement and disagreement patterns among politicians. The final stance predictions of M3 are improved over all issues except Guns, Iran, and TPP. For Guns, the stance prediction remains the same throughout all models, meaning party information does not boost the initial predictions determined from Twitter based behaviors. For Iran, the addition of M1 and M2 lower the accuracy, but the temporal features from M3 are able to restore it to the original prediction. For TPP, this trend is likely due to the fact that all models incorporate party information and the issue of TPP is the most heavily divided within and across parties, with 8 Republicans and 4 Democrats in support of TPP and 8 Republicans and 12 Democrats opposed. Even in cases where M1 and/or M2 remained steady or lowered the initial baseline result (e.g., stance for Religion and Pay), the final prediction by M3 is still higher than that of the baseline.

Table 3.8.
Overall Accuracy for Stance (St) and Agreement (Ag) Prediction. Global represents the accuracy over all politicians, while Rep and Dem refer to Republicans or Democrats only.

	Global		Rep		Dem	
	St	Ag	St	Ag	St	Ag
LB	68.36	52.49	66.80	49.10	69.92	44.86
M1	81.25	76.34	75.39	75.16	87.11	85.44
M2	85.16	87.30	81.25	84.26	89.06	91.37
M3	89.84	87.76	87.11	85.35	92.58	91.49



(a) ACA (b) Abortion

Figure 3.3. Temporal Twitter Activity by Party. The red and blue lines represent the temporal overlaps, or lack thereof, of Republicans and Democrats (respectively) in Twitter activity one week before and after a major event.

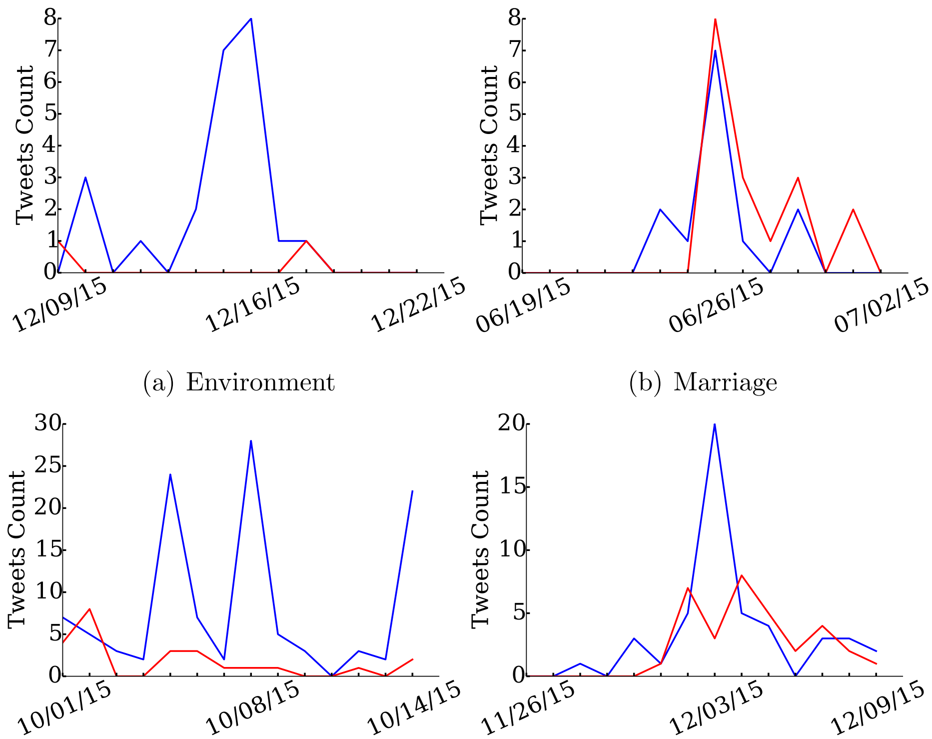
Overall Results. Table 3.8 presents our overall results for stance and agreement prediction in terms of accuracy. The Global score is the overall average for all politicians, while Rep and Dem consider only Republicans or Democrats, respectively. Each model increases the accuracy of the previous model's prediction, showing that additional Twitter behavioral features can help overcome the strong party line biases captured by M1.

3.4.3 Effects of Framing and Temporal Activity Patterns

As shown in Table 3.7, performance for *some* issues does not improve in M3. Upon investigation, we found that for all issues, except Abortion which improves in agreement, one or both of the top frames for the party are shared across party lines. For example, for ACA both Republicans and Democrats have the *Economic* and *Health and Safety* frames as their top two frames. For TPP, both parties share the *Economic* frame. In addition to similar framing overlap, the Twitter timeline for ACA also exhibits overlap, as shown in Figure 3(a). This figure highlights one week before and after the Supreme Court ruling to uphold the ACA. The peak of Twitter activity is the day of the ruling, 6/25/2015.

Conversely, Abortion, which shares no frames between parties (Democrats frame Abortion with *Constitutionality* and *Health and Safety* frames; Republicans use *Economic* and *Capacity and Resources* frames), exhibits a timeline with greater fluctuation. The peak of Figure 3(b) is 8/3/2015, which is the day that the budget was passed to include funding for Planned Parenthood. Despite sharing a peak, both parties have different patterns over this time frame, allowing M3 to extract enough information to increase agreement prediction accuracy by 1.61%.

Figure 3.4(a) shows an example of one event for the Environment issue: when the mayor of Flint, Michigan declared a state of emergency over lead in the city's water supply. Due to different temporal patterns and frames for such events, the Environment accuracy improves across all models, as shown in Table 3.7. Similarly, Figure 3.4(b) shows the week before and after the Supreme Court ruled to uphold the legality of same-sex marriage. The two central peaks are shared by both parties, but each party also has one peak before (Democrats) or after (Republicans) the event. Additionally, both parties share the *Constitutionality* frame as their top frame, but the second most used frame is *Morality* for Republicans and *Fairness and Equality* for Democrats. These slight differences allow the M3 model to improve over the M2 prediction. Finally, Figure 3.4(c) shows the week before and after Democratic



(a) Environment (b) Marriage

(c) Guns (d) Guns

Figure 3.4. Temporal Twitter Activity by Party for Three Issues.

Senators pushed for gun control legislation after the Umpqua Community College shooting and Figure 3.4(d) shows tweets around the Inland Regional Center in San Bernadino shooting. For these events, both parties exhibit different timeline patterns and frames. Consequently, these behavioral features dominate the stance prediction and allow agreement accuracy to reach 99.59%.

3.5 Chapter Summary

In this chapter, we present our framework for modeling the dynamic nature of political discourse on Twitter. Initially, we focus on a small set of politicians and issues to study the benefits of weakly supervised modeling and incorporation of behavioral features. Contrary to previous works, which typically focus on a single aspect of this complex microblogging behavior, we build a holistic model connecting party line biases, temporal behaviors, and issue framing into a single predictive model which identifies fine-grained stances and agreement patterns. Despite having no direct supervision and using only intuitive local classifiers to bootstrap our global model, our approach results in a strong predictive model which helps shed light on political discourse within and across party lines.

4 POLITICAL DISCOURSE FRAME PREDICTION

In this chapter, we focus on political framing, a very nuanced political discourse analysis task, on a variety of issues frequently discussed on Twitter. Framing (Entman, 1993; Chong and Druckman, 2007) is employed by politicians to bias the discussion towards their stance by emphasizing specific aspects of the issue. For example, the debate around increasing the minimum wage can be framed as a *quality of life* issue or as an *economic* issue. While the first frame supports increasing minimum wage because it improves workers' lives, the second frame, by conversely emphasizing the costs involved, opposes the increase. Using framing to analyze political discourse has gathered significant interest over the last few years (Tsur et al., 2015; Card et al., 2015; Baumer et al., 2015) as a way to automatically analyze political discourse in congressional speeches and political news articles. Different from previous works which focus on these longer texts or single issues, our dataset includes tweets authored by all members of the U.S. Congress from both parties, dealing with several policy issues (e.g., immigration, ACA, etc.). These tweets were annotated by adapting the annotation guidelines developed by Boydstun et al. (2014) for Twitter.

Twitter issue framing is a challenging multilabel prediction task. Each tweet can be labeled as using one or more frames, out of seventeen possibilities, while only providing 280 characters as input to the classifier. The main contribution of this chapter is to evaluate whether the *social and behavioral information* available on Twitter is sufficient for constructing a reliable classifier for this task. We approach this framing prediction task using a weakly supervised collective classification approach which leverages the dependencies between tweet frame predictions based on the interactions between their authors.

These dependencies are modeled by connecting Twitter users who have social connections or behavioral similarities. Social connections are directed dependencies that represent the followers of each user as well as retweeting behavior (i.e., user A retweets user B's content). Interestingly, such social connections capture the flow of influence within political parties; however, the number of connections that cross party lines is extremely low. Instead, we rely on capturing behavioral similarity between users to provide this information. For example, users whose Twitter activity peaks at similar times tend to discuss issues in similar ways, providing indicators of their frame usage for those issues. In addition to using social and behavioral information, our approach also incorporates each politician's party affiliation and the frequent phrases (e.g., bigrams and trigrams) used by politicians on Twitter.

These lexical, social, and behavioral features are extracted from tweets via weakly supervised models and then declaratively compiled into a graphical model using Probabilistic Soft Logic (PSL), the probabilistic modeling framework introduced in Section 2.2. As described in Section 4.3, PSL specifies high level rules over a relational representation of these features. These rules are then compiled into a graphical model called a hinge-loss Markov random field (Bach et al., 2013), which is used to make the frame prediction. Instead of direct supervision we take a bootstrapping approach by providing a small seed set of keywords adapted from Boydstun et al. (2014), for each frame.

Our experiments show that modeling social and behavioral connections improves F_1 prediction scores in both supervised and unsupervised settings, with double the increase in the latter. We apply our unsupervised model to our entire dataset of tweets to analyze framing patterns over time by both party and individual politicians. Our analysis provides insight into the usage of framing for identification of *aisle-crossing politicians*, i.e., those politicians who vote against their party.

Table 4.1.
Frames and Descriptions. The first 14 frames are taken from Boydstun et al. (2014) and the last 3 are our proposed Twitter-specific frames. Boydstun's original Frame 15 (*Other*) is omitted from this chapter.

Frame Number, Frame Name, and Brief Description of Frame
1. Economic: *Pertains to the economic impacts of a policy*
2. Capacity & Resources: *Pertains to lack of or availability of resources*
3. Morality & Ethics: *Motivated by religious doctrine, righteousness, sense of responsibility*
4. Fairness & Equality: *Of how laws, punishments, resources, etc. are distributed among groups*
5. Legality, Constitutionality, & Jurisdiction: *Including court cases, restriction and expressions of rights*
6. Crime & Punishment: *Policy violation and consequences*
7. Security & Defense: *Threats or defenses/preemptive actions to protect against threats*
8. Health & Safety: *Includes care access and effectiveness*
9. Quality of Life: *Effects on individual and community life*
10. Cultural Identity: *Norms, trends, and customs of cultures*
11. Public Sentiment: *Pertains to opinions, polling, and demographics*
12. Political Factors & Implications: *Efforts, stances, filibusters, lobbying, references to other politicians*
13. Policy Description, Prescription, & Evaluation: *Discusses effectiveness of current or proposed policies*
14. External Regulation and Reputation: *Interstate and international relationships of the U.S.*
15. Factual: *Expresses a pure fact, with no detectable political spin*
16. (Self) Promotion: *Promotes another person or the author in some way, e.g., television appearances*
17. Personal Sympathy & Support: *Expresses sympathy, emotional response, or solidarity with others*

(1) Twenty million and counting... #ACAWorks [13]

(2) Six years later health care costs have skyrocketed [1] and millions have lost access to their doctors. [8] #RepealObamacare

(3) Too many families feel the devastation of gun violence [9]. The American people deserve action from Congress [12] #WearOrange

(4) This is a great day for Americans who believe in individual rights [5] and women's health care (Frame: 4). #StopTheSham #SCOTUS

(5) We must bolster the security of our borders [7] and craft an #immigration policy that grows our economy. [1]

(6) #500Days ago Boko Haram kidnapped 276 schoolgirls, 219 are still missing. #BringBackOurGirls [15]

(7) I look forward to talking w/Bill Edwards on @1290WTKS at 8:44AM. We'll talk budget, Syria, ISIL, and @GaPorts. Tune in: [16]

(8)Senseless violence has no place in the world and especially not at church. My prayers are with Emanuel AME and everyone in Charleston today.[17]

Figure 4.1. Tweets Highlighting Frame Classification Difficulty. The superscript number after each tweet or color section indicates the frame. Different colors highlight phrases that indicate possible framing dimensions. No highlight indicates that the entire tweet falls under one frame.

4.1 Political Frames

4.1.1 The Policy Codebook Frames

In this section, we describe in greater detail the fourteen frames adapted from the Policy Frames Codebook of Boydstun et al. (2014), shown in Table 4.1. The Codebook provides details of how each frame can be used as well as a variety of examples demonstrating how to classify texts as having one frame over another. The codebook was designed to guide the annotation of newspaper articles that discuss policy issues. We highlight the main aspects of each frame here, as well as examples of how this frame typically appears in our tweets dataset or is interpreted in our annotation process.

Frame 1: Economic. A text is expressed with an *Economic* frame if it discusses an issue in terms of the monetary or financial effects of the issue on an individual or group of people, such as a family, community, or the entire economy. This frame is also used to describe an issue in terms of its economic impacts on various areas of the economy including job wages, trade, employment, and taxes. For example, a tweet discussing the abortion issue in terms of whether or not the cost of the procedure should be covered by health insurance providers uses this frame. Specific examples of this frame in use are shown in tweets (2) and (5) in Figure 4.1.

Frame 2: Capacity & Resources. This frame is used for an issue that is discussed in terms of the availability, or lack thereof, of physical or financial resources. For example, a tweet that discusses an issue in terms of the strain on existing infrastructure would fall under this frame. It is easy to confuse this frame with the *Economic* frame and the two often overlap. An easy way to determine the difference between the two is illustrated in the following example. If a tweet discusses the cost of a policy, then it is framed with the *Economic* frame. However, if it discusses the lack of sufficient resources to fund some part of the policy, then it is framed with the *Capacity & Resources* frame. This frame does not occur frequently in our dataset.

Frame 3: Morality. If a tweet discusses an issue by emphasizing its moral impacts, such as religious or ethical aspects of the issue, or a sense of personal responsibility or duty about an issue, then that tweet is expressed with the *Morality* frame. This frame is often associated with religious-based morality, for example, a tweet that states people should follow religious laws in order to be good people. However, it can also appear in a non-religious context, such as an argument about how supporting programs such as Medicare is "the right thing" for society as a whole.

Frame 4: Fairness & Equality. This frame is used to emphasize how policies are distributed among individuals and groups. It is typically used to emphasize the effects of the policy on minority groups, such as different races or genders. For example, if a tweet discusses “women’s health care” as opposed to “health care”, then it would be classified under this frame. Because tweets are limited to 280 characters, a politician’s choice to include this extra word indicates their bias towards one frame over another.

Frame 5: Legality, Constitutionality, & Jurisdiction. This frame is used to express the legal or constitutional impacts of an issue. When a tweet references laws or court cases that are either current or upcoming, it would fall under this frame. Examples of how this frame typically appears in our dataset include discussions of whether or not President Obama exceeded the power of his office, upcoming Supreme Court cases, and discussions of gun ownership in terms of personal rights. An example of this frame is shown in tweet (4) in Figure 4.1.

Frame 6: Crime & Punishment. The *Crime & Punishment* frame is another frame that does not appear frequently in our data set. It is used to discuss issues in terms of infractions and punishment for infractions of policies. This frame is typically used to frame acts of violence as murder or discuss how such actions should be punished.

Frame 7: Security & Defense. Issues that focus on a threat to people or the country as a whole or how to handle the threat fall under this frame. The key difference between this frame and Frame 8 is that this frame includes *preemptive measures* taken to defend against a threat, for example, actions taken to prevent gun violence. Examples of this frame in our dataset include discussions of stopping immigration by building a wall (e.g., tweet (5) in Figure 4.1) or deploying troops to combat ISIS.

Frame 8: Health & Safety. This frame appears in tweets that emphasize health care issues such as access, diseases, mental health, or hospitals, as well as safety issues. In contrast to Frame 7, handling matters of safety *after* an event (such as a shooting or widespread pandemic) would fall under this frame. This frame often appears in our dataset in tweets that discuss health insurance access (e.g., tweet (2) in Figure 4.1) and effectiveness when discussing the ACA.

Frame 9: Quality of Life. This frame is used when a tweet discusses how an issue affects the quality of life of people, specifically effects on happiness and community life. For example, in the context of the immigration debate, this frame is used to discuss how immigration reforms will impact immigrants' family lives. Another frequent occurrence is the use of this frame to describe the negative effects of gun violence on families or communities, as shown in tweet (3) in Figure 4.1.

Frame 10: Cultural Identity & Pop Culture. When an issue is discussed using social norms, cultural values, or stereotypes, it falls under this frame. This can also be used to reference known values of political or religious groups, people, or famous politicians. When referring to famous politicians, the tweet is also often labeled as having Frame 12. This frame frequently appears in our dataset when politicians emphasize American values when discussing an issue.

Frame 11: Public Sentiment. When a tweet refers to public opinion or is used to describe the results of polls and demographics on an issue, it is classified as having this frame. It can also include references to a politician's own party or supporters, in which case the tweet would also be expressed with Frame 12. Examples in our dataset include tweets that quote poll results or reference what actions Americans want politicians to take.

Frame 12: Political. Tweets that discuss the politics of an issue, such as filibusters, lobbying, bipartisan movements, and political strategies are considered to be framed with the *Political* frame. When political parties, such as @HouseDems or @HouseGOP, or specific politicians such as President Obama are mentioned, then the tweet also expresses the issue in terms of its political impacts. Tweet (3) in Figure 4.1 shows an example of this frame in use.

This frame frequently appears in conjunction with other frames in our dataset. This is due to the inherent social network structure of Twitter, which is not present in newspaper articles. The Codebook was designed for annotating the latter, which are also much longer than tweets and therefore often discuss policies in more depth. For newspaper articles or speeches, Frame 13 is more likely to appear with other frames.

Frame 13: Policy Prescription & Evaluation. When a tweet describes the details of an issue beyond its basic components, such as pros and cons of the issue or ways to improve its effects, then the tweet is expressed with this frame. For example, this frame applies if a tweet discusses exactly how the ACA has affected people instead of just mentioning the ACA. Another example of this would be discussing different lengths of waiting periods to buy guns to determine which time frame is best. Tweet (1) in Figure 4.1 shows an example of this frame used as a hashtag.

Frame 14: External Regulation & Reputation. This frame is used to focus on the reputation of the United States and its relationships with other nations or relationships between states within the country. For example, if a tweet refers to opinions about the U.S. military in Middle Eastern countries then it would be expressed with this frame.

4.1.2 Proposed Twitter-specific Frames.

In Boydstun et al. (2014), Frame 15 is the *Other* frame and is designed to capture articles that cannot be classified under the first fourteen frames. When labeling our dataset we noticed that tweets labeled as Frame 15 could be divided into three types of frames. Therefore we dropped the *Other* frame from our analysis and proposed the following three frames as *Twitter-specific* frames because they may not be applicable in traditional media settings.

Frame 15: Factual. This frame is used when tweets discuss an issue with no detectable political twist or spin. They are not verified as facts. These tweets typically discuss numerical information about a policy, such as prices or percentages. For example, the number of Americans who are insured because of the ACA frequently appears in tweets and when it does without mentioning anything else, we label it as the *Factual* frame. However, if such a tweet mentioned that the ACA cost taxpayers some amount of money, the emphasis would be on the cost and it would therefore fall under the *Economic* frame. Tweet (1) in Figure 4.1 is an example of this scenario. The first part of the tweet ("Twenty million") alone would indicate the *Factual* frame. However, by adding "and counting... #ACAworks", the author of the tweet indicates the growth and outreach of ACA, resulting in the *Policy* frame (Frame 13) being chosen.

Frame 16: (Self) Promotion. Politicians typically use Twitter to promote their own or their political allies' political actions and public appearances, either on TV or the radio, on their social media accounts. Sometimes these tweets also mention the issues that they will be discussing during their appearance. When this happens, the tweet may have secondary frames. Tweet (7) in Figure 4.1 shows an example of this frame in use.

Frame 17: Personal Sympathy & Support. This frame is used to express personal emotions or sentiments on an issue. It often appears in tweets following gun violence events, in which politicians state their "thoughts and prayers" are with the victims. It is also used to express solidarity, in which the politician states that they "stand with" a group in support of their movement. An example of this frame is shown in tweet (8) in Figure 4.1.

4.2 Data Collection, Preprocessing, and Annotation

Data Collection and Preprocessing. We collected 184,914 of the most recent tweets of members of the 114$^{\text{th}}$ U.S. House of Representatives and Senate (collectively referred to as the U.S. Congress). Using an average of ten keywords per issue, we filtered out tweets not related to the following six issues of interest: (1) limiting or gaining access to abortion, (2) debates concerning the Affordable Care Act (i.e., ACA or Obamacare), (3) the issue of gun rights versus gun control, (4) effects of immigration policies, (5) acts of terrorism, and (6) issues concerning the LGBTQ community. Table 4.2 lists the keywords or phrases used to filter the entire dataset to only those tweets related to the six issues studied in this paper. In order to reduce noise, these tweets are further processed by removing the following attributes: capitalization, stop words, URLs, and punctuation.

Forty politicians (ten Republican Senators, ten Republican Representatives, ten Democratic Senators, and ten Democratic Representatives), were chosen randomly for annotation. Table 4.3 presents the overall distribution of our Congressional Tweets Dataset, which is available for use by the NLP community.[1]

Data Annotation. Two graduate students were trained in the use of the Policy Frames Codebook developed by Boydstun et al. (2014) for annotating each tweet with a frame. The general aspects of each frame are listed in Table 4.1 and examples of corresponding tweets are shown in Figure 4.1. Frames are designed to generalize across issues and overlap of multiple frames is possible. Additionally, the Codebook is typically applied to congressional speeches or newspaper articles where discussion of policy (Frame 13) can encompass other frames within the text. Consequently, annotators using the Codebook are advised to be careful when assigning Frame 13 to a text. For similar reasons, as well as the inherent social network structure of Twitter,

[1]The dataset and PSL scripts are available at: `http://purduenlp.cs.purdue.edu/projects/twitterframing`.

Table 4.2.
Keywords or Phrases Used to Filter Tweets for Issue.

ISSUE	KEYWORDS OR PHRASES
ABORTION	*abortion, pro-life, pro-choice, Planned Parenthood, StandWithPP, Hobby Lobby, birth control, women's choice, women's rights, women's health*
ACA	*patient protection, affordable care act, ACA, Obamacare, health care, healthcare, Burwell, Medicare, Medicaid, repeal and replace*
GUNS	*Charleston, gun, shooting, Emanuel, Second Amendment, Oregon, San Bernadino, gun violence, gun control, 2A, NRA, Orlando, Pulse*
IMMIGRATION	*immigration, immigrants, illegal immigrants, border, amnesty, wall, Dreamers, Dream Act*
LGBTQ	*equality, marriage, gay, transgender, marriage equality, same-sex, gay marriage, religious freedom, RFRA, bathroom bill*
TERRORISM	*terrorism, terrorists, terror network, ISIS, ISIL, Al Qaeda, Boko Haram, extremist*

Table 4.3.
Statistics of Collected Tweets. The abbreviations in the table correspond to the following: REP for Republican, DEM for Democrat, ABORT for Abortion, IMM for Immigration, and TER for Terrorism.

TWEETS	BY PARTY		BY ISSUE					
	REP	DEM	ABORT	ACA	GUNS	IMM	TER	LGBTQ
ALL	48504	43953	6467	35854	15532	13442	15205	6046
LABELED	894	1156	170	564	543	233	446	183

we further advise annotators to be cautious when labeling tweets with Frame 12 which covers situations such as politicians calling each other out by name, party, or group.

Table 4.4.
Count of Each Type of Frame Per Issue in Labeled Dataset. The abbreviations in the table correspond to the following: REP for Republican, DEM for Democrat, ABORT for Abortion, IMM for Immigration, and TER for Terrorism.

ISSUE	FRAMES																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ABORT	4	7	23	55	40	0	2	32	10	0	4	46	20	0	1	13	8
ACA	65	9	6	28	24	0	3	128	21	3	18	116	174	2	21	100	15
GUNS	2	2	37	16	30	21	93	8	36	14	49	166	65	0	5	55	147
IMM	16	7	6	6	42	3	15	0	29	19	7	81	52	1	1	32	2
LGBTQ	0	0	9	99	23	2	2	3	10	17	7	39	14	1	2	11	48
TER	6	4	46	3	11	10	115	1	6	13	14	69	68	35	6	99	57

Based on this guidance and the difficulty of labeling tweets (as shown in Figure 4.1 and as discussed in Card et al. (2015)), annotators were instructed to use the following procedure: (1) attempt to assign a primary frame to the tweet if possible, (2) if not possible, assign two frames to the tweet where the first frame is chosen as the more accurate of the two frames, (3) when assigning frames 12 through 17, ensure that the tweet cannot be assigned to any other frames. Annotators spent one month labeling the randomly chosen tweets. For all tweets with more than one frame, annotators met to come to a consensus on whether the tweet should have one frame or both. The labeled dataset has an inter-annotator agreement, calculated using Cohen's Kappa statistic, of 73.4%.

Figure 4.2 shows the coverage of the labeled frames used by each political party. From this, general patterns can be observed. For example, Republicans use Frames 12 (*Political Factors & Implications*) and 17 (*Personal Sympathy & Support*) more frequently than Democrats, while Democrats tend to use Frames 4 (*Fairness & Equality*), 9 (*Quality of Life*), 10 (*Cultural Identity*), and 11 (*Public Sentiment*) more often than Republicans. Lastly, Table 4.4 shows the number of each type of frame that appears in each issue in our labeled dataset. The total counts are not equally distributed

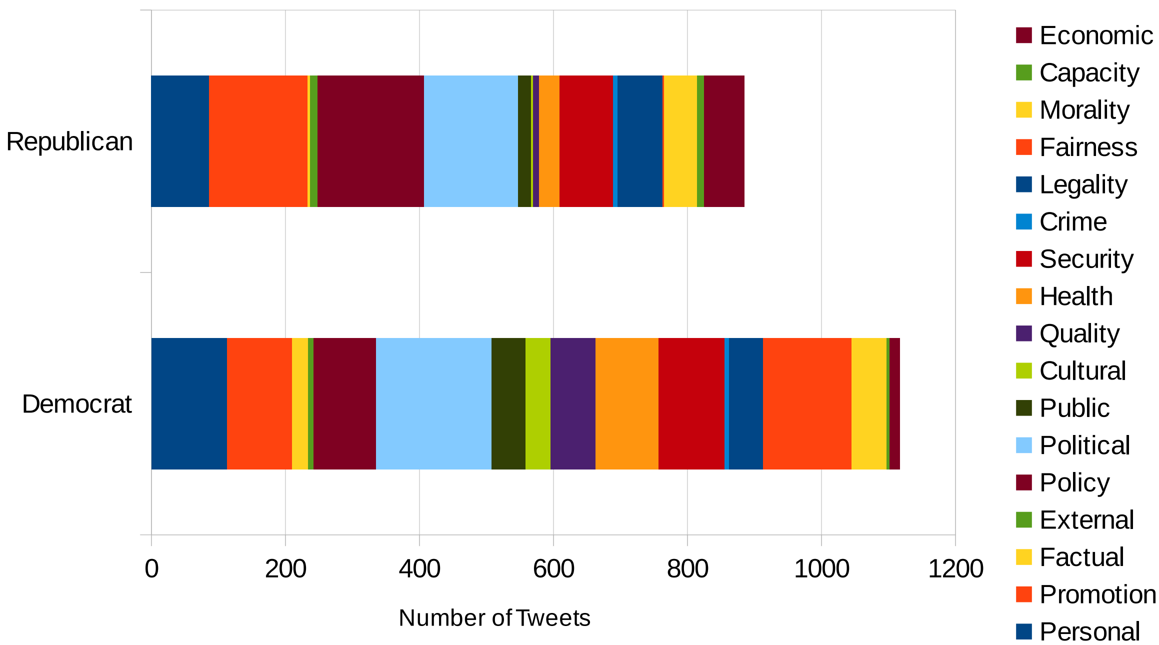


Figure 4.2. Coverage of Frames by Party.

across issues or frames for two reasons: (1) these counts accurately reflect politician's framing choices on Twitter because, for example, politicians infrequently discuss issues in terms of resources (Frame 2) and (2) we randomly chose which politicians to label without controlling for the issues they discussed. Due to the difficulty and time required to manually annotate the tweets, as well as our desire to analyze frames in as realistic a setting as possible, we conducted our experiments with the dataset as it is and did not try to find more instances of specific frames in order to balance the dataset.

Extensions of the Codebook for Twitter Use. The first fourteen frames outlined in Table 4.1 are directly applicable to the tweets of U.S. politicians. In our labeled set, Frame 15 (*Other*) was never used. Therefore, we drop its analysis from this paper. From our observations during annotation, we propose the addition of the three frames listed at the bottom of Table 4.1 specifically for Twitter analysis:

Factual (Frame 15), *(Self) Promotion* (Frame 16), and *Personal Sympathy & Support* (Frame 17). Tweets that present a fact, with no detectable political spin or twists, are labeled as having the *Factual* frame (15). Tweets that discuss a politician's appearances, speeches, statements, or refer to political friends are considered to have the *(Self) Promotion* frame. Finally, tweets where a politician offers their "thoughts and prayers", condolences, or stands in support of others, are considered to have the *Personal Sympathy & Support* frame. Section 4.1.2 described each of these new frames in more detail.

We find that for most tweets, one frame is not enough. This is caused by the compound nature of many tweets (e.g., tweets (3) and (5) in Figure 4.1). Some tweets are two separate sentences, with each sentence having a different frame. Other tweets begin with one frame and end with another (e.g., tweet (2)). A final problem, that may also be relevant to longer text articles, is that of subframes within a larger frame, as shown in tweet (5) of Figure 4.1. In this tweet, two frames are identifiable: Frame 7 (*Security & Defense*) is highlighted in yellow and Frame 1 (*Economic*) is highlighted in green. However, the tweet as a whole could fall under Frame 13 (*Policy*) if this tweet was a rebuttal point about an immigration policy. The lack of available context for short tweets can make it difficult to determine if a tweet should have one primary frame or is more accurately represented by multiple frames.

4.3 Global Models of Twitter Language and Activity

Due to the dynamic nature of political discourse on Twitter, our approach is designed to require as little supervision as possible. We implement a variety of weakly supervised classifiers which are defined over domain information. Some of these classifiers are based on directly observable information which is easy to infer, such as political party affiliation or the issue of the tweet. Others are designed to extract non-trivial linguistic information from tweets, for example, bigrams used by one party for a certain issue, or more complex behaviors such as retweet patterns. Thus, the only weak sources of supervision these models require include: unigrams related to the issues (taken from Table 4.2 in Section 4.2), unigrams adapted from the Boydstun et al. (2014) Codebook for frames (shown in Table 4.5), and political party of the author of the tweets. Once this information is extracted, it is then formatted into input for PSL predicates.

Table 4.5.: Frame and Corresponding Unigrams Used for Initial Supervision.

Frame Number & Name	Corresponding Unigrams
1. Economic	*premium(s), small, business(es), tax(es), economy, economic, cost(s), employment, market, spending, billion(s), million(s), company, companies, funding, regulation, benefit(s), health*
2. Capacity & Resources	*resource(s), housing, infrastructure, IRS, national, provide(s), providing, fund(s), funding, natural, enforcement*
3. Morality & Ethics	*moral, religion(s), religious, honor(able), responsible, responsibility, illegal, protect, god(s), sanctity, Islam, Muslim, Christian, radical, violence, victim(s), church*
4. Fairness & Equality	*fair(ness), equal(ity), inequality, law(s), right(s), race, gender, class, access, poor, civil, justice, social, women(s), LGBT, LGBTQ, discrimination, decision(s)*
5. Legality, Constitutionality, & Jurisdiction	*right(s), law(s), executive, ruling, constitution(al), amnesty, decision(s), reproductive, legal, legality, court, SCOTUS, immigration, amendment(s), judge, authority, precedent, legislation*

continued on next page

Table 4.5.: *Frame and Corresponding Unigrams, continued.*

Issue	Frame and Corresponding Unigrams
6. Crime & Punishment	*crime(s), criminal(s), gun(s), violate(s), enforce(s), enforced, enforcement, civil, tribunals, justice, victim(s), civilian(s), kill, murder, hate, genocide, consequences*
7. Security & Defense	*security, secure, defense, defend, threat(s), terror, terrorism, terrorist(s), gun(s), attack(s), wall, border, safe, safety, violent, violence, ISIS, ISIL, suspect(s), domestic, prevent, protect*
8. Health & Safety	*health(y), care, healthcare, Obamacare, access, disease(s), mental, physical, affordable, coverage, quality, (un)insured, disaster, relief, unsafe, cancer, abortion*
9. Quality of Life	*quality, happy, social, community, life, benefit(s), adopt, fear, deportation, living, job(s), activities, family*
10. Cultural Identity	*identity, social, value(s), Reagan, Lincoln, conservative(s), liberal(s), nation, America, American(s), community, communities, country, dreamers, immigrants, refugees, history, historical*
11. Public Sentiment	*public, sentiment, opinion, poll(s), turning, survey, support, American(s), reform, action, want, need, vote*
12. Political Factors & Implications	*politic(s), political, stance, view, (bi)partisan, filibuster, lobby, Republican(s), Democrat(s), House, Senate, Congress, committee, party, POTUS, SCOTUS, administration, GOP*

continued on next page

Table 4.5.: *Frame and Corresponding Unigrams, continued.*

Issue	Frame and Corresponding Unigrams
13. Policy Description, Prescription, & Evaluation	*policy, fix(ing), work(s), working, propose(d), proposing, proposal, solution, solve, outcome(s), bill, law, amendment, plan, support, repeal, reform*
14. External Regulation and Reputation	*regulation, US, ISIS, ISIL, relations, international, national, trade, foreign, state, border, visa, ally, allies, united, refugees, leadership, issues, Iraq, Iran, Syria, Russia, Europe, Mexico, Canada*
15. Factual	*health, insurance, affordable, deadline, enroll, sign, signed, program, coverage*
16. (Self) Promotion	*statement, watch, discuss, hearing, today, tonight, live, read, floor, talk, tune, opinion, TV, oped*
17. Personal Sympathy & Support	*victims, thoughts, prayer(s)(ing), family, stand, support, tragedy, senseless, condolences*

These predicates (presented in Table 4.6) are therefore weakly supervised and are further combined into the probabilistic rules of each global PSL model as shown in Table 4.7. PSL allows us to build connections between all of the weaker models (e.g., models using unigram features only) in order to improve our overall prediction.

Our overall prediction task is to design a PSL model capable of predicting the frame(s) of a given tweet. We define this prediction goal as a target predicate in PSL notation: Frame(T, F). Here, T represents a tweet and F represents one of the 17 frames listed in Table 4.1.

Table 4.6.
Descriptions of PSL Predicates. Model Basis describes the general type of features represented by the listed predicates. No. is used to identify which predicates are combined together into the PSL Models of Table 4.7. Key Features lists the tweet information represented by these predicates. The final column shows the features represented in PSL predicate notation.

Model Basis	No.	Key Features	PSL Predicates
Linguistic	1	Unigrams	$\textsc{Unigram}_F$(T, U)
	2	Unigram Similarity	MaxSim (T, F)
	3	Bigrams (Party)	$\textsc{Bigram}_P$(T, B)
	4	Trigrams (Party)	$\textsc{Trigram}_P$(T, TG)
	5	Bigrams (Party & Issue)	$\textsc{Bigram}_{IP}$(T, B)
	6	Trigrams (Party & Issue)	$\textsc{Trigram}_{IP}$(T, TG)
Directly Observable	7	Party	Party(T, P)
	8	Issue	Issue(T)
Social Behavioral	9	Temporal Activity	SameTime(T1, T2)
	10	Retweet Patterns	Retweets(T1, T2)
	11	Following Network	Follows(T1, T2)

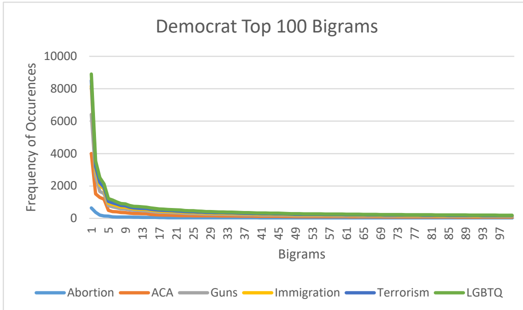
4.3.1 Directly Observed Information

The two classifiers in this section are based on easy-to-observe information: political party affiliation and the issue of the tweet. Party affiliation is domain dependent and issues can be easily extracted from tweets using keywords or phrases.

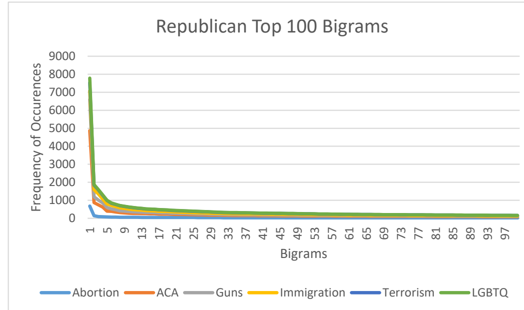
Political Party Affiliation of Author. Framing behavior can be indicative of ideology and party affiliation. To investigate if political party knowledge can improve frame prediction, we use the predicate: Party(T, P), which indicates that tweet T was written by a politician in party P.

Table 4.7.
Examples of PSL Model Combinations and Rules. Each model adds to the rules of the previous model. Model Basis indicates the source of information. Dir Obs represents directly observed information including political party affiliation and issue. Ling represents linguistic features. Soc represents social and Behav represents behavioral features. Model indicates the model label that appears in the tables of Section 4.4. Predicate Numbers represent the predicates by number (Column 2 of Table 4.6) that appear in the rules of this model. The last column provides an example of what rules in this model look like in PSL notation.

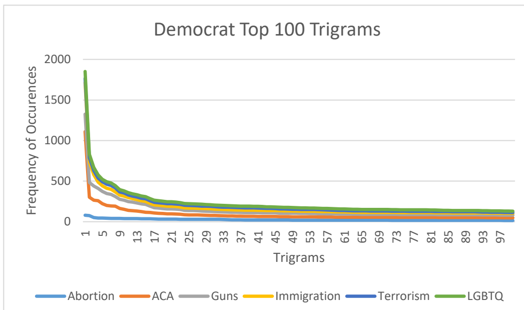
Model Basis	Model	Predicate Number	Examples of PSL Rules
Ling	Uni	1	Unigram_F(T, U) → Frame(T, F)
	Sim	1, 2	Unigram_F(T, U) ∧ MaxSim(T, F) → Frame(T, F)
	U+S	1, 2	MaxSim(T, F)) → Frame(T, F)
	Big_P	1, 2, 3, 4, 7, 8	Unigram_F(T, U) ∧ Bigram_P(T, B) → Frame(T, F)
	Tri_P	1, 2, 3, 4, 7, 8	Bigram_P(T, B) ∧ Trigram_P(T, TG) → Frame(T, F)
	Big_{IP}	1, 2, 3	Unigram_F(T, U) ∧ Bigram_{IP}(T, B) → Frame(T, F)
	Tri_{IP}	1, 2, 3, 4	Unigram_F(T, U) ∧ Trigram_{IP}(T, TG) → Frame(T, F)
Dir Obs	Pol	1, 7	Party(T, P) → Frame(T, F)
	Iss	1, 7, 8	Issue(T, I) → Frame(T, F)
Soc Behav	Temp	1, 2, 3, 4, 9	SameTime(T1, T2) ∧ Frame(T1, F) → Frame(T2, F)
	RTs	1, 2, 3, 4, 9, 10	Retweets(T1, T2) ∧ Frame(T1, F) → Frame(T2, F)
	Fol	1, 2, 3, 4, 9, 10, 11	Follows(T1, T2) ∧ Frame(T1, F) → Frame(T2, F)



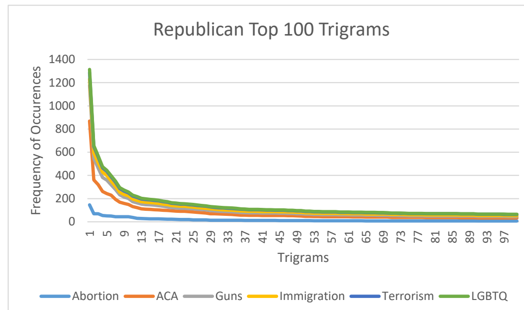
(a) Democrat Bigrams



(b) Republican Bigrams



(c) Democrat Trigrams



(d) Republican Trigrams

Figure 4.3. Distributions of Bigrams and Trigrams by Party.

Issue Discussed in Tweet. We observe that politicians from different parties will present issues differently. For example, on the issue of gun control, Republicans are known for discussing the issue in terms of an individual's rights (Frame 5), while Democrats frame the issue as a need for safety (Frame 7). This is represented by the predicate: ISSUE(T, I), where I represents the issue of tweet T. In Section 4.2, we simultaneously filtered and labeled our collected tweets to only those tweets which discuss our six issues of interest. For some tweets, issues may overlap. For example, if a tweet discusses gun control measures in the context of an ISIS-credited shooting, then the tweet will have two labels: guns and terrorism.

Table 4.8.: Top 20 Bigrams Lists of Democrats. All lists the issue-independent bigrams. The other six issues correspond to issue-dependent bigrams.

Issue	Top 20 Democrat Bigrams
All	*gun violence, health care, Affordable Care, Care Act, LGBT community, suspected terrorists, women's health, Planned Parenthood, background checks, gun safety, thoughts prayers, health insurance, House floor, keep guns, take action, Supreme Court, common sense, gun reform, victims families, action gun*
Abo	*women's health, Planned Parenthood, birth control, reproductive rights, attack women's, #hobbylobby decision, attacking women's, reproductive health, right choose, women's reproductive, woman's right, defund #PPFA, health services, Supreme Court, health decisions, protect women, #notmybossbusiness act, attacks women, select committee*
ACA	*health care, health insurance, affordable care, care act, thanks ACA, affordable health, million Americans, women's health, access health, repeal ACA, care reform, millions Americans, quality affordable, mental health, open enrollment, uninsured rate, thanks affordable, Medicare Medicaid, Medicaid expansion*
Guns	*gun violence, thoughts prayers, gun safety, background checks, House floor, victims families, keep guns, violence prevention, gun control, prayers victims, must act, commonsense gun, gun laws, reduce gun, common sense, Congress must, prayers go, suspected terrorists, end gun, prevent gun*

continued on next page

Table 4.8.: *Top 20 Democrat Bigrams, continued.*

Issue	Top 20 Democrat Bigrams
Immig	*immigration reform, immigration system, comprehensive immigration, broken immigration, fix broken, homeland security, border security, Senate passed, immigration bill, President Obama, bipartisan immigration, reform #timeisnow, pass immigration, pass comprehensive, discuss immigration, town hall, families together, Obama's #immigrationaction, nation immigrants, action immigration*
LGBTQ	*marriage equality, LGBT Americans, LGBT community, LGBT equality, LGBT rights, discrimination LGBT, LGBT youth, LGBT discrimination, protect LGBT, support LGBT, Supreme Court, LGBT people, civil rights, proud stand, samesex marriage, gay marriage, proud join, equality LGBT, #lgbtequality day, stand bullying*
Ter	*Boko Haram, Middle East, #bringbackourgirls #joinrepwilson, Iraq War, Iraq Syria, Syrian refugees, terrorist attacks, suspected terrorists, terrorist attack, fight ISIS, Iraq Afghanistan, terror suspects, watch list, Congress must, last 11, 11 years, weapons US, bought weapons, Syria Iraq, American people*

Table 4.9.: Top 20 Bigrams Lists of Republicans. All lists the issue-independent bigrams. The other six issues correspond to issue-dependent bigrams.

Issue	Top 20 Republican Bigrams
All	*health care, @HouseAppropsGOP @RepHalRogers, terrorist attacks, thoughts prayers, Planned Parenthood, Obamacare repeal, Middle East, repeal Obamacare, terrorist attack, immigration laws, illegal immigrants, victims families, radical Islam, fix health, #betterway fix, Supreme Court, Syrian refugees, @SenateMajLdr McConnell, executive amnesty, next year*
Abo	*Planned Parenthood, defund Planned, unborn children, unborn child, Protection Act, Child Protection, pain capable, 20 weeks, Hobby Lobby, taxpayer funding, funding Planned, protect unborn, bill protect, capable unborn, women's health, prolife bill, funding abortion, unborn babies, protect children, every child*
ACA	*health care, care law, health insurance, repeal Obamacare, Care Act, Affordable Care, healthcare law, care reform, due Obamacare, American people, Obamacare website, watch live, House floor, care plan, small businesses, President's health, employer mandate, delay Obamacare, care costs, Obama admin*
Guns	*thoughts prayers, prayers go, gun control, victims families, prayers victims, prayers family, prayers families, loved ones, prayers affected, terrorist attack, prayers people, #2a rights, gun laws, family friends, gun violence, please keep, Fort Hood, first responders, background checks, last night*

continued on next page

Table 4.9.: *Top 20 Republican Bigrams, continued.*

Issue	Top 20 Republican Bigrams
Immig	*immigration reform, border security, illegal immigrants, immigration laws, illegal immigration, executive amnesty, secure border, homeland security, border crisis, immigration bill, President Obama's, immigration system, immigration law, executive action, Obama's executive, southern border, immigration actions, Obama's immigration, President Obama, action immigration*
LGBTQ	*religious freedom, protect religious, gay marriage, religious freedoms, Supreme Court, bathroom directive, Obama administration, Jefferson's statue, Virginia passed, statute protect, life religious, freedom military, years since, transgender bathroom, right religious, since Virginia, marriage penalty, defend religious, military religious, school bathroom*
Ter	*Middle East, terrorist attack, terrorist attacks, Syrian refugees, defeat ISIS, fight ISIS, terror attacks, Iraq Syria, terror attack, national security, radical Islamic, ISIS threat, Gitmo terrorists, watch live, President Obama, terrorists US, Paris attacks, United States, strategy defeat, President Obama's*

Table 4.10.: Top 20 Trigrams Lists of Democrats.

Issue	Top 20 Democrat Trigrams
All	*Affordable Care Act, action gun violence, common sense gun, prevent gun violence, gun violence #nobillnobreak, terrorists buying guns, sense gun reform, end gun violence, keep guns hands, address gun violence, basic health plan, terror watch list, suspected terrorists buying, time Congress act, gun violence prevention, quality affordable health, care act enrollment, gun violence enough, comprehensive immigration reform, acting victims families*
Abo	*attacking women's health, attack women's health, women's reproductive rights, women's health #standwithpp, Planned Parenthood funding, woman's right choose, Roe v Wade, Americans don't want, another #gopshutdown planned, #gopshutdown Planned Parenthood, want another #gopshutdown, don't want another, attacks women's health, SCOTUS #hobbylobby decision, defund Planned Parenthood, defend reproductive rights, #housedemocrats recommit #stopthesham, recommit #stopthesham defend, #stopthesham defend reproductive, WWH v Hellerstedt*
ACA	*Affordable Care Act, health care reform, affordable health care, access health care, women's health care, thanks Affordable Care, mental health care, quality affordable health, health care law, health care coverage, quality health care, health insurance coverage, health care decisions, health care millions, vote repeal ACA, affordable health insurance, sign health insurance, enroll health insurance, gained health coverage, health care services*

continued on next page

Table 4.10.: *Top 20 Democrat Trigrams, continued.*

Issue	Top 20 Democrat Trigrams
Guns	*gun violence prevention, thoughts prayers victims, reduce gun violence, thoughts prayers go, prevent gun violence, end gun violence, gun violence #nobillnobreak, background checks gun, action gun violence, prayers victims families, keep guns hands, thoughts prayers family, common sense gun, Congress must act, address gun violence, victims gun violence, thoughts prayers families, gun violence epidemic, checks gun sales, thoughts prayers people*
Immig	*comprehensive immigration reform, broken immigration system, fix broken immigration, immigration reform #timeisnow, pass immigration reform, since Senate passed, bipartisan immigration reform, pass comprehensive immigration, President Obama's #immigrationaction, immigration reform #cir, immigration reform bill, support immigration reform, keep families together, Senate passed bipartisan, need comprehensive immigration, need immigration reform, homeland security funding, fix immigration system, president's actions immigration, passed immigration reform*
LGBTQ	*discrimination LGBT Americans, support LGBT youth, end discrimination LGBT, LGBT Pride Month, right side history, support marriage equality, take stand bullying, civil rights time, fight civil rights, rights time write, next chapter fight, LGBT equality next, equality next chapter, time write books, protect LGBT Americans, chapter fight civil, books law equality, protects LGBT Americans, end LGBT discrimination*

continued on next page

Table 4.10.: *Top 20 Democrat Trigrams, continued.*

ISSUE	TOP 20 DEMOCRAT TRIGRAMS
TER	*terror suspects bought, suspects bought weapons, bought weapons US, last 11 years, terrorist watch list, protect suspected terrorists, time Congress act, weapons US time, abducted Boko Haram, US time Congress, Congress act #nomoresilence, take action protect, US take action, @repesty @speakerryan #nobillnobreak, expect US take, @speakerryan #nobillnobreak American, action protect suspected, American people expect, people expect US, suspected terrorists criminals*

Table 4.11.: Top 20 Trigrams Lists of Republicans.

ISSUE	TOP 20 REPUBLICAN TRIGRAMS
ALL	*fix health care, #betterway fix health, defund Planned Parenthood, @HouseApropsGOP @RepHalRogers bill, health care system, kept Senate passes, calling full declassification, health care law, detainees terrorist activity, prayers victims families, GTMO detainees terrorist, Orlando terrorist attack, Child Protection Act, Obamacare repeal bill, radical Islamic terrorism, declassification review GTMO, unborn child protection, terrorist attacks Paris, review GTMO detainees, promise kept Senate*

continued on next page

Table 4.11.: *Top 20 Republican Trigrams, continued.*

Issue	Top 20 Republican Trigrams
Abo	*defund Planned Parenthood, unborn child protection, Child Protection Act, funding Planned Parenthood, pain capable unborn, capable unborn child, protect children disabilities, prolife bill protect, every child #blessingnotburden, bill protect children, supports prolife bill, taxpayer funding abortion, Roe V Wade, defunding Planned Parenthood, federal funding Planned, Planned Parenthood video, #paincapable unborn child, Obamacare defund Planned, abortions 20 weeks, legislation defund Planned*
ACA	*health care law, Affordable Care Act, health care reform, health care plan, President's health care, health care costs, health care bill, health care system, like health care, mental health care, special inspector general, medical device tax, Sigma bill create, supporting Sigma bill, care plan keep, another day another, Obamacare special inspector, bill create Obamacare, create Obamacare special, another Obamacare delay*
Guns	*thoughts prayers go, thoughts prayers victims, thoughts prayers family, prayers victims families, thoughts prayers families, thoughts prayers affected, thoughts prayers people, send thoughts prayers, prayers go victims, sending thoughts prayers, thoughts prayers everyone, prayers go family, prayers go families, begun 10 minutes, 10 minutes debate, Orlando terrorist attack, prayers family friends, don't need gun, need gun laws*

continued on next page

Table 4.11.: *Top 20 Republican Trigrams, continued.*

Issue	Top 20 Republican Trigrams
Immig	*enforce immigration laws, Obama's executive amnesty, executive action immigration, Senate immigration bill, broken immigration system, President Obama's executive, President's executive amnesty, President Obama's immigration, discuss immigration reform, executive actions immigration, amnesty illegal immigrants, discussing immigration reform, nations immigration laws, immigration system broken, enforcement immigration laws, illegal immigrant children, criminal illegal immigrants, comprehensive immigration reform, fix broken immigration, immigration executive action*
LGBTQ	*protect religious freedom, statute protect religious, Jefferson's statute protect, years since Virginia, since Virginia passed, religious freedom military, right religious freedom, life religious freedom, military religious freedom, passed Jefferson's statute, defend religious freedom, victory religious freedom, marks 229 years, protecting religious freedom, rights religious freedom, religious freedom #letthemserve, promote religious freedom, religious freedom today, attack religious freedom, school bathroom directive*
Ter	*strategy defeat ISIS, military action Syria, radical Islamic terrorism, Syrian refugees US, Benghazi terrorist attack, Gitmo terrorists US, military intervention US, plan defeat ISIS, radical Islamic terrorists, state sponsor terror, terrorists US soil, House foreign chairman, VISA waiver program, military force Syria, use social media, state sponsor terrorism, terrorist attacks Brussels, use military force, ally Middle East, terrorist attacks Paris*

4.3.2 Linguistic Information

The classifiers of this section represent linguistic information that may indicate the frame of the tweet. This includes unigrams that may indicate the frame, words that are similar to these unigrams, and political slogans represented as bigrams and trigrams.

Unigrams and Similarity of Unigrams. Using the guidelines provided in the Policy Frames Codebook of Boydstun et al. (2014), we adapted a list of expected unigrams for each frame. These lists are shown in Table 4.5. For example, unigrams that should be related to Frame 12 (*Political Factors & Implications*) include filibuster, lobby, Democrats, and Republicans, because this frame deals with political maneuvering or strategies, such as filibusters, lobbying, or appealing to the party or constituency. We expect that if a tweet and frame contain a matching unigram, then that frame is likely used to express that tweet. The information that tweet T has expected unigram U of frame F is represented with the PSL predicate: $\textsc{Unigram}_F(\mathrm{T},\ \mathrm{U})$. This knowledge is then used as input to PSL Model $\textsc{Uni}$ via the rule: $\textsc{Unigram}_F(\mathrm{T},\ \mathrm{U}) \rightarrow\ \textsc{Frame}(\mathrm{T},\ \mathrm{F})$ (shown in line one of Table 4.7).

However, not every tweet will have a unigram that matches those in these lists. Under the intuition that at least one unigram in a tweet should be *similar* to a unigram in the list, we designed the following *MaxSim* metric to compute the maximum similarity between a word in a tweet and a word from the list of unigrams per frame.

$$\textsc{MaxSim}(\mathrm{T},\ \mathrm{F}) = \underset{\mathrm{u} \in F, \mathrm{w} \in T}{\arg\max}\, \textsc{Similarity}(\textsc{w},\ \textsc{u}) \tag{4.1}$$

T is a tweet, $\textsc{w}$ is each word in T, and $\textsc{u}$ is each unigram in the list of expected unigrams (per frame). $\textsc{Similarity}$ is the computed `word2vec` similarity (using pretrained embeddings) of each word in the tweet with every unigram in the list of unigrams for each frame. The frame F of the maximum scoring unigram is input to the PSL predicate: $\textsc{MaxSim}_F(\mathrm{T},\ \mathrm{F})$, which indicates that tweet T has the highest similarity to frame F.

Bigrams and Trigrams. In addition to unigrams, we also explored the effects of political party *slogans* on frame prediction. Slogans are common catch phrases or sayings that people typically associate with different U.S. political parties. For example, Republicans are known for using the phrase "repeal and replace" when they discuss the ACA. Similarly, in the 2016 U.S. presidential election, Secretary Hillary Clinton's campaign slogan became "Love Trumps Hate". We hypothesized that such slogans would either directly correspond to frames or be indicative of the frames used to express issues associated with these slogans.

To visualize slogan usage by parties for different issues, we used the *entire* tweets dataset, including all unlabeled tweets, to extract the top bigrams and trigrams per party for each issue. The histograms in Figure 4.3 show these distributions for the top 100 bigrams and trigrams. Based on these results, we use the top twenty bigrams shown in Tables 4.8 and 4.9 (e.g., *women's healthcare* and *immigration reform*) and the top twenty trigrams shown in Tables 4.10 and 4.11 (e.g., *prevent gun violence*) as input to PSL predicates $\textsc{Bigram}_{IP}(\mathrm{T},\ \mathrm{B})$ and $\textsc{Trigram}_{IP}(\mathrm{T},\ \mathrm{TG})$. These predicates represent that tweet T has bigram B or trigram TG from the respective issue I phrase lists of either party P.

Since frames are designed to generalize across issues, we also explore the top twenty bigrams and trigrams used by each party, regardless of the issue. These lists appear in the first row of Tables 4.8, 4.9, 4.10, and 4.11. This information is represented by the PSL predicates $\textsc{Bigram}_{P}(\mathrm{T},\ \mathrm{B})$ and $\textsc{Trigram}_{P}(\mathrm{T},\ \mathrm{TG})$. These predicates represent that tweet T has bigram B or trigram TG from the respective issue-independent phrase lists of either party P.

4.3.3 Twitter Behavior and Social Information

In addition to directly observed information and language based features of tweets, we also exploit the behavioral and social features of Twitter including similarities between temporal activity and network relationships.

Temporal Similarity. We construct a temporal histogram for each politician which captures their Twitter activity over time. Politicians are most likely to tweet about an event within hours of its occurrence. Similarly, most politicians tweet about the event most frequently the day of the event and this frequency decreases over time. From these temporal histograms, we observed that the frames used the day of an event were similar and gradually changed over time. For example, once the public is notified of a shooting, politicians respond with the *Personal Sympathy & Support* frame (Frame 17) to offer sympathy to the victims and their families. Over the next days or weeks, both parties slowly transition to using additional frames. For example, Democrats use Frame 7 to argue for gun control legislation. To capture this behavior we use the PSL predicate SameTime(T1, T2). This indicates that tweet T1 occurs around the same time as tweet T2. We conducted experiments with different hour and day limits and found that using a time frame of one hour results in the best accuracy while also limiting noise. This information is used in PSL Model Temp via rules such as: SameTime(T1, T2) & Frame(T1, F) $\rightarrow$ Frame(T2, F), as shown in line four of Table 4.7.

Network Similarity. Finally, we expect that politicians who share ideologies, and thus are likely to frame issues similarly, will retweet and/or follow each other on Twitter. Due to the compound nature of tweets, retweeting with additional comments can add more frames to the original tweet. Additionally, politicians on Twitter are more likely to follow members of their own party or similar non-political entities than those of the opposing party. To capture this network-based behavior we use two PSL predicates: Retweets(T1, T2) and Follows(T1, T2). These predicates indicate that the content of tweet T1 includes a retweet of tweet T2 and that the author of T1 follows the author of T2 on Twitter, respectively. The last two lines of Table 4.7 show examples of how network similarity is incorporated into PSL rules.

4.4 Experimental Results

One overall goal of this dissertation is to provide a modeling framework that can easily adapt to the dynamic nature of political discourse on Twitter. Because of the possibility of frame overlap, the frame prediction task is a 17-class multilabel classification task. As described in Section 4.2, manual annotation of frames is both difficult and time-consuming. Furthermore, annotations of tweets written in 2016 may not be as useful in the future as administrations and relevant issues change. Therefore, we designed the weakly supervised classifiers of Section 4.3 to capture basic, observed features as well as higher levels of abstraction of features which will generalize to future datasets.

These features are combined into a variety of PSL models which are tested in two settings, referred to in this section as supervised and unsupervised. In the supervised experiments, we train and test with the labeled 2,050 tweets of the Congressional Tweets Dataset to choose our best features. In the unsupervised setting, we learn the weights of the models using the unlabeled set of these tweets.

Experimental Settings. We provide an analysis of our PSL models under both supervised and unsupervised settings. In the supervised experiments, we used five-fold cross validation with randomly chosen splits, while also ensuring that all frames were represented in the splits. The goal of our supervised experiments was to learn how different attributes of politicians and their tweets interact with each other to contribute to the prediction score.

Traditionally, this task can be viewed as text categorization, typically approached with a classifier, such as an SVM, using bag-of-words features. The results of this approach on our dataset are shown in column 2 of Table 4.12. In this scenario, the SVM tends to prefer the majority class, which results in many incorrect labels. Column 3 shows the results of using an SVM with bag-of-words features to perform multilabel classification. This approach decreases the F_1 score for a majority of frames. Both SVMs also result in F_1 scores of 0 for some frames, further lowering the

Table 4.12.
Baseline and Skyline Micro-weighted Average F_1 Scores. SVM Indiv. is the SVM trained to predict one frame. SVM Multi. is the multiclass SVM. PSL U+S is the adapted unigram PSL model. PSL Fol is the collective network using linguistic, social, and behavioral features.

Setting	SVM Indiv.	SVM Multi.	PSL U+S	PSL Fol
Supervised	28.67	18.90	66.02	77.79
Unsupervised	—	—	37.14	58.66

overall performance. Finally, columns 4 and 5 show the results of using our worst and best PSL models, respectively. PSL Model U+S, which uses our adapted unigram and unigram similarity features (predicates 1 and 2 in Table 4.6) instead of the bag-of-words features for multilabel classification, serves as our strongest unigram-based baseline to improve upon. Additionally, the PSL model of the supervised, collective network setting represents the best results we can achieve.

We also explore the results of our PSL models in an unsupervised setting because the highly dynamic nature of political discourse on Twitter makes it unrealistic to expect annotated data to generalize to future discussions. The only source of supervision comes from the initial unigrams lists and party information as described in Section 4.3. The labeled tweets are used for evaluation only. As seen in Table 4.12, we are able to improve the best unsupervised model to within an F_1 score of 7.36 points of the unigram baseline of 66.02, and 19.13 points of the best supervised score of 77.79.

4.4.1 Detailed Analysis of Linguistic Indicators

Our weakly supervised classifiers of Section 4.3 extract information that rely on domain knowledge, such as the Policy Frames Codebook (Boydstun et al., 2014). Bigrams and trigrams, which correspond to political slogans, can also be extracted.

However, both types of linguistic features are noisy and do not perform well in the unsupervised setting, as described in greater detail in the following analyses sections.

Tables 4.13 and 4.15 present our supervised and unsupervised (never previously published) results using directly observed or linguistic based features alone: party, issue, unigrams, unigram similarity, and bigrams and trigrams by party only. These features correspond to predicate numbers 1, 2, 3, 4, 7, and 8 in Table 4.6. These predicates are used to construct the rules of the following non-collective PSL models in Table 4.7: UNI which uses rules based on the adapted unigrams, POL which uses adapted unigrams and political party information, ISS which uses the previous features as well as the issue of the tweet, SIM which uses the previous features plus unigram similarity, and BIG$_P$ and TRI$_P$ which add issue-independent bigrams and trigrams to all previous rules, respectively.

As described in the following analyses sections, our observations led us to hypothesize that by conditioning these linguistic features on other information, for example, the issue of the tweet, we could improve the accuracy of these features. Based on this, we created the first three PSL models (U+S, BIG$_{IP}$, and TRI$_{IP}$) shown in Tables 4.14 and 4.16. These PSL models combine linguistic features with political party and issue information to improve the overall accuracy of the linguistic models. Model U+S combines adapted unigrams and unigram similarity into one model. Model BIG$_{IP}$ combines rules using unigrams, unigram similarity and bigrams conditioned on both issue and political party. Similarly, Model TRI$_{IP}$ combines these unigrams and bigrams with issue and party dependent trigrams. This model augments the features used in the best performing linguistic model TRI$_P$ shown in Tables 4.13 and 4.15.

Based on improvements from these models, we then conditioned the linguistic features on social and behavioral information, which is represented by the PSL models TEMP which incorporates similar temporal activity, RTS which includes retweet patterns, and FOL which adds in the follower network information. As can be seen when comparing Tables 4.13 and 4.15 with Tables 4.14 and 4.16, more precise features conditioned on social and behavioral information result in more accurate predictions, especially in the unsupervised setting.

4.4.2 Analysis of Supervised Experiments

Table 4.13 compares the results of our models in the supervised, non-collective setting. Overall, prediction improves as the model has access to more information. We note that, unlike simple text-categorization problems which can often achieve near-optimal performance using bag-of-words features alone, frame prediction requires more nuanced information. Model UNI, which uses features similar to bag-of-words, achieves an F_1 score of only 52.21. However, our experiments show that connecting linguistic features with additional information (e.g., party affiliation, issue relevance, and party associated key phrases) improves the per-model F_1 score dramatically, up to 75.95 for Model TRI_P. This trend occurs in both the overall weighted average F_1 score and for most frame types individually.

For 13 of the 17 frames, the F_1 score of Model TRI_P exceeds the average inter-annotator agreement. Interestingly, for some frames (e.g., *Capacity & Resources* (Frame 2), *Quality of Life* (Frame 9), *Cultural* (Frame 10), *External Regulation* (Frame 14), *Factual* (Frame 15)) the addition of party bigram information does not improve upon the prediction of the previous model. However, the addition of party trigram information in Model TRI_P is able to further improve the results, indicating that trigrams are more useful for frame classification than bigrams or unigrams.

Table 4.14 shows the results of our supervised, collective experiments. Here we can see that by adding Twitter behavior (beginning with Model TEMP), our behavior-based models achieve the best F_1 scores across all frames. Model TEMP achieves the highest results on two frames, suggesting retweeting and network follower information do not help improve the prediction score for these frames. Similarly, Model RTS achieves the highest prediction for five of the frames, suggesting the network follower information of Model FOL cannot further improve the score for these frames. Overall,

Table 4.13.
F_1 Scores of Supervised, Non-collective PSL Models Using Directly Observed and Linguistic Based Features. The highest prediction per frame is marked in bold. The non-collective setting does not exploit the social network dependencies present in our data.

Frame No.	Frame	Results of Supervised PSL Model Frame Predictions					
		Uni	Pol	Iss	Sim	Big_P	Tri_P
1	Economic	72.13	73.68	79.63	81.32	81.63	**85.11**
2	Capacity	14.29	14.29	44.44	66.67	66.67	**82.35**
3	Morality	39.58	39.17	45.25	57.78	66.67	**88.46**
4	Fairness	63.56	67.83	65.19	69.91	79.53	**82.35**
5	Legality	57.96	58.91	63.32	63.27	60.24	**67.57**
6	Crime	60.0	60.0	60.87	60.87	61.54	**63.64**
7	Security	60.0	60.49	65.16	72.9	75.86	**83.12**
8	Health	63.41	66.94	67.42	70.13	72.47	**75.68**
9	Quality	30.19	31.82	45.1	55.17	55.17	**76.47**
10	Cultural	20.0	31.58	47.06	66.67	66.67	**88.89**
11	Sentiment	12.25	15.25	24.62	24.24	26.24	**29.41**
12	Political	57.23	58.25	60.76	65.22	69.57	**73.92**
13	Policy	31.25	32.7	39.23	40.94	44.34	**65.43**
14	External	50.0	56.15	64.71	72.73	72.73	**85.71**
15	Factual	64.0	68.97	71.43	81.82	81.82	**82.35**
16	Promotion	68.52	69.51	75.91	76.81	77.1	**82.05**
17	Personal	70.34	72.58	69.15	71.53	76.92	**91.07**
	Weighted Avg.	52.21	54.3	59.0	63.54	66.37	**75.95**

the Twitter behavior based models are able to outperform language based models alone, including the best performing language model (Model Tri_{IP}) which combines unigrams and issue-dependent bigrams and trigrams together to collectively infer the correct frames.

Table 4.14.
F_1 Scores of Supervised, Collective PSL Models Using Linguistic, Social, and Behavioral Features. The highest prediction per frame is marked in bold.

Frame No.	Frame	Results of Supervised PSL Model Frame Predictions					
		U+S	BIG_{IP}	TRI_{IP}	TEMP	RTs	FOL
1	ECONOMIC	85.19	85.19	86.73	87.72	87.72	**89.88**
2	CAPACITY	55.38	61.54	76.71	77.11	77.11	**79.55**
3	MORALITY	73.39	80.52	86.95	87.5	**87.43**	**87.43**
4	FAIRNESS	63.56	67.83	65.19	69.91	79.53	**82.35**
5	LEGALITY	80.41	80.78	80.79	**83.33**	81.79	82.16
6	CRIME	54.55	54.55	66.67	**76.92**	**76.92**	**76.92**
7	SECURITY	84.40	82.14	84.10	86.67	86.67	**88.48**
8	HEALTH	73.50	75.76	75.59	77.46	**79.71**	**79.71**
9	QUALITY	69.39	68.00	69.39	72.34	72.34	**82.93**
10	CULTURAL	75.86	78.57	81.25	81.25	81.25	**85.71**
11	SENTIMENT	12.25	15.25	24.62	24.24	26.24	**29.41**
12	POLITICAL	54.21	63.31	74.33	74.42	**74.52**	**74.52**
13	POLICY	55.75	58.87	60.25	61.54	64.06	**65.06**
14	EXTERNAL	60.71	59.15	64.71	74.35	74.35	**85.71**
15	FACTUAL	66.56	68.00	71.43	81.82	80.82	**82.85**
16	PROMOTION	85.71	86.46	86.58	87.34	87.33	**91.76**
17	PERSONAL	71.79	71.71	74.73	75.00	**77.55**	**77.55**
	WEIGHTED AVG.	66.02	68.78	72.49	74.40	75.71	**77.79**

4.4.3 Analysis of Unsupervised Experiments

Table 4.15 shows that the overall trend for the unsupervised PSL models is similar for most frames: adding additional information to the linguistic features continually increases the F_1 score. This trend, however, does not hold across all models for all frames, and for some frames (e.g., *Security & Defense*, *Political Factors*, *Policy Description*, and *(Self) Promotion)* the final linguistic Model TRI_P prediction is not the highest. The reduced performance is due to the complexity added to each successive model, which can be exploited when learning in a supervised setting but not when no external supervision is available.

Interestingly, we expected the addition of author party information (Model Pol) and issue of the tweet (Model Iss) to improve performance beyond what was observed. However, our experiments show that party information only becomes useful when considered in conjunction with popular phrases by party (i.e., bigrams or trigrams, used in Models Big_P and Tri_P). Additionally, for frames whose prediction drops under Models Pol, Iss, and Sim, the addition of party bigrams and trigrams (Models Big_P and Tri_P, respectively) allows the models to recover from this, further supporting the usefulness of party-based trigrams as found by the supervised PSL models.

Conversely, the additional information provided by Models Big_P and Tri_P lowered performance in some cases (e.g., *Security & Defense*, *Political Factors*, *Policy Description*, *(Self) Promotion*). Upon investigation, we found that a majority of tweets from these frames match to phrase indicators from *both* parties, introducing contradictory noise (i.e., the model information clashes with that of Models Pol, Iss, and Sim) into the model and lowering results. This is also reflected in the weighted averages of the F_1 scores shown in the last line of Table 4.15, which drop in Models Sim and Big_P but are the highest in Model Tri_P.

Finally, the contribution of adding issue information is very limited. This is to be expected, as frames are designed to generalize across all issues of interest. Following the analysis of Boydstun et al. (2014) and Card et al. (2015) which focused on three issues (immigration, smoking bans, and same-sex marriage), our study looks into six issues and the results appear to confirm that the frames *do generalize across issues.*

Overall, we are able to achieve F_1 scores above 50% (where random chance is 5.8%) for half of the frames in an unsupervised, non-collective setting.

As shown in Table 4.16, Model Fol, the combination of language and Twitter behavior features achieves the best results on 16 of the 17 issues. There are a few interesting aspects of the unsupervised setting which differ from the supervised setting.

Table 4.15.
F_1 Scores of Unsupervised, Non-collective PSL Models Using Linguistic Features. The highest prediction per frame is marked in bold. The non-collective setting does not exploit the social network dependencies present in our data.

Frame No.	Frame	Results of Unsupervised PSL Model Frame Predictions					
		Uni	Pol	Iss	Sim	Big_P	Tri_P
1	Economic	31.82	31.47	31.9	34.25	49.18	**70.0**
2	Capacity	23.38	23.38	23.38	24.49	60	**66.67**
3	Morality	28.63	28.63	28.92	28.86	32.65	**35.16**
4	Fairness	33.49	33.53	29.15	48.55	42.59	**53.99**
5	Legality	44.58	44.58	45.02	44.75	60.15	**66.67**
6	Crime	7.88	7.88	7.88	7.88	31.57	**60.0**
7	Security	42.5	42.87	**41.1**	36.75	26.09	22.22
8	Health	48.36	48.89	48.54	52.55	60.08	**70.33**
9	Quality	17.82	18.06	25.17	21.75	35.55	**40.0**
10	Cultural	15.38	15.83	15.42	16.95	18.6	**25.81**
11	Sentiment	15.22	15.72	15.28	16.67	18.18	**19.35**
12	Political	49.06	49.06	49.68	**49.75**	31.35	28.57
13	Policy	39.88	40.0	**40.36**	32.63	19.4	15.38
14	External	12.66	12.66	12.71	14.29	32.0	**44.44**
15	Factual	24.64	25.0	25.37	34.34	60.0	**66.67**
16	Promotion	50.11	49.89	51.09	**54.04**	26.47	48.0
17	Personal	45.36	45.37	57.38	35.34	71.86	**78.79**
	Weighted Avg.	38.26	38.36	39.59	38.61	37.61	**43.33**

Six of the frame predictions do worse in Model Big_{IP}, which is double that of the supervised version. This is likely due to the presence of overlapping bigrams across frames and issues. For example, "women's healthcare" could appear in both Frames 4 and 8 and the issues of ACA and abortion. However, all six are able to improve with the addition of trigrams (Model Tri_{IP}), whereas only one of three frames improves in the supervised setting. This further supports that bigrams may not be as useful

Table 4.16.
F_1 Scores of Unsupervised, Collective PSL Models Using Linguistic, Social, and Behavioral Features. The highest prediction per frame is marked in bold.

Frame No.	Frame	Results of Unsupervised PSL Model Frame Predictions					
		U+S	Big$_{IP}$	Tri$_{IP}$	Temp	RTs	Fol
1	Economic	31.82	31.52	69.57	72.22	72.22	**73.23**
2	Capacity	23.38	28.51	40.00	**41.18**	**41.18**	**41.18**
3	Morality	28.63	29.41	47.67	53.98	43.06	**53.99**
4	Fairness	33.49	47.19	59.15	63.50	63.50	**64.74**
5	Legality	44.58	46.93	58.02	60.64	60.63	**64.54**
6	Crime	7.89	7.62	73.33	75.00	75.00	**76.92**
7	Security	42.50	40.24	51.83	62.09	61.68	**64.09**
8	Health	48.36	48.79	79.43	86.49	86.49	**86.67**
9	Quality of Life	17.82	21.99	48.89	52.63	52.63	**54.35**
10	Cultural	15.38	15.67	51.22	52.63	52.63	**55.56**
11	Sentiment	15.22	15.72	50.79	53.97	41.03	**54.69**
12	Political	49.06	48.20	50.29	46.99	46.99	**47.23**
13	Policy	39.88	39.39	37.02	42.77	42.77	**43.79**
14	External	12.66	14.22	44.44	66.67	66.67	**71.43**
15	Factual	24.64	19.21	70.95	70.37	70.41	**78.95**
16	Promotion	40.11	46.41	48.16	50.96	50.96	**52.89**
17	Personal	45.36	46.15	59.66	62.99	62.13	**71.20**
	Weighted Avg.	37.14	38.79	53.13	56.49	55.54	**58.66**

as trigrams in an unsupervised setting. Finally, in Model RTs, which adds retweet behaviors, we notice that five of the frames decrease in F_1 score and eleven of the frames have the same score as the previous model. These results suggest that retweet behaviors are not as useful as the follower network relationships in an unsupervised setting. However, this may also be due to fewer retweets present in our dataset, since politicians do not retweet each other as often as the general public.

4.5 Qualitative Analysis

In this section, we explore the ability of our models to locate framing trends which can be used to analyze political discourse on Twitter concerning real world events and voting behaviors. The detailed real world event analysis included here does not appear in our previous publications. Johnson et al. (2017b) presents a high-level overview of the Orlando analysis and Johnson et al. (2017a) discusses the voting trends also covered here.

We first learned the weights of our best performing PSL models using the labeled data and performed MPE inference on the 90,407 remaining *unlabeled* tweets to obtain their predicted frames. We used these predictions to analyze the political discourse on Twitter by focusing on three real world events and the voting behaviors on two issues: the ACA and terrorism.

4.5.1 Framing Trends of Real World Events

In Figures 4.4, 4.6, and 4.8 we show the frame predictions for three events: the shooting at the Pulse Nightclub in Orlando, Florida (June 12, 2016), the shooting at the Inland Regional Center in San Bernadino, California (December 2, 2015), and the shooting at the Emanuel African Methodist Episcopal (EAME) Church in Charleston, South Carolina (June 17, 2015). In each figure the top panel shows the frame trends for Republicans, while the bottom panel shows those of Democrats. For the first two events, we also present a close-up view of the dates of the shootings to better visualize the frame coverage (shown in Figures 4.5 and 4.7).

Orlando. In Figure 4.4 we see gun related tweets from one day before the Orlando shooting and the thirteen days after. There are three interesting peaks of activity: June 12$^{\text{th}}$, June 15$^{\text{th}}$, and June 22$^{\text{nd}}$. We use the predictions of our model to analyze the discourse on Twitter at these dates and connect it with relevant events around those dates.

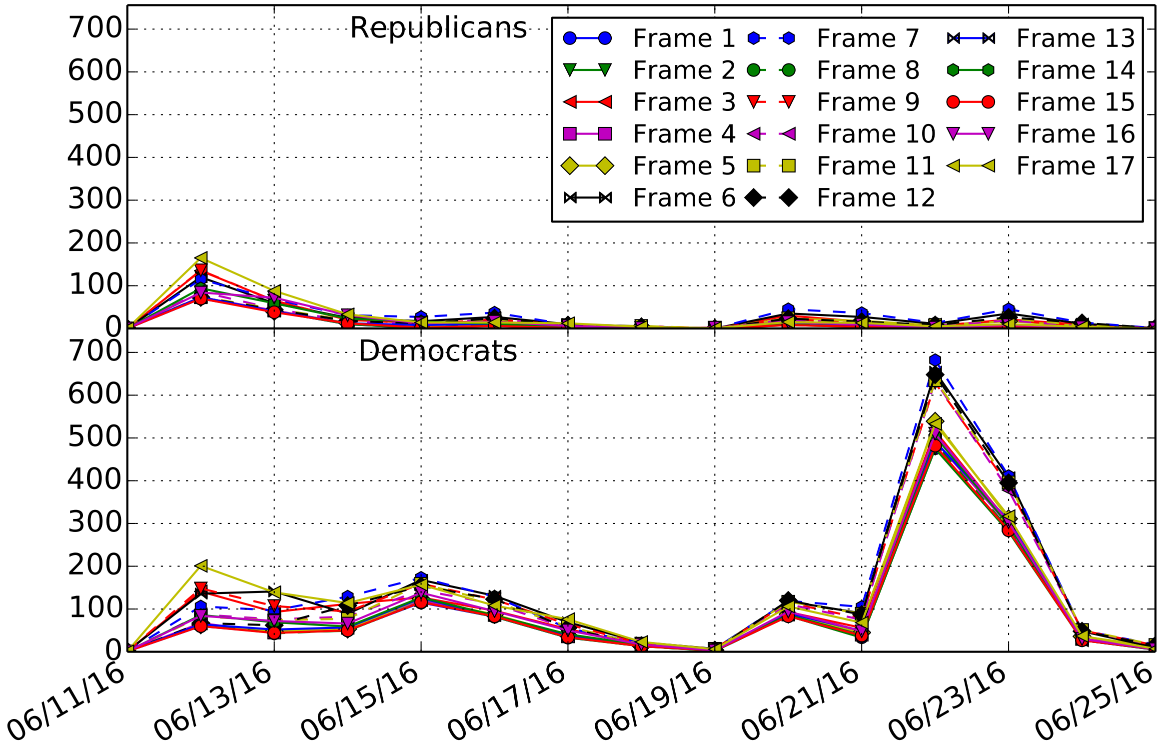


Figure 4.4. PSL Prediction Around Orlando Pulse Nightclub Shooting.

Initial Response (June 12th). The peak on the day of the shooting is shown in more detail in Figure 4.5, which highlights the following top three frames for Republicans and Democrats: 17 (*Personal Sympathy & Support*), 9 (*Quality of Life*), and 3 (*Morality & Ethics*). Frame 17 reflects politicians tweeting that their "thoughts and prayers" are with the community, as seen in the first line of Table 4.17. Offers of prayers and sympathy are used by both parties as the initial response the day both shootings occurred. This can be considered both as a reflection of the politicians' immediate emotional reaction to the shooting but also to support other agendas, as Frame 17 also appears in tweets that use other frames, specifically Frames 9 and 3. Interestingly, Republicans and Democrats use these frames in nuanced ways to promote different agendas.

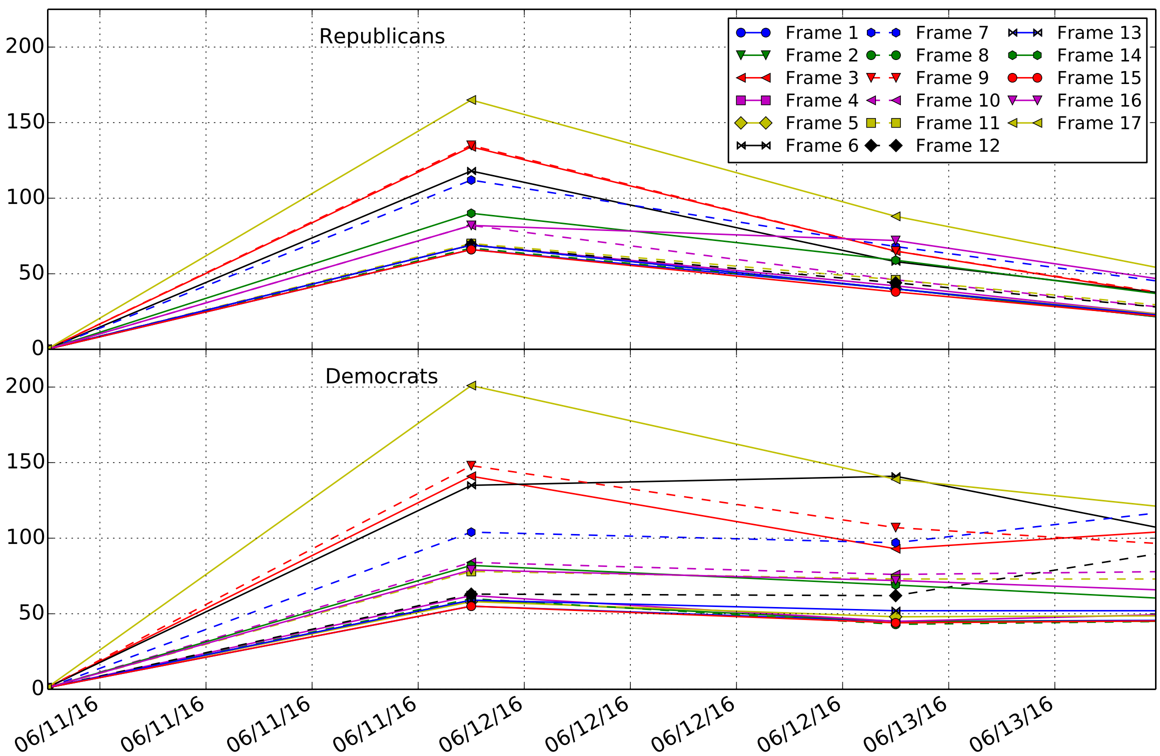


Figure 4.5. Close-up of June 12, 2016 Prediction.

The impacts of the shooting on the quality of life of the community (or nation as a whole) are discussed in tweets having Frame 9. For example, line two of Table 4.17 shows a Democrat tweet calling for action to keep gun violence tragedies from affecting communities. In these tweets, Republicans are more likely to refer to the "Orlando community" while Democrats are more likely to reference the "LGBT community."

Republicans used Frame 3, often in combination with Frame 17, to discuss the shooting as an act of evil or terrorism as well as to suggest links between the shooter and ISIS (examples of these tweets are shown in lines three and four of Table 4.17). Democrats, however, used Frame 3 to express a sense of responsibility on their part to take actions to prevent gun violence (e.g., line five of Table 4.17) or refer to the shooting as a hate crime or act of terror (e.g., line six of Table 4.17).

Political Action (June 15th). Two days after the shooting (June 14$^{\text{th}}$) the discussion on Twitter addressed political action initiated by Democrats. This is shown in the second peak on June 15$^{\text{th}}$, which corresponds to the day Democrats held a filibuster to push for a vote on gun control.

The top frame that day for both parties is Frame 7 (*Security & Defense*), however it is used differently. Democrats frame the need for gun control laws as a preemptive measure that will prevent gun violence (e.g., line seven of Table 4.17). Republicans use Frame 7 to discuss the need to prevent threats posed by ISIS (possibly due to the shooter's association with ISIS) as shown in line eight of Table 4.17. Additionally, some Republicans promote bipartisan efforts to stop the sale of guns to known terrorists (line eight).

The model also shows Democrats using Frame 11 (*Public Sentiment*) among their top frames, which is used to cite the American people's desire for gun control as the motivation for their filibuster (e.g., the last line of Table 4.17).

Bipartisan Political Action (June 22nd). Finally, June 22, 2016 was the day Senators proposed bipartisan political action: a ban for gun sales to people registered on the "no fly" list. Both parties use Frame 7 (*Security & Defense*) as their top frame that day, but the tweets reference defending against general gun violence (Democrats) or terrorist threats (Republicans), reflecting the same pattern seen one week before on June 15$^{\text{th}}$.

Overall, the Democrats have maintained the same level of discussion about the shooting five days after its occurrence. Eventually they move the initial discussion of sympathy towards one which publicizes their concrete political actions, shown in rapidly increasing Twitter activity around June 21$^{\text{st}}$. After the initial response, Republicans appear to become more silent about the issue, based on tweet quantities.

San Bernadino. As a second example, Figure 4.6 shows tweets beginning one day before the San Bernadino, California shooting (December 2, 2015) and up to four

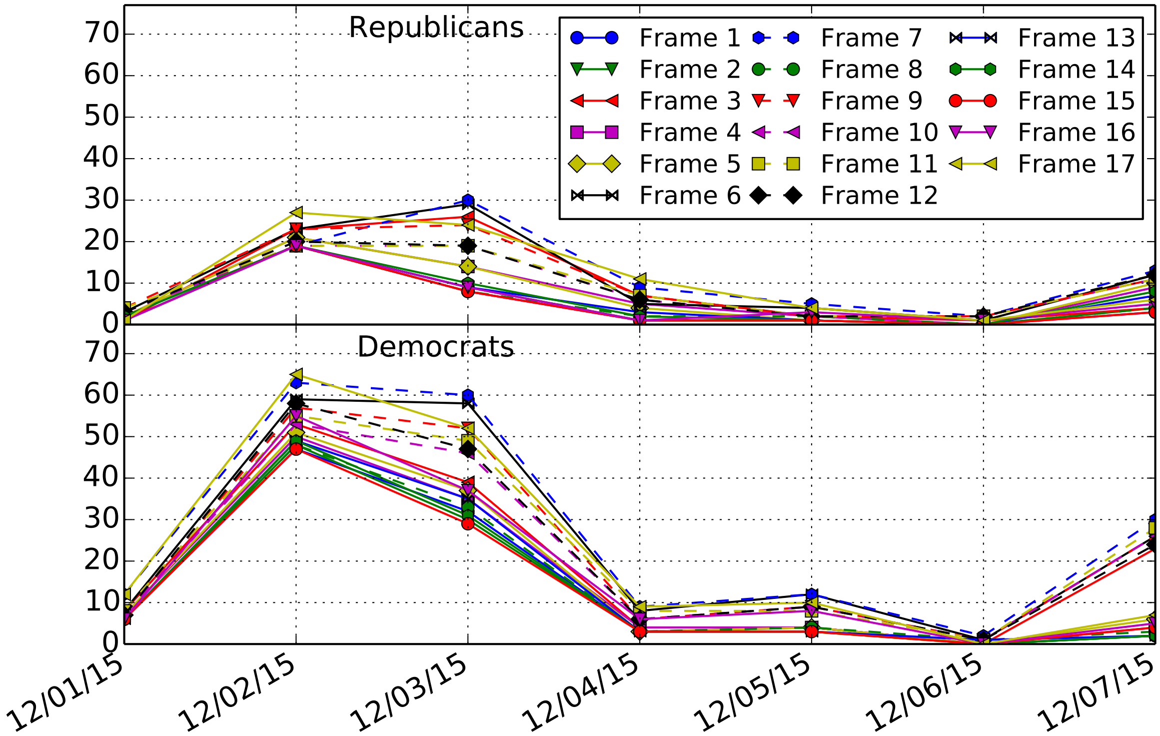


Figure 4.6. PSL Prediction for San Bernadino Shooting.

days after. The peaks for this event (December 2nd and 3rd) are shown in more detail in Figure 4.7.

Republicans have used Frame 17 to express their thoughts and prayers for those affected, as shown in the first line of Table 4.18. Their next top frames are tied and include Frame 3, which also appears the following day and is used to refer to the shooting as religious-motivated terrorism (line four of Table 4.18).

Democrats use Frame 17 similarly to express sympathy the day of the shooting (e.g., line two of Table 4.18), however the usage of this frame drops over the following days. Their second top frame is Frame 7 which presents the need for gun control as a matter of safety (line three of Table 4.18).

Frame 7 becomes the top frame of both Republicans and Democrats on the following day, however it is used differently by the two sides. Democrats use Frame

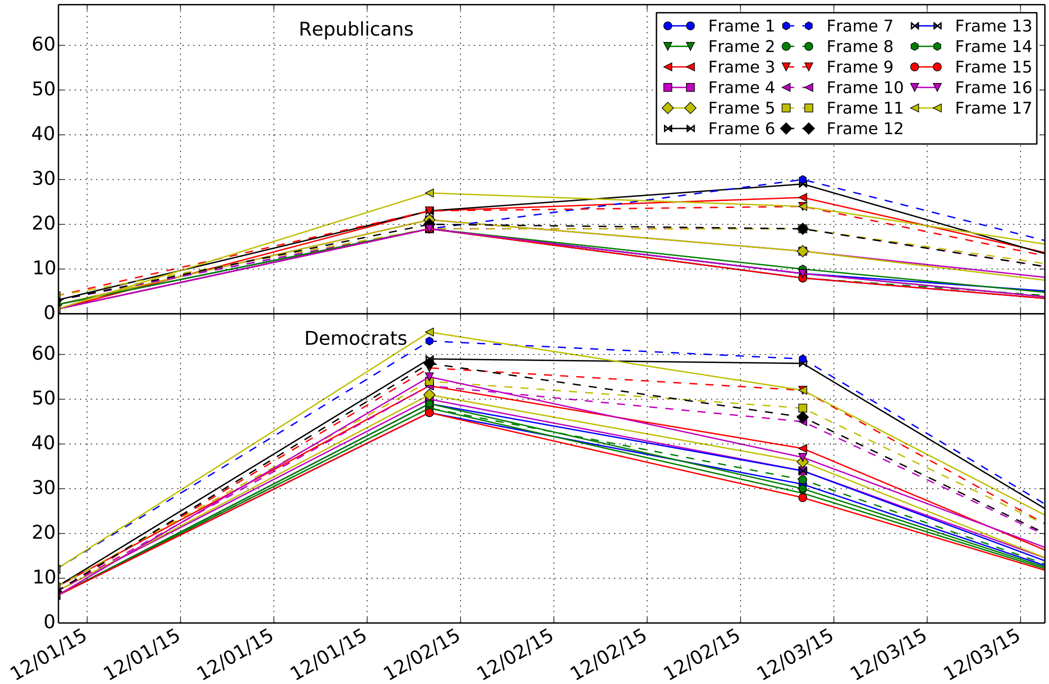


Figure 4.7. Close-up of December 2, 2015 Prediction.

7 to continue their push for gun control legislation, while Republicans use Frame 7 *in response to Democrats*, to express that there is a greater need for action against terrorist threats. The last two lines of Table 4.18 show examples of these tweets.

Charleston. Our final example is the shooting at the Emanuel African Methodist Episcopal Church in Charleston, South Carolina on the evening of June 17, 2015. One key difference between this event and the previous events is that most tweets occurred the *day after the event*, as shown in Figure 4.8. This is likely due to the timing of the shooting, which took place around 9:05 p.m. Similar to the previous events, we see that the top frames include Frames 17, 9, and 3. *Both* Republicans and Democrats use the frames for this event in *similar* ways. Frame 17 is used to convey sympathy for the victims and city, as shown in lines one and two of Table 4.19. Frame 9 is used to express the effects of the shooting on the community (lines three and four in Table 4.19). Frame 6 is more prevalent in the discussion than Frame 7,

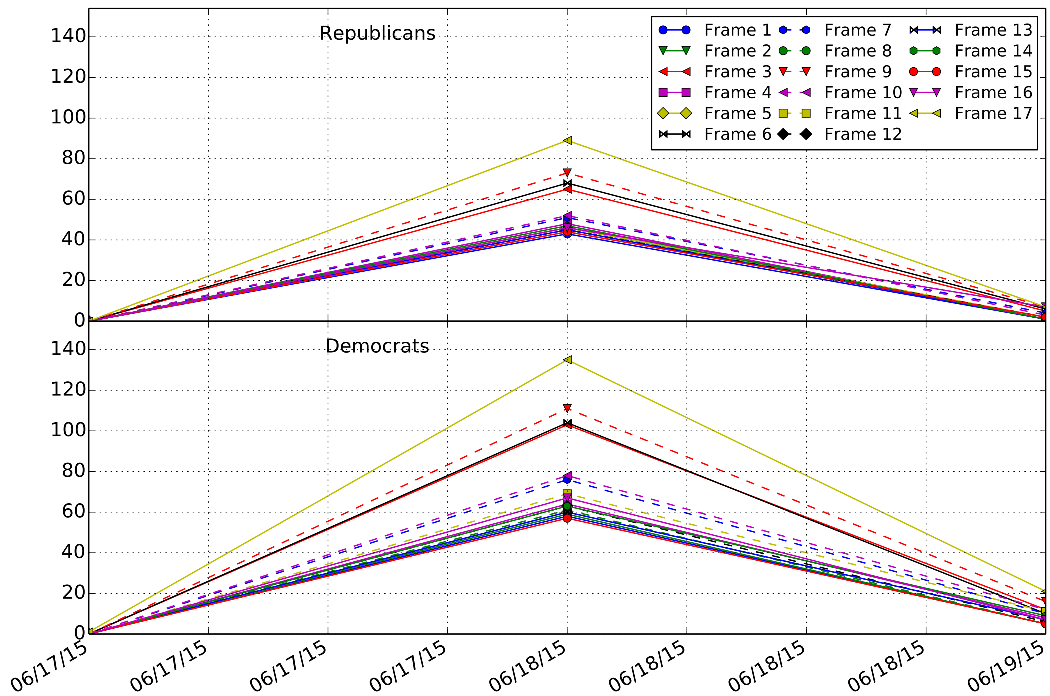
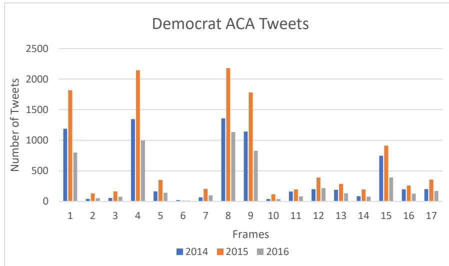


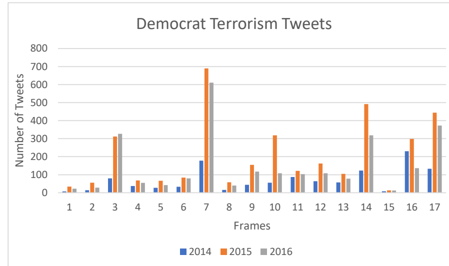
Figure 4.8. PSL Prediction for the Charleston EAME Church Shooting.

while the opposite is true for the other two events, and is used to show hopes for the capture/punishment of the shooter (e.g., lines five and six of Table 4.19). Finally, Frame 3 is used to juxtapose the violence of the shooting with the location in which it occurred: a church, or place of worship.

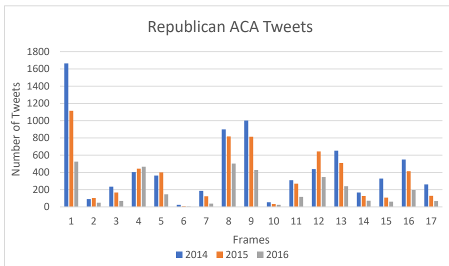
Overall Analysis. In the first two examples, we can see the general trend of politicians tweeting most frequently on the day the event occurs and gradually becoming more silent over time until another event occurs. We also see that for both parties, the initial response is to show sympathy to the victims of the attacks (Frame 17), which declines over the following days or is combined with additional frames. For the San Bernadino shooting, the secondary frames that day are the same frames used the day after the Orlando shooting, indicating similarities in the frames used for gun violence events over time. For all events, the top frames include Frames 17, 9, 3, 7, and 6; however, the frequency of each frame varies across events and days after those



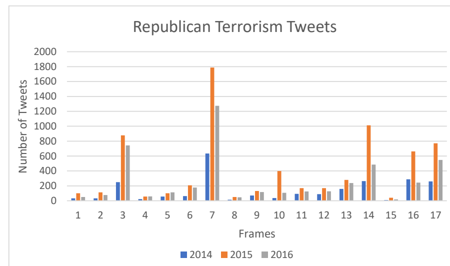
(a) Democrat ACA Frames



(b) Democrat Terrorism Frames



(c) Republican ACA Frames



(d) Republican Terrorism Frames

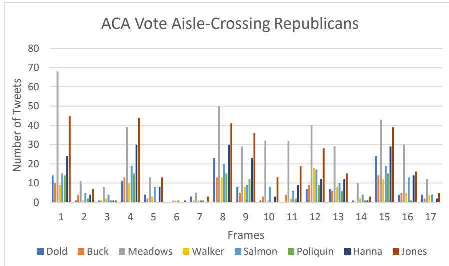
Figure 4.9. Predicted Frames for Tweets from 2014 to 2016 by Party for ACA and Terrorism Issues.

events, possibly due to politicians changing their focus as new information about the situation becomes available or in response to other politicians.

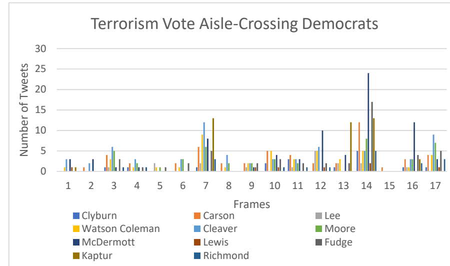
4.5.2 Framing Trends of Voting Behaviors

Figure 4.9 shows the results of our frame analysis for both parties over time for two issues: ACA and terrorism. We highlight these two issues because they are among the most frequently discussed issues in our dataset. To explore how frames can shed light on voting behaviors, we compiled the predicted frames for tweets from 2014 to 2016 for each party. Figure 4.10 presents the results of frame prediction for 2015 tweets of aisle-crossing individual politicians for these two issues.

Party Frames. From Figure 4.9(a) we can see that Democrats mainly use Frames 1, 4, 8, 9, and 15 to discuss ACA, while Figure 4.9(c) shows that Republicans pre-



(a) Aisle-crossing Republicans on ACA Votes.



(b) Aisle-crossing Democrats on Terrorism Votes.

Figure 4.10. Predicted Frames for Tweets of Aisle-crossing Politicians in 2015.

dominantly use Frames 1, 8, 9, 12, and 13. Though the parties use similar frames, they are used to express different agendas. For example, Democrats use Frame 8 to indicate the positive effect that the ACA has had in granting more Americans health care access. Republicans, however, use Frame 8 (and Frame 13) to indicate their party's agenda to replace the ACA with access to different options for health care. Additionally, Democrats use the *Fairness & Equality* frame (Frame 4) to convey that the ACA gives minority groups a better chance at accessing health care. They also use Frame 15 to express statistics about enrollment of Americans under the ACA. Finally, Republicans use Frames 12 and 13 to bring attention to their own party's actions to "repeal and replace" the ACA with different policies.

Figures 4.9(b) and 4.9(d) show the party-based framing patterns over time for terrorism related tweets. For this issue both parties use similar frames: 3, 7, 10, 14, 16, and 17, but to express different views. For example, Democrats use Frame 3 to indicate a moral responsibility to fight ISIS. Republicans use Frame 3 to frame terrorists or their attacks as a result of "radical Islam". An interesting pattern to note is seen in Frames 10 and 14 for both parties. In 2015 there is a large increase in the usage of this frame. This seems to indicate that parties possibly adopt new frames *simultaneously or in response to the opposing party*, perhaps in an effort to be in control of the way the message is delivered through that frame.

Individual Frames. In addition to entire party analysis, we were interested in seeing if frames could shed light on the behavior of *aisle-crossing* politicians. These are politicians who do not vote the same as the majority vote of their party (i.e., they vote the same as the opposing party). Identifying such politicians can be useful in governments which are heavily split by party, for example, governments such as the recent U.S. Congress (2015 to 2017) where politicians tend to vote the same as the rest of their party members. For this analysis, we collected five 2015 votes from the House of Representatives on both issues and compiled a list of the politicians who voted opposite to their party. The most important descriptor we noticed was that all aisle-crossing politicians *tweet less frequently on the issue* than their fellow party members. This is true for both parties. This behavior could indicate lack of desire to draw attention to one's stance on the particular issue.

Figure 4.10(a) shows the framing patterns of aisle-crossing Republicans on ACA votes from 2015. Recall from Figure 4.9 that Democrats mostly use Frames 1, 4, 8, 9, and 15, while Republicans mainly use Frames 1, 8, and 9. In this example, these Republicans are considered aisle-crossing votes because they have voted the same as Democrats on this issue. The most interesting pattern to note here is that these Republicans use the same framing patterns as the Republicans (Frames 1, 8, and 9), but they also use the frames that are *unique to Democrats*: Frames 4 and 15. These latter two frames appear significantly less in the Republican tweets of our entire dataset as well. These results suggest that to predict aisle-crossing Republicans it would be useful to check for usage of typically Democrat-associated frames, especially if those frames are infrequently used by Republicans.

Figure 4.10(b) shows the predicted frames for aisle-crossing Democrats on terrorism-related votes. We see here that there are very few tweets from these Democrats on this issue and that overall they use the same framing patterns as seen previously: Frames 3, 7, 10, 14, 16, and 17. However, given the small scale of these tweets, we can also consider Frames 12 and 13 to show peaks for this example. This suggests that for aisle-crossing Democrats the use of additional frames not often used by their party for discussing an issue might indicate potentially different voting behaviors.

4.6 Chapter Summary

In this chapter we presented models for solving the problem of issue-independent framing analysis of U.S. politicians on Twitter. We have proposed three new Twitter-specific frames and have compiled and annotated a Congressional Tweets Dataset for use by the NLP community. We provide weakly supervised models to extract and format tweet information which is then used as input to both non-collective and collective global PSL models. We show that by incorporating Twitter behaviors, such as similar activity times and similar networks, we can increase F_1 score prediction. We provide an analysis of our approach in both supervised and unsupervised settings, as well as a real world analysis of framing trends over time. These models serve as an interesting exploratory tool to study the changing trends in framing patterns of political discourse on Twitter and their ramifications in the real world.

Table 4.17.
Example Tweets Associated With the Orlando Pulse Nightclub Shooting.

Date	Politician	Political Party	Tweet	Predicted Frame(s)
6/12/2016	Alex Mooney	Republican	My thoughts and prayers are with the people of #Orlando, the victims, and their families.	17
6/12/2016	Brad Ashford	Democrat	As authorities investigate the Orlando shooting, we must pray for the victims and act swiftly to keep these tragedies out of our communities.	9
6/12/2016	Lisa Murkowski	Republican	What happened in Orlando was an absolute tragic act of terrorism spawned by an ideology of hate being pushed by ISIS.	3
6/12/2016	Bob Goodlatte	Republican	The attack in #Orlando was an act of pure evil. My prayers are w/ the families of victims and the injured. We will continue seeking answers.	3, 17
6/12/2016	David Cicilline	Democrat	Voters should absolutely hold us accountable for what we're doing or not doing to address gun violence.	3
6/12/2016	Yvette Clark	Democrat	I am deeply saddened by the act of hate and terror enacted on the lives of Orlando's LGBT Community and I #StandWithOrlando	3, 17
6/15/2016	Jeanne Shaheen	Democrat	Joining @ChrisMurphyCT on the Senate floor to say #Enough and call for reforms 2 prevent gun violence.	7, 12
6/15/2016	Mark Kirk	Republican	Americans need to know Washington is listening - We must keep guns out of the hands of suspected terrorists.	7
6/15/2016	Kirsten Gillibrand	Democrat	As we mourn victims of yet another tragedy, time to finally act on commonsense gun safety reforms supported by the American people.	11, 12

Table 4.18.
Example Tweets Associated With the Inland Regional Center of San Bernadino Shooting on December 2, 2015.

Date	Politician	Political Party	Tweet	Predicted Frame(s)
12/2/2015	Rand Paul	Republican	My thoughts and prayers are with the victims, families, and brave first responders during this unspeakable tragedy.	17
12/2/2015	Jeff Merkley	Democrat	Heartbroken about yet another horrific shooting My thoughts and prayers are with the #SanBernadino community	17
12/2/2015	Harry Reid	Democrat	Gun violence has become a cancer on this nation. We must make common sense gun reforms that keep weapons out of dangerous hands.	7
12/3/2015	Jeff Duncan	Republican	Wish Obama and Clinton would speak out about terrorism and radical Islamic jihad as quickly as they call for gun control here in America	3, 12
12/3/2015	Adam Schiff	Democrat	Must require background check for every gun sale, make easier to preclude those w serious mental health probs from gaining access weapons.	7
12/3/2015	Mike Kelly	Republican	Mr. President: Instead of climate control or gun control, we need terror control and serious American leadership for a world out of control.	7, 12

Table 4.19.
Example Tweets Associated With the EAME Church of Charleston, South Carolina Shooting on June 17, 2015.

Date	Politician	Political Party	Tweet	Predicted Frame(s)
6/18/2015	Garret Graves	Republican	Our thoughts and prayers are with the city of Charleston this morning. #CharlestonShooting	17
6/18/2015	Raul Ruiz	Democrat	I am deeply saddened to hear of the tragic shooting in #Charleston. My family and I are keeping all those affected in our thoughts and prayers.	17
6/18/2015	Mick Mulvaney	Republican	The shooting in Charleston that took the lives of nine people has shaken that community and the state as a whole.	9
6/18/2015	Seth Moulton	Democrat	My thoughts and prayers are with Charleston, SC today. We must work to end these senseless acts of violence in our communities.	9, 17
6/18/2015	Pete King	Republican	Our thoughts and prayers must be with 9 innocent men and women murdered in Charleston, SC. Every effort must be made to capture the killer. RIP	6, 17
6/18/2015	Terri Sewell	Democrat	My thoughts are with those impacted by the #CharlestonShooting. I pray that the perpetrator is brought to justice soon.	6
6/18/2015	Patrick McHenry	Republican	My thoughts and prayers are with the #Charleston community this morning. Horrific to see this violence anywhere, much less a house of worship.	3, 17
6/18/2015	Stephen Lynch	Democrat	My thoughts and prayers go out to the victims and their families in #Charleston. A church is a house of the Lord, a place of peace.	3, 17

5 MORAL FOUNDATIONS CLASSIFICATION

By using social media platforms politicians are able to express their stances on issues and by selectively using certain political slogans, reveal their underlying political ideologies and moral views on an issue. Previous works in political and social science have shown a correlation between political ideology, political stances, and the moral convictions used to justify these stances (Graham et al., 2009). For example, consider the following tweet by a prominent member of the U.S. Congress:

> We are permitting the incarceration and shooting of thousands of black and brown boys in their formative years.

Figure 5.1. Example Tweet Highlighting Classification Difficulty.

The text expresses concern about the fate of young individuals (i.e., *incarceration, shooting*), specifically for vulnerable members of minority groups. The Moral Foundations Theory (MFT) (Haidt and Joseph, 2004; Haidt and Graham, 2007) provides a theoretical framework for explaining these nuanced distinctions. The theory suggests that there are five basic moral values which underlie human moral perspectives, emerging from evolutionary, social, and cultural origins. These are referred to as the Moral Foundations (MF), and include *care/harm, fairness/cheating, loyalty/betrayal, authority/subversion*, and *purity/degradation* (Table 5.1 provides a more detailed explanation). The above example reflects the moral foundations that shape the author's perspective on the issue, which in this case are *care/harm* and *fairness/cheating*.

Traditionally, analyzing text based on MFT relied on a lexical resource, the Moral Foundations Dictionary (MFD) (Haidt and Graham, 2007; Graham et al., 2009). The MFD, similar to LIWC (Pennebaker et al., 2001; Tausczik and Pennebaker, 2010), associates a list of related words with each one of the moral foundations. Analyzing text amounts to counting the number of occurrences of words related to each one of the moral foundations. Given the highly abstract and generalized nature of the moral foundations, this approach often falls short of dealing with the highly ambiguous text politicians use to express their perspectives on specific issues. Consider the following tweet, by another prominent member of the U.S. Congress. The tweet reflects the author's use of both the *care/harm* and *fairness/cheating* moral foundations.

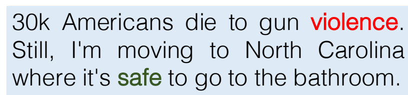


Figure 5.2. Example Tweet Highlighting Classification Difficulty.

While the first foundation can be directly identified using word choice, the second requires first identifying the sarcastic expression referring to LGBTQ rights and then using extensive world knowledge, that the tweet refers to intended legislation about transgender bathroom access restrictions, to determine the appropriate moral foundation.

In this chapter, we aim to solve this challenge by suggesting a data-driven approach to moral foundation identification in text. Previous work (Garten et al., 2016) has looked at classification-based approaches over tweets specifically related to Hurricane Sandy, augmenting the textual content with background knowledge using entity linking (Lin et al., 2017). Different from this and similar works, we look at tweets extracted from U.S. politicians over several years, discussing a large number of events, and touching on several different political issues. Our approach is guided by the intuition that the abstract moral foundations will manifest differently in text, depending on the specific characteristics of the events discussed in the tweet. As a result, it is necessary to correctly model the relevant contextualizing information.

Specifically, we are interested in exploring how political ideology, language, and framing interact to represent morality on Twitter. We examine the interplay of political slogans (for example *"repeal and replace"* when referring to the Affordable Care Act), and policy framing techniques (Boydstun et al., 2014; Johnson et al., 2017a) as features for predicting the underlying moral values which are expressed in politicians' tweets. In addition, we identify high-level phrases characterizing the main point of the tweet, which allows the model to identify the author's perspective on specific issues and generalize over the specific wording used (for example, if the tweet mentions `Religion` or `Political Maneuvering`).

This information is incorporated into global probabilistic models using Probabilistic Soft Logic (PSL), a graphical probabilistic modeling framework (Bach et al., 2013). PSL specifies high level rules over a relational representation of these features, which are compiled into a graphical model called a hinge-loss Markov random field that is used to make the final moral foundation prediction. Our experiments show the importance of modeling contextualizing information, leading to very significant improvements over dictionary driven approaches and purely lexical methods.

In summary, this chapter makes the following contributions: (1) This chapter is among the first to explore jointly modeling language and political framing techniques, as well as social and behavioral information, for the classification of moral foundations used by U.S. politicians on Twitter. (2) We provide a description of our annotation guidelines and an annotated dataset of 2,050 tweets[1]. (3) We suggest easily-adaptable computational models for classifying the moral foundations present in tweets across a variety of policy issues.

5.1 Moral Foundations Theory

The Moral Foundations Theory (Haidt and Graham, 2007) was proposed by sociologists and psychologists as a way to understand how morality develops, as well as its similarities and differences across cultures. The theory consists of the five moral foundations shown in Table 5.1. The goal of this work is to classify the moral foundation implied in the tweets of the Congressional Tweets Dataset (Johnson et al., 2017a).

[1]The data is available at: `purduenlp.cs.purdue.edu/projects/twittermorals`.

Table 5.1.
Brief Descriptions of Moral Foundations.

Moral Foundation and Brief Description
1. Care/Harm: Care for others, generosity, compassion, ability to feel pain of others, sensitivity to suffering of others, prohibiting actions that harm others.
2. Fairness/Cheating: Fairness, justice, reciprocity, reciprocal altruism, rights, autonomy, equality, proportionality, prohibiting cheating.
3. Loyalty/Betrayal: Group affiliation and solidarity, virtues of patriotism, self-sacrifice for the group, prohibiting betrayal of one's group.
4. Authority/Subversion: Fulfilling social roles, submitting to authority, respect for social hierarchy, leadership, fellowship, respect for traditions, prohibiting rebellion against authority.
5. Purity/Degradation: Associations with the sacred and holy, disgust, contamination, underlies religious notions of striving to live in an elevated way, prohibiting violated the sacred.
6. Non-moral: Does not fall under any of the other foundations.

5.2 Dataset Annotation

5.2.1 Congressional Tweets Dataset

We first attempted to use Amazon Mechanical Turk for annotation, but found that most Mechanical Turkers would choose the Care/Harm or Fairness/Cheating label a majority of the time. Additionally, annotators preferred choosing first the foundation branch (i.e., Care/Harm) and then its sentiment (positive or negative) as opposed to the choice of each foundation separately, i.e., given the choice between Harm or Care/Harm-Negative, annotators preferred the latter. Based on these observations, two annotators, one liberal and one conservative, manually annotated a subset of tweets, agreed on general guidelines, and then labeled the remaining tweets of the dataset. The overall distribution, distributions by political party, and distributions per issue of the labeled dataset are presented in Table 5.2.

Table 5.2.
Distributions of Moral Foundations. All is across the entire dataset. Party is the Republican (Rep) or Democrat (Dem) specific distributions. Issue lists the six issue-specific distributions (Abortion, ACA, Guns, Immigration, LGBTQ, Terrorism).

Morals	All	Party		Issue					
		Rep	Dem	Abo	ACA	Gun	Imm	LGBTQ	Ter
Care	524	156	368	37	123	215	33	34	113
Harm	355	151	204	26	64	141	19	34	101
Fairness	268	55	213	41	81	19	11	86	39
Cheating	82	37	45	14	27	11	10	9	13
Loyalty	303	63	240	28	29	128	36	38	58
Betrayal	53	25	28	10	4	9	6	3	22
Authority	192	62	130	24	44	50	38	10	34
Subversion	419	251	168	34	169	75	73	25	60
Purity	174	86	88	24	3	102	5	24	41
Degradation	66	34	32	5	0	31	0	4	31
Non-moral	334	198	136	17	143	28	47	7	96

Labeling tweets presents several challenges. First, tweets are short and thus lack the context that is often necessary for choosing a moral viewpoint. Tweets are often ambiguous, e.g., a tweet may express care for people who are being harmed by a policy. One major challenge was overcoming the political bias of the annotator. For example, if a tweet discusses opposing Planned Parenthood because it provides abortion services, the liberal annotator typically viewed this as Harm (i.e., taking services away from women and thus hurting them), while the conservative annotator tended to view this as Purity (i.e., all life is sacred). To overcome this bias, annotators were given the political party of the politician who wrote the tweets and instructed to choose the moral foundation *from the politician's perspective*. Finally, as noted in the previous chapter, tweets present a compound problem: tweets often present two thoughts, some of which can even be contradictory. This results in one tweet having multiple moral foundations or even two opposing foundations. Annotators chose a primary moral foundation whenever possible, but were allowed a secondary foundation if the tweet presented two differing thoughts.

Several recurring themes continued to appear throughout the dataset including "thoughts and prayers" for victims of gun shooting events or rhetoric against the opposing political party. The annotators agreed on the following general guidelines for these repeating topics: (1) The Purity label is used for tweets that relate to prayers or the fight against ISIL/ISIS. (2) Loyalty is for tweets that discuss "stand(ing) with" others, American values, American troops or allies, or reference a demographic that the politician belongs to, e.g. if the politician tweeting is a woman and she discusses women-related issues. (3) At the time the dataset was collected, the President was Barack Obama and the Republican party controlled Congress. Therefore, any tweets specifically attacking Obama or tweets against the controlling party were labeled as Subversion. (4) Tweets discussing health or welfare were labeled as Care. (5) Tweets which discussed limiting or restricting laws or rights were labeled as Cheating. (6) Sarcastic attacks, typically against the opposing political party, were labeled as Degradation.

5.2.2 Senate Tweets 2016

In addition to the Congressional Tweets Dataset, we also compiled two smaller datasets for use in qualitative analysis. Using a combination of web scraping and the Twitter API, we collected the available tweets of all Senators during the year 2016. This approach allows users to overcome the recovery limit of the Twitter API by scraping for available tweet IDs, while still adhering to the terms of service, i.e., if a politician deletes a tweet, we are *unable* to recover it.

5.2.3 CongressTweets

CongressTweets is a collection of the tweets of all congressional members in 2018 [2]. To facilitate comparison with the Senate Tweets 2016 dataset, we only used the tweets of senators from this collection. This dataset and the Senate Tweets 2016 dataset are used in the qualitative application of the models for the analysis of real world political behavior.

5.3 Language-based Models

For this work, we designed extraction models and PSL models that were capable of adapting to the dynamic language used on Twitter and predicting the moral foundation of a given tweet. Our approach uses weakly supervised extraction models, whose only initial supervision is a set of unigrams and the political party of the tweet's author, to extract features for each PSL model. These features are represented as PSL predicates and combined into the probabilistic rules of each model, as shown in Table 5.3, which successively build upon the rules of the previous model.

For each aspect of information that composes the PSL models, scripts are written to first identify and then extract the correct information from the tweets. Once extracted, this information is formatted into PSL predicate notation and input to the PSL models. Table 5.3 presents the information that composes each PSL model, as well as an example of how PSL rules in this model appear.

[2] The dataset is available for download at: https://github.com/alexlitel/congresstweets/tree/master/data.

Table 5.3.
Examples of Moral PSL Model Rules Using Gold Standard Frames. Each model builds successively on the rules of the previous model.

Model No.	Information Used	Example of PSL Rule
M1	Unigrams (MFD or AR)	Unigram_M(T, U) → Moral(T, M)
M2	Model 1 + Party	Unigram_M(T, U) ∧ Party(T, P) → Moral(T, M)
M3	Model 2 + Issue	Unigram_M(T, U) ∧Party(T, P) ∧ Issue(T, I) → Moral(T, M)
M4	Model 3 + Phrase	Unigram_M(T, U) ∧ Party(T, P) ∧Phrase(T, PH) →Moral(T, M)
M5	Model 4 + Frame	Unigram_M(T, U) ∧ Phrase(T, PH) ∧ Frame(T, F) → Moral(T, M)
M6	Model 5 + Party-Bigrams	Unigram_M(T, U) ∧ Party(T, P) ∧ Bigram_P(T, B) → (Moral(T, M)
M7	Model 6 + Party-Issue-Bigrams	Unigram_M(T, U) ∧ Party(T, P) ∧ Bigram_{PI}(T, B) → Moral(T, M)
M8	Model 7 + Phrase	Bigram_{PI}(T, B) ∧ Phrase(T, PH) → Moral(T, M)
M9	Model 8 + Frame	Bigram_{PI}(T, B) ∧ Frame(T, F) → Moral(T, M)
M10	Model 9 + Party-Trigrams	Unigram_M(T, U) ∧Party(T, P) ∧Trigram_P(T, TG) →Moral(T, M)
M11	Model 10 + Party-Issue-Trigrams	Unigram_M(T, U) ∧ Party(T, P) ∧ Trigram_{PI}(T, TG) → Moral(T, M)
M12	Model 11 + Phrase	Trigram_{PI}(T, TG) ∧ Phrase(T, PH) → Moral(T, M)
M13	Model 12 + Frame	Trigram_{PI}(T, TG) ∧ Frame(T, F) → Moral(T, M)

Table 5.4.
Examples of Joint Moral and Frame PSL Model Rules. For these models, the FRAME predicate is *not* initialized with known values, but is predicted jointly with the Moral predicate.

M2: UNIGRAMS + PARTY
$\text{UNIGRAM}_M(\text{T, U}) \wedge \text{PARTY}(\text{T, P}) \wedge \text{FRAME}(\text{T, F}) \rightarrow \text{MORAL}(\text{T, M})$
$\text{UNIGRAM}_M(\text{T, U}) \wedge \text{PARTY}(\text{T, P}) \wedge \text{MORAL}(\text{T, M}) \rightarrow \text{FRAME}(\text{T, F})$

M13: ALL FEATURES
$\text{TRIGRAM}_{PI}(\text{T, TG}) \wedge \text{PHRASE}(\text{T, PH}) \wedge \text{FRAME}(\text{T, F}) \rightarrow \text{MORAL}(\text{T, M})$
$\text{TRIGRAM}_{PI}(\text{TG, B}) \wedge \text{UNIGRAM}_M(\text{T, U}) \wedge \text{MORAL}(\text{T, M}) \rightarrow \text{FRAME}(\text{T, F})$

5.3.1 Unigrams

Works studying the Moral Foundations Theory typically assign a foundation to a body of text based on a majority match of the words in the text to the Moral Foundations Dictionary (MFD), a predefined list of unigrams associated with each foundation. These unigrams capture the conceptual idea behind each foundation. However, annotators noted that when choosing a foundation they typically used a small phrase or the entire tweet, not a single unigram. Based on this, we compiled all of the annotators' phrases per foundation into a unique set to create a new list of unigrams for each foundation. These unigrams are referred to as "Annotator's Rationale (AR)" throughout the remainder of this dissertation. The PSL predicate $\text{UNIGRAM}_M(\text{T, U})$ is used to input any unigram U from tweet T that matches the MFD or AR lists of unigrams, M, into the PSL models. An example of a rule using this predicate can be seen in the first row of Table 5.3.

During annotation, we observed that often a tweet has only one match to a unigram, if any, and therefore a majority count approach tends to fail. Further, as shown in Figure 5.2, many tweets have one unigram that matches one foundation and another unigram that matches a different foundation. In such cases, the correct foundation cannot be determined from unigrams alone. Based on these observations and the annotators' preference for using phrases, we incorporate the most frequent bigrams and trigrams for each political party (BIGRAM$_P$(T, B) and TRIGRAM$_P$(T, TG)) and for each party on each issue (BIGRAM$_{PI}$(T, B) and TRIGRAM$_{PI}$(T, TG)). As shown in Johnson et al. (2017a), these top 20 bigrams and trigrams produce a more accurate prediction than unigrams alone.

5.3.2 Ideological Information

Previous works have shown a strong correlation between ideology and the moral foundations (Haidt and Graham, 2007), as well as between ideology and policy issues (Boydstun et al., 2014). Annotators were able to agree on labels when instructed to label from the ideological point of view of the tweet's author, even if it opposed their own views. Based on these positive correlations, we incorporate both the issue of the tweet (ISSUE(T, I)) and the political party of the author of the tweet (PARTY(T, P)) into our PSL models. Examples of how this information is represented in the PSL models are shown in rows two and three of Table 5.3.

5.3.3 Abstract Phrases

As described above, annotators reported that phrases were more useful than unigrams in determining the moral foundation of the tweet. This is likely due to the observation that politicians are known for repeating certain slogans in both their tweets and speeches. These key phrases *indirectly* indicate the political figures' core beliefs and ideological stances, two aspects which are intertwined with morality. Identification of these phrases automatically decomposes the framing strategy of a tweet into more specific categories, and can be used to disambiguate predictions in which tweets expressed with different moralities fall under the same framing strategy.

Consider the following two tweets:

1. *POTUS exec. order on guns is a gross overreach of power that tramples on the rights of law abiding Americans and our Constitution*
2. *With this ruling #SCOTUS has upheld a critical freedom for women to make their own decisions about their bodies*

In the first tweet, several phrases indicate the frame: "exec. order", "overreach of power", "rights of law abiding Americans", "our constitution". In the second tweet, the relevant phrases are "this ruling" and "upheld a critical freedom". All of these phrases indicate that the same frame is being used in both tweets. However, analyzing the specific terminology in each case and the context in which it appears helps capture the moral foundations underlying ideological similarities and differences. For example, in the context of gun-rights debates, phrases highlighting "law and order" and references to the constitution tend to reflect a conservative ideology and authority moral foundation, while phrases highlighting women's ability to choose in the abortion debate tend to reflect a liberal ideology and fairness moral foundation.

However, due to the dynamic nature of language and trending issues on Twitter, it is impracticable to construct a list of all possible slogans or phrases one can expect to appear in tweets. These phrases must be *abstracted* into higher-level phrases that are more stable over time and thus easier to identify and extract.

For example, a tweet discussing "President Obama's signing a bill", has two possible concrete phrases: President Obama's signing and signing a bill. Each phrase falls under two possible abstractions: political maneuvering (Obama's actions) and mentions legislation (signing of a bill). In this paper we use the following high-level abstractions: `legislation or voting`, `rights and equality`, `emotion`, `sources of danger or harm`, `positive benefits or effects`, `solidarity`, `political maneuvering`, `protection and prevention`, `American values or traditions`, `religion`, and `promotion`. For example if a tweet mentions "civil rights" or "equal pay", then these phrases indicate that the `rights and equality` abstraction is being used to express morality. Some of these abstractions correlate with the corresponding moral foundation or frame, e.g. the `religion` abstraction is highly correlated with the Purity foundation.

To match phrases in tweets to these abstractions, we use the embedding-based model of Lee et al. (2017). This phrase similarity model was trained on the Paraphrase Database (PPDB) (Ganitkevitch et al., 2013) and incorporates a Convolutional Neural Network (CNN) to capture sentence structures. This model generates the embeddings of our abstract phrases and computes the cosine similarities between phrases and tweets as the scores. The input tweets and phrases are represented as the average word embeddings in the input layer, which are then projected into a convolutional layer, a max-pooling layer, and finally two fully-connected layers. The embeddings are thus represented in the final layer. The learning objective of this model is:

$$\begin{aligned}\min_{W_c,W_w}\Big(\sum_{<x_1,x_2>\in X} & max(0,\delta - cos(g(x_1),g(x_2)) \qquad +cos(g(x_1),g(t_1))) \\ & +max(0,\delta - cos(g(x_1),g(x_2))) \qquad +cos(g(x_2),g(t_2))\Big) \\ & +\lambda_c||W_c||^2 + \lambda_w||W_{init} - W_w||^2,\end{aligned}$$

where X is all the positive input pairs, δ is the margin, $g(\cdot)$ represents the network, λ_c and λ_w are the weights for L2-regularization, W_c is the network parameters, W_w is the word embeddings, W_{init} is the initial word embeddings, and t_1 and t_2 are negative examples that are randomly selected.

All tweet-phrase pairs with a cosine similarity over a given threshold are used as input to the PSL model via the predicate PHRASE(T, PH), which indicates that tweet T contains a phrase that is similar to an abstracted phrase (PH). Rows four, seven, and eleven of Table 5.3 show examples of the phrase rules as used in our modeling procedure.

5.3.4 Nuanced Framing

Framing is a political strategy in which politicians carefully word their statements in order to bias public opinion towards their stance on an issue. This technique is a fine-grained view of how issues are expressed. Frames are associated with both issue, political party, and ideologies. For example, if a politician emphasizes the economic burden a new bill would place on the public, then they are using the *Economic* frame. Different from this, if they emphasize how people's lives will improve because of this bill, then they are using the *Quality of Life* frame.

In this chapter, I study frames in two settings: where the frames are known and when they are unknown in a joint prediction with the moral foundations. Using the Congressional Tweets Dataset as the true labels for 17 policy frames, this information is input to PSL using the FRAME(T, F) predicate as shown in Table 5.3. Conversely, the same predicate can be used as a joint prediction target, with no initialization, as shown in Table 5.4.

5.4 Social and Behavior-based Models

The LANGUAGE model shown in Table 5.5 consists of language-based features only. These include the unigrams based on the Moral Foundations Dictionary, *political slogans* represented by bigrams and trigrams associated with each party for each issue, ideological phrase indicators (abstractions of the slogans), and frames. Each feature is described in detail in Section 5.3.

The first row of Table 5.3 shows the use of unigram indicators from the Moral Foundations Dictionary (MFD_M(T, U)) and ideological phrases (PHRASE(T1, S)). For example, the predicate MFD_M(T, U) indicates that this tweet T has unigram U from the Moral Foundations Dictionary (MFD) list of unigrams for an expected Moral Foundation M. The rule in this first row would therefore read as: if tweet T has unigram U from the MFD list for moral M and has slogan S that belongs to a group of phrases, then we expect moral M is implied in tweet T.

Table 5.5.
Examples of PSL Model Rules. Each row shows an example of how the model combines rules from previous models to build an increasingly comprehensive model. The LANGUAGE model uses language only features. +RETWEETS adds retweet information to the language rules. Similarly, +FOLLOWING adds social network information and +TEMPORAL adds time patterns.

PSL MODEL	EXAMPLE OF PSL RULE
LANGUAGE	MFD_M(T, U) $\wedge$ PHRASE(T1, S) $\rightarrow$ MORAL(T, M)
+RETWEETS	RETWEETS(T1, T2) $\wedge$ MORAL(T1, M) $\rightarrow$ MORAL(T2, M)
+FOLLOWING	FOLLOWS(T1, T2) $\wedge$ MORAL(T1, M) $\rightarrow$ MORAL(T2, M)
+TEMPORAL	TEMPORAL(T1, T2) $\wedge$ FOLLOWS(T1, T2) $\rightarrow$ MORAL(T1, M)

The next model, RETWEETS, builds upon the language-based baseline by adding retweet information into the prediction. Retweets are useful because they are both textual indicators and miniature representations of the network structure inherent in the political sphere of Twitter. This feature is therefore able to simultaneously capture both the impact of language and social connections.

The FOLLOWING model takes this one step further and incorporates the actual social network into the PSL model. This predicate, FOLLOWS(T1, T2), indicates that the author of tweet T1 follows the author of tweet T2. Since politicians are likely to follow other politicians or Twitter accounts that share similar ideologies and ideology has been shown to be associated with moral foundations, this PSL model can exploit the social network relationships of politicians to detect similar moral foundations patterns.

Lastly, the TEMPORAL PSL model adds information about similar time activity between tweets. Rules in this model indicate if tweets occur within the same time frame as one another. For this work, a time window of one day was used. This feature is motivated by the observation that most politicians tweet about an event on the day it occurs, and discussion of the event declines over time. Therefore, if two politicians share similar moral viewpoints, we expect them to use the same moral foundations to discuss an event at the same time.

5.5 Experimental Results

5.5.1 Analysis of Supervised Experiments for Language-based Models

We conducted supervised experiments using five-fold cross validation with randomly chosen splits on the labeled portion of the dataset. Table 5.8 shows an overview of the average results of our supervised experiments for five of the PSL models. The first column lists the PSL model. The second column presents the results of a given model when using the MFD as the source of the unigrams for the initial model (M1). The final column shows the results when the AR unigrams are used as the initial source of supervision.

This table highlights a subset of the results to show the overall trends of the full results shown in Tables 5.6 and 5.7. As can be seen in all three tables, as we add more information with each PSL model, the overall results continue to improve, with the final model (M13) achieving the highest F_1 score for both sources of unigrams.

An interesting trend to note is that the AR unigrams result in better average performance for most of the models until Model 9. Models 9 and above incorporate the most powerful features: bigrams and trigrams with phrases and frames. This suggests that the AR unigrams, designed specifically for the political Twitter domain, are more useful than the MFD unigrams, when only unigrams are available. Conversely, the MFD unigrams are designed to *conceptually* capture morality, and therefore have weaker performance in the unigram-based models, but achieve higher performance when combined with the more powerful features of the higher models.

Table 5.6.
F_1 Scores of PSL Models Using The Moral Foundations Dictionary (MFD). The highest prediction per moral is marked in bold. Average represents the macro-weighted average F_1 score over all morals.

Moral Fdn.	Results of Non-Joint PSL Model Predictions												
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
Care	16.61	52.51	43.34	53.24	53.38	53.59	55.64	62.40	66.00	66.48	67.32	67.59	**67.78**
Harm	12.57	47.62	42.58	50.39	57.24	55.29	60.06	67.06	71.58	71.58	72.39	**73.68**	73.54
Fairness	24.68	52.22	45.16	50.22	51.50	50.86	61.54	71.13	74.00	74.50	75.32	**75.48**	**75.48**
Cheating	0.00	0.00	0.00	0.00	0.00	0.00	0.00	21.05	51.85	51.85	56.14	**60.00**	**60.00**
Loyalty	18.29	44.53	41.49	43.87	43.59	44.22	47.65	59.15	62.82	63.75	63.75	63.95	**64.20**
Betrayal	0.00	0.00	10.00	20.00	20.00	20.00	18.18	34.78	66.67	66.67	68.42	**70.00**	**70.00**
Authority	0.00	30.93	30.19	33.10	35.53	33.96	45.52	55.29	62.50	65.91	67.78	69.23	**69.61**
Subversion	3.77	32.69	13.39	25.90	24.66	42.36	59.29	72.66	77.29	78.08	78.41	79.22	**79.61**
Purity	0.00	8.89	4.88	9.88	9.76	56.12	63.86	70.86	72.13	74.16	76.09	79.14	**80.41**
Degradation	2.99	15.38	9.52	10.00	10.00	8.00	20.69	52.94	61.54	61.54	68.09	**73.47**	**73.47**
Non-moral	0.00	0.00	1.60	3.51	12.70	12.31	54.55	71.14	80.90	81.82	82.35	82.54	**83.33**
Average	7.17	25.89	22.01	27.28	28.94	34.25	44.27	58.04	67.93	68.76	70.55	72.21	**72.49**

Table 5.7.
F_1 Scores of PSL Models Using Annotator's Rationale (AR). The highest prediction per moral is marked in bold. Average represents the macro-weighted average F_1 score over all morals.

Moral Fdn.	Results of Non-Joint PSL Model Predictions												
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
Care	7.29	29.72	30.51	30.86	30.62	35.66	46.41	54.17	61.77	62.16	62.91	64.79	**64.91**
Harm	2.25	8.89	19.31	21.89	26.18	26.09	37.28	52.40	62.18	62.18	63.74	64.67	**64.86**
Fairness	9.15	26.43	27.12	28.70	30.43	31.92	53.56	69.88	72.52	72.52	74.26	**74.63**	**74.63**
Cheating	4.76	13.33	25.45	25.45	38.71	39.34	40.68	51.61	62.16	62.16	64.94	**65.82**	**65.82**
Loyalty	2.61	19.66	23.85	25.10	27.31	29.57	38.06	47.73	54.30	55.22	55.59	57.34	**57.91**
Betrayal	0.00	0.00	0.00	6.25	12.12	11.76	18.18	28.57	60.47	60.47	62.22	**65.22**	**65.22**
Authority	13.59	40.19	48.40	51.82	56.25	56.14	57.04	63.30	66.45	66.67	67.32	**67.53**	**67.53**
Subversion	4.79	40.69	42.34	43.21	43.93	44.03	47.20	55.12	56.47	56.47	57.07	57.53	**57.65**
Purity	5.62	13.64	19.78	23.16	30.00	60.38	69.66	76.67	79.35	79.35	80.21	81.82	**82.52**
Degradation	16.66	31.37	37.74	44.83	51.61	51.61	57.14	68.75	73.53	73.53	77.33	**78.95**	**78.95**
Non-moral	28.78	52.99	60.48	61.33	64.72	66.00	73.62	79.41	82.25	82.25	82.55	82.78	**83.20**
Average	8.68	25.17	30.45	32.96	37.44	41.14	48.98	58.87	66.50	66.63	68.01	69.19	**69.38**

Table 5.8.
Overview of Macro-average F_1 Scores of PSL Models. The baseline represents the traditional, majority vote approach.

PSL Model	MFD	AR
Baseline	12.5	10.86
M1	7.17	8.68
M3	22.01	30.45
M5	28.94	37.44
M9	67.93	66.50
M13	72.49	69.38

5.5.2 Analysis of Joint Experiments

In addition to studying the effects of each feature on the ability of the models to predict moral foundations, we also explored the joint prediction of both the policy frames and moral foundations. These two tasks are highly related as shown by the large increase in score between the baseline and skyline measurements in Table 5.9 once frames are incorporated into the models.

Both moral foundations and frame classification are challenging multi-label classification tasks, the former using 11 possible foundations and the latter consisting of 17 possible frames. Furthermore, joint learning problems are harder to learn due to larger numbers of parameters, which in turn affect learning and inference.

Table 5.9 shows the macro-average F_1 scores for three different models. The baseline model refers to Model 13 with all features *except frames*. The joint model is Model 13 designed to predict both the moral foundation and frame of a tweet simultaneously (as shown in Table 5.4). Finally, the skyline model is Model 13 with all features, where the frames are initialized with their known values.

The joint model using AR unigrams outperforms the baseline, showing that there is some benefit to modeling both moral foundations and frames together, as well as using domain-specific unigrams. However, it is unable to beat the MFD-based unigrams model. This is likely due to the large amount of noise introduced by incorrect frame predictions into the joint model. As expected, the joint model does not outperform the skyline model which is able to use the known values of the frames in order to accurately classify the moral foundations associated with the tweets.

Finally, the predictions for the frames in the joint model were quite low, going from an average F_1 score of 26.09 in MODEL 1 to an average F_1 score of 27.99 in MODEL 13. We believe this to be because the frames are predicted with *no initialization*. In the previous chapter, we initialized the frame prediction models with a set of unigrams expected to occur for each frame. Different from this approach, the only information these models provide to the frames are political party, issue, associated bigrams and trigrams, and the *predicted values for the moral foundations* from this information. The F_1 score of 27.99 with little initialization indicates that there is indeed a relationship between policy frames and the moral foundations expressed in tweets worth exploring in future work.

Table 5.9.
Overview of Macro-average F_1 Scores of Joint PSL Models.

PSL MODEL	MFD	AR
BASELINE	55.49	55.88
JOINT	51.22	58.75
SKYLINE	72.49	69.38

5.5.3 Analysis of Supervised Experiments for Social and Behavioral Models

The first column of Table 5.10 shows the results when using only language-based features in the PSL models. This baseline model corresponds to M13 in Table 5.3. Since we are interested in showing the benefits of modeling social network and behavioral features in addition to language features, we use this as our baseline to show improvement against. The second column presents results when politician retweet information, i.e., when politicians retweet each other, is included into the language model. Similarly, the third column is when following information, i.e., when politicians are following other politicians, is used in the prediction. Finally, the last column indicates the results when features related to the timing of tweets are incorporated into the model.

This table shows that for all moral foundations adding features of social or behavioral information extracted from politician's Twitter networks improves the overall prediction, with a 9.14 point increase in average F_1 score over all foundations.

For most foundations however, incorporation of retweet information did not increase the score, and in some cases lowered the score. This could be due to two likely reasons: first, there is a low quantity of retweet information in this dataset, resulting in too little social information to increase the score, or second, many retweets are a copy of the original tweet with little new information added. In such cases, the model would only have access to the language-based features used in the baseline. However, based on the results of Table 5.10, retweet information is a useful predictor of the Subversion moral foundation. This is reflected in the data in tweets where a politician from one political party retweets a politician from the opposite party in order to criticize their statement in the original tweet.

5.5.4 Analysis of Unsupervised Experiments

Prior related works do not provide unsupervised analyses for their approaches for classifying moral foundations in tweets. Therefore, we used the language-based features as our language only baseline PSL model (shown in column one of Table 5.11). The remaining columns of Table 5.11 correspond to the addition of each social-behavioral network feature, similar to the supervised testing approach.

Table 5.10.
F_1 Scores of Supervised Experiments. Numbers in boldface indicate the highest prediction. The average is the macro-weighted average F_1 score over all moral foundations.

Moral Fdn.	Results of PSL Model Predictions			
	Baseline	+Retweets	+Following	+Temporal
Care	67.78	67.78	69.75	**75.59**
Harm	73.68	73.64	73.32	**77.65**
Fairness	75.48	75.48	80.14	**85.40**
Cheating	60.00	60.00	61.02	**65.81**
Loyalty	64.20	64.19	65.57	**75.10**
Betrayal	70.00	70.00	71.67	**72.11**
Authority	69.61	69.62	70.67	**71.43**
Subversion	79.61	81.19	85.82	**88.58**
Purity	80.41	80.43	81.29	**85.95**
Degradation	73.47	72.30	72.83	**74.42**
Non-moral	83.33	83.35	88.27	**92.31**
Average	72.49	74.16	76.02	**81.63**

These results support the overall findings of this dissertation that the addition of social and behavioral information results in the best prediction in an unsupervised setting. The final combined model has an improved average F_1 score of 12.06 points over the language-only baseline. Furthermore, approximately half of the predictions exceed the reported inter-annotator agreement of 67.2% for this dataset, calculated using Cohen's Kappa coefficient (Johnson and Goldwasser, 2018), suggesting that weakly supervised models incorporating social and behavioral information can help overcome the need for annotation, even in an unsupervised approach.

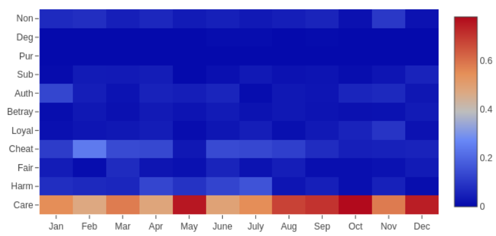
Table 5.11.
F_1 Scores of Unsupervised Experiments. Numbers in boldface indicate the highest prediction. The average is the macro-weighted average F_1 score over all moral foundations.

Moral Fdn.	Results of PSL Model Predictions			
	Baseline	+Retweets	+Following	+Temporal
Care	55.49	56.37	63.99	**67.23**
Harm	53.11	53.21	55.07	**64.40**
Fairness	56.22	56.22	64.78	**68.80**
Cheating	38.06	40.00	44.29	**47.92**
Cheating	49.91	50.34	54.82	**59.09**
Loyalty	50.00	50.00	51.79	**57.78**
Betrayal	52.32	52.73	56.43	**58.15**
Authority	55.80	57.61	62.04	**64.40**
Subversion	62.11	62.54	63.422	**67.50**
Purity	52.34	52.34	57.27	**60.95**
Degradation	57.51	57.88	71.01	**73.98**
Average	52.69	53.57	61.20	**64.75**

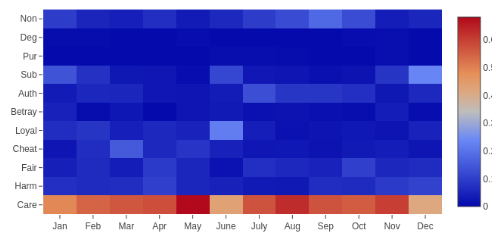
5.6 Qualitative Analysis

In this section, we present two case studies showing the usefulness of the weakly supervised models in an unsupervised setting for the analysis of the relationships between moral foundations used in social media discourse and real world political behavior. Predicted moral foundations were obtained by running the tweets from the two Senate collections of 2016 and 2018, as described in Section 5.2, through the unsupervised PSL model.

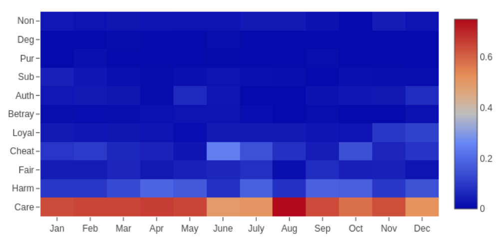
Figure 5.3 shows the predicted moral foundations for each political party over the two years of 2016 and 2018. Figures 5.4 through 5.6 show the distributions of moral



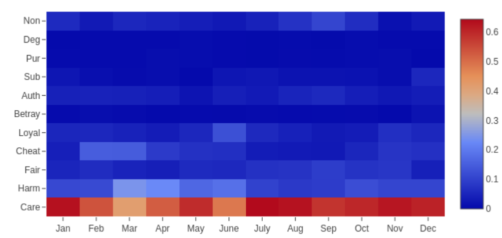
(a) Republican Tweets 2016.



(b) Republican Tweets 2018.



(c) Democrat Tweets 2016.



(d) Democrat Tweets 2018.

Figure 5.3. Monthly Coverage of Moral Foundations in Republican and Democrat Tweets.

foundations used by each party in tweets discussing specific events. Last, we present a final qualitative study showing how combining multiple aspects of political discourse, including ideology, stance, framing strategies, and moral foundations, leads to a more refined prediction.

5.6.1 Moral Foundations Trends Across Years

Figure 5.3(a) and Figure 5.3(b) show the predicted moral foundations of Republicans' tweets in 2016 and 2018, respectively, concerning the six issues studied in this work: health care, women's rights, gun violence, immigration, terrorism, and LGBTQ rights. From these two figures, we can see that Republicans favor the Care foundation, but still use the other foundations as well throughout the year. However, there is a greater concentration of tweets expressing Care in 2016 compared to 2018, in which use of this foundation drops. Consequently, the use of other moral foundations increases in 2018 and is more evenly spread out throughout the year.

In Figure 5.3(a), there are two areas with peak use of the Care foundation during 2016. The first is around June and corresponds to increased Twitter activity discussing *Whole Woman's Health v. Hellerstedt*, a Supreme Court case concerning women's rights to health care, and the Orlando Pulse Nightclub shooting, an event related to both terrorism and gun violence. The second peak is during the months of September and October and corresponds to increased activity in the months proceeding November in which the midterm elections were held. Figure 5.3(b) also reflects this peak in the months proceeding the midterm elections for 2018. Furthermore, activity in this time frame spiked in July due to the Brett Kavanaugh confirmation hearings. Figures 5.3(c) and 5.3(d) similarly show the predicted moral foundations of Democrats' tweets in 2016 and 2018, respectively. Figure 5.3(c) shows that Democrats favor the first four moral foundations (Care, Harm, Fairness, and Cheating) more evenly. This only changes during a spike in activity in June, over the same issues which caused an increase in Republican activity. However, the lower frequency of foundations used in 2016 correlates with the more infrequent use of Twitter by Democratic Senators. This changes dramatically in Figure 5.3(d), which shows that Democratic activity discussing these issues on Twitter *triples*. Additionally, more moral foundations are used throughout 2018 by Democrats.

Similar to Republicans in 2018, Democrats also show a spike in activity and moral foundations during the months of July to October. Tweets from these months also correspond to the Kavanaugh hearings and pre-election activity. An interesting point between the two 2018 heatmaps is that both Republicans and Democrats use the Care foundation in their tweets in similar proportions during these months, but their use of other foundations is more varied.

5.6.2 Event-specific Moral Foundations Trends

We have observed that when events occur, such as a shooting, Twitter activity discussing the event peaks on the day of the event and gradually diminishes over the following weeks. Figures 5.4 through 5.6 highlight key events in 2016 and 2018 for three different policy issues: gun violence, women's rights, and LGBTQ rights. Each heat map shows the frequency of each moral foundation used by Republicans and Democrats to discuss these specific events, for one month after the event occurs.

Gun Violence. Figure 5.4 shows the predicted moral foundations for tweets discussing two events related to gun violence. The first is the June 12, 2016 shooting at the Pulse Nightclub in Orlando, Florida. The first column of the heat map shows Republican moral foundations used to discuss this shooting. The second column shows the foundations used by Democrats. Columns three and four are the Republican and Democrat foundations used to discuss the Marjory Stoneman Douglas High School shooting on February 14, 2018. For both parties, over both years, the first four moral foundations (i.e., Care, Harm, Fairness, and Cheating) are used more frequently than all others. Similar to the yearly trends, Care is the most used foundation to discuss these events. This is to be expected because after shootings both parties express

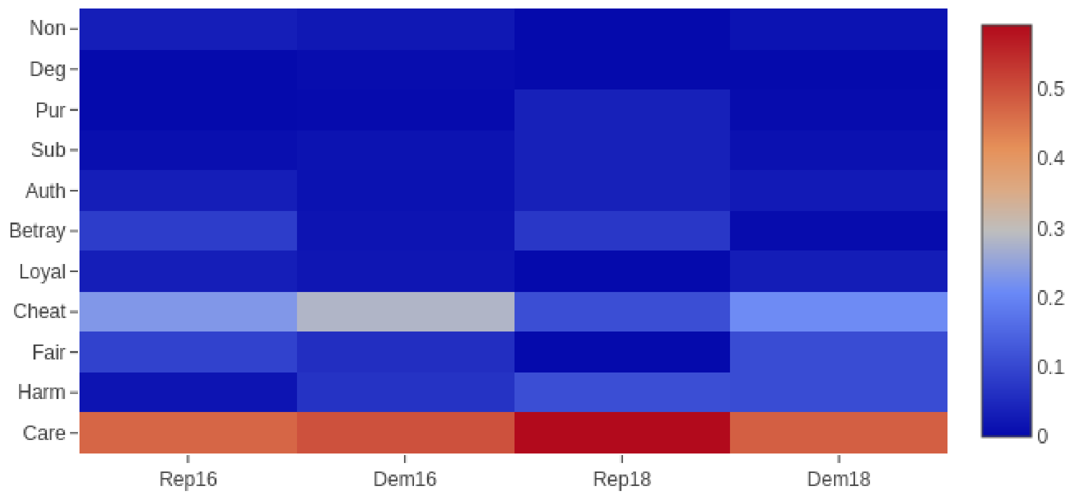


Figure 5.4. Moral Foundations of Tweets Discussing Shooting Events. The two columns on the left are predictions for tweets one month after the Orlando Pulse Nightclub shooting. The two columns on the right are predictions for tweets one month after the Marjory Stoneman Douglas High School shooting.

their concern for the victims and families and offer their “thoughts and prayers” to those affected. Two interesting trends are shown in this heat map: (1) an increase from 2016 to 2018 in the use of the Care foundation by Republicans and the Harm and Fairness foundations by Democrats, and (2) increased use of the Cheating moral foundation when compared to other events. This foundation appears in tweets related to a lack of justice for the victims of the shootings and their families, as well as tweets discussing the need for blood donations for the Orlando victims being hindered by unjust blood donor restrictions.

Women’s Rights. Figure 5.5 presents a similar heat map for two events related to women’s rights. The first two columns are the predicted moral foundations of Republican and Democrat tweets for the *Whole Woman’s Health v. Hellerstedt* Supreme

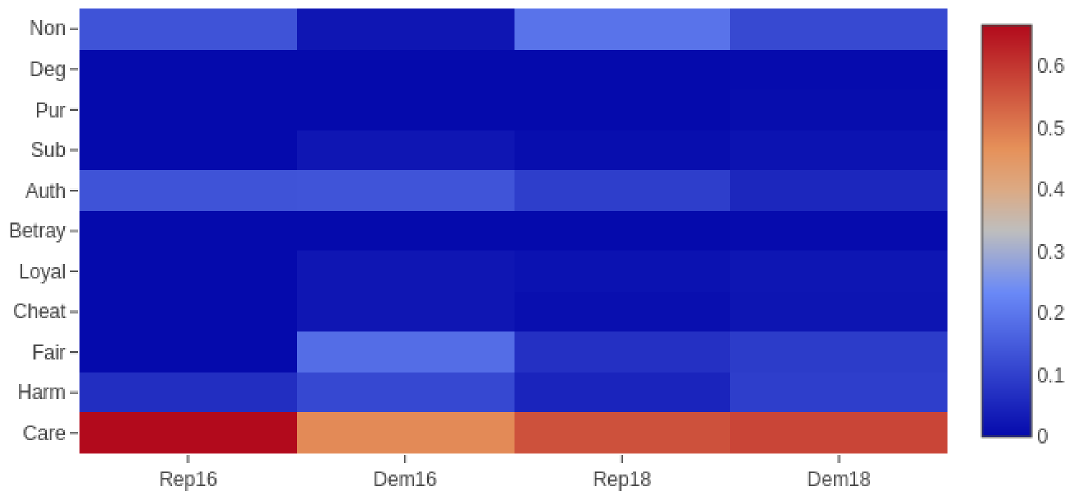


Figure 5.5. Moral Foundations of Tweets Discussing Events Related to Women's Rights and the Supreme Court. The two columns on the left are predictions for tweets one month after the *Whole Women's Health v. Hellerstedt* Supreme Court case. The two columns on the right are predictions for tweets during the month of testimonies from the Brett Kavanaugh confirmation hearing.

Court case which determined that laws enacted by Texas placed an undue burden on women seeking a legal abortion, and thus were unconstitutional. The second two columns correspond to predicted foundations for tweets discussing the testimony of Dr. Christine Blasey Ford in the Brett Kavanaugh Supreme Court nomination hearing. For both parties and years, the top moral foundations used are Care, Harm, Authority, and Non-moral. Interestingly, Democrats in 2016 discuss this issue in terms of Fairness, but the use of Fairness in 2018 declines and is replaced with Non-moral arguments. In 2016, both parties use the Authority foundation to discuss support or lack thereof for the Supreme Court and President Obama on this issue. However, in 2018, there is a significant decrease in the use of this foundation, while the use of the Non-moral foundation increases for both parties. For Republicans in

2018, the top foundations are Care and Authority, reflected in tweets which discuss a simultaneous care and support for the hearing proceedings and Kavanaugh's reputation. Democrats, however, use Care, Harm, and Fairness as their top foundations to express concern about the potentially harmful effect on legislation pertaining to women's rights that his nomination to the Supreme Court might cause.

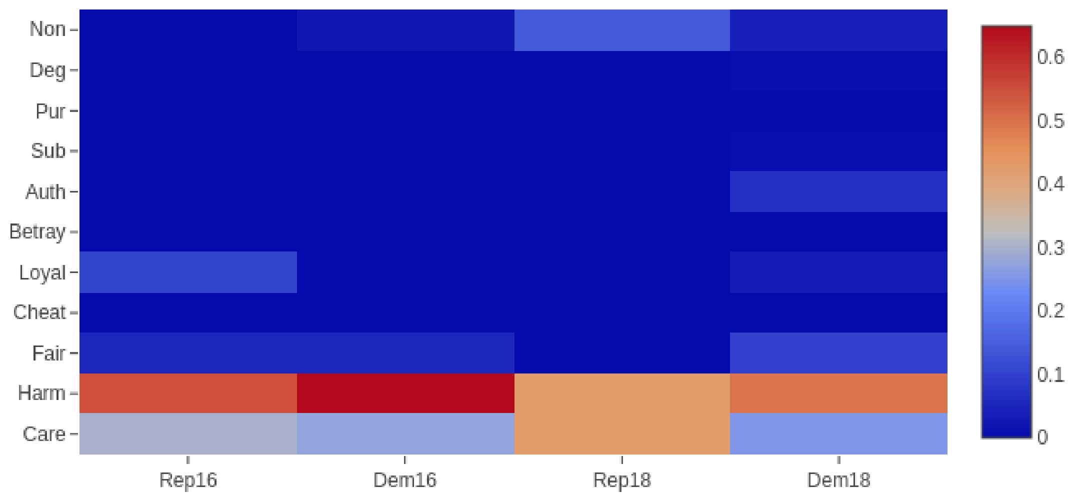


Figure 5.6. Moral Foundations of Tweets Discussing Events Related to Transgender Rights. The two columns on the left are predictions for tweets one month after the North Carolina "bathroom bill". The two columns on the right are predictions for tweets one month after the current administration announced transgender people would not be allowed to serve in the military.

LGBTQ Rights. Figure 5.6 presents a heat map of predicted moral foundations concerning two events related to transgender rights. The leftmost columns represent tweets discussing the passage of the *Public Facilities Privacy & Security Act* in North Carolina which constrains transgender people to only access bathrooms corresponding to their gender at birth. The rightmost columns represent tweets discussing the current administration's proposed ban prohibiting transgender people from serving in the military.

For this issue, both parties use a dual Care-Harm foundation to express concern over how the legislation will harm differing populations. Different from most issues, there is a greater emphasis on the harm such legislation could cause, as evidenced by the significantly higher representation of Harm foundation predictions for all groups, except the Republicans in 2016.

5.6.3 Disambiguation of the Frames and Moral Foundations Underlying Political Discourse on Twitter

Throughout this dissertation, my findings have shown that modeling more abstract representations of language, such as framing strategies or ideological phrases, instead of low level words directly extracted from tweets, results in a more accurate prediction for the political discourse tasks of frames and moral foundations classification. However, since policy frames are designed to be issue independent and generalize across political parties and over time, it is common to have opposing political ideologies represented by similar framing techniques.

By combining the predictions of all of the models presented in this dissertation, it is possible to develop a more comprehensive view of the tweets in order to understand the politician's stance on the issue, how they are framing this stance, and what moral foundation is motivating their framing strategy or stance, even if the tweets may express different messages under similar labels. Each of these aspects of the tweet could be wrongly classified using traditional NLP-based approaches, but by using a holistic view of language and behavior patterns on Twitter, our models are able to accurately disambiguate political obfuscation to reveal the political strategies used to deliver political messages.

Consider Tables 5.12 and 5.13 which present tweets taken from Tables 4.17, 4.18, and 4.19 in Chapter 4. The event column refers to three mass shooting events: the

Orlando PULSE nightclub shooting on June 12, 2016 (OPN), the Inland Regional Center of San Bernadino shooting on December 2, 2015 (IRC), and the Emanuel African Methodist Episcopal (EAME) Church shooting on June 17, 2015. The ideology column lists the politician's political party affiliation, either conservative (Cons) or liberal (Lib). The NLP PF column lists the frames that are predicted using a traditional unigram-based approach. To predict a frame, the unigrams in the tweet are matched to unigrams associated with each frame and the highest number of matches is chosen as the predicted label. However, these approaches often only choose one label. The NLP PM column is calculated in a similar manner: by counting the number of unigrams in the tweet that match unigrams listed in the Moral Foundations Dictionary. Again, traditional models using this frequency-matching approach often assign only the highest matching label. The PSL PF and PSL PM columns are the frames and moral foundations that are predicted by the best performing PSL models of Chapters 3 and 4. When interpreting these tables it is important to note that the NLP columns list multiple labels, but in traditional unigram frequency counting approaches, the highest match would be the only label selected. If there is a tie, as is the case in our examples since each frame or foundation matches one word in the tweet, it is either broken arbitrarily or the tweet is not labelled.

Tables 5.12 and 5.13 highlight the usefulness of the weakly supervised modeling approach presented in this dissertation by providing examples of when traditional language-based baselines fail to capture the entire essence of a tweet. Contrary to this failure, a weakly supervised modeling approach combining language, social, and behavioral features does give a more accurate prediction of the frames and morals within a tweet, further providing a more comprehensive view of the tweet's message.

The first key problem is that traditional approaches study framing and moral foundations techniques with unigram frequency matching. Unlike moral foundations, policy frames do not have a set dictionary of unigrams associated with each frame. As shown in the two tables, this leads to inconsistent results: half of the example tweets would be misclassified using a traditional unigram approach, four out of the twelve tweets only have one of the possible labels correct, and only one of the tweets is correctly predicted by both methods. From these tables we can also see that the frame predictions often misclassify the moral or religious overtones of a tweet because these tones tend to use words that would match to Frame 3 (*Religion & Morality*). There is also a greater emphasis on Frame 7 (*Security & Defense*) to focus on prevention of gun violence, and greater confusion of this frame with Frame 8 which typically handles health and recovery after a shooting. Frame misclassification also occurs frequently due to a unigram in the tweet, e.g., cancer, matching an expected unigram for a frame, e.g., the Health & Safety frame. Consider the tweet by Harry Reid, in the last row of Table 5.13: traditional approaches would label this tweet as Frame 8, when it actually has nothing to do with health.

For moral foundations predictions in these two tables, only two tweets have more than one unigram match to a foundation: Lynch (row 4, Table 5.12; this tweet is also the only one where all predictions match) and Duncan (row 5, Table 5.13). Of the twelve example tweets, only these two would have been useful for experiments.

Another interesting trend in these two tables is that the NLP models are able to detect one unigram match to a certain moral foundation, and the PSL model is able to match to a different, but more correct foundation, e.g., the tweet by Lisa Murkowski (row 1, Table 5.13). Additionally, the PSL models, which use abstractions of language or behavioral information, are able to detect moral foundations that the NLP approach misses because there are no direct unigram matches.

This subset of results highlights the overall findings of this dissertation: that using abstracted language, social, and behavioral features instead of traditional language features results in predictions that can be used to understand the comprehensive political meaning of a tweet.

5.7 Chapter Summary

Moral foundations and policy frames can be employed as political strategies in which politicians use these techniques to garner support from the public. Politicians carefully word their statements to express their moral and social position on issues, while maximizing their base's response to their message. In this chapter we presented global PSL models for the classification of moral foundations expressed in political discourse on microblogs, specifically Twitter. We show the benefits and drawbacks of both traditionally used MFD unigrams and domain-specific unigrams for initialization of the models. We provide an initial approach to the joint modeling of policy frames and moral foundations. We also show experimental results demonstrating the effectiveness of social and behavioral information extracted from tweets and the political networks of Twitter for accurate moral foundations classification.

Table 5.12.
Example Tweets Associated With Mass Shooting Events: Orlando Pulse Nightclub (OPN), Inland Regional Center of San Bernadino (IRC), and the EAME Church of Charleston (EAME). Ideol represents a conservative (Cons) or liberal (Lib) ideology; PF are the predicted frames using a traditional NLP method or our new PSL method; PM are the predicted moral foundations using an NLP or PSL method.

Event	Politician	Ideol	Tweet	NLP PF	PSL PF	NLP PM	PSL PM
OPN	Bob Goodlatte	Cons	The attack in #Orlando was an act of pure evil. My prayers are w/ the families of victims and the injured. We will continue seeking answers.	3, 7	3, 17	2, 5, 9, 11	1, 5, 9
OPN	Yvette Clark	Lib	I am deeply saddened by the act of hate and terror enacted on the lives of Orlando's LGBT Community and I #StandWithOrlando	7, 9, 10	3, 17	5	2, 5
EAME	Patrick McHenry	Cons	My thoughts and prayers are with the #Charleston community this morning. Horrific to see this violence anywhere, much less a house of worship.	7, 9, 10	3, 17	2, 5	2, 5, 9
EAME	Stephen Lynch	Lib	My thoughts and prayers go out to the victims and their families in #Charleston. A church is a house of the Lord, a place of peace.	3	3, 17	1, 5, 9	1, 5, 9
EAME	Mick Mulvaney	Cons	The shooting in Charleston that took the lives of nine people has shaken that community and the state as a whole.	9	9	5	11
EAME	Seth Moulton	Lib	My thoughts and prayers are with Charleston, SC today. We must work to end these senseless acts of violence in our communities.	7, 9	9, 17	2, 5	1, 5, 9

Table 5.13.
Example Tweets Associated With Mass Shooting Events. Continued from Table 5.12.

Event	Politician	Ideol	Tweet	NLP PF	PSL PF	NLP PM	PSL PM
OPN	Lisa Murkowski	Cons	What happened in Orlando was an absolute tragic act of terrorism spawned by an ideology of hate being pushed by ISIS.	7	3	6	2, 10
OPN	David Cicilline	Lib	Voters should absolutely hold us accountable for what we're doing or not doing to address gun violence.	7, 10	3	2	1, 7
OPN	Mark Kirk	Cons	Americans need to know Washington is listening - We must keep guns out of the hands of suspected terrorists	6, 7, 10	7	6	1, 9
OPN	Kirsten Gillibrand	Lib	As we mourn victims of yet another tragedy, time to finally act on commonsense gun safety reforms supported by the American people.	7, 8, 10	11, 12	1	1, 5, 7
IRC	Jeff Duncan	Cons	Wish Obama and Clinton would speak out about terrorism and radical Islamic jihad as quickly as they call for gun control here in America	7, 8	3, 12	6, 7	2, 3, 8
IRC	Harry Reid	Lib	Gun violence has become a cancer on this nation. We must make common sense gun reforms that keep weapons out of dangerous hands.	7, 8, 10	7	2, 5	1, 2, 7

6 CONCLUSION

In this dissertation, I have presented a framework for modeling the dynamic nature of political discourse on Twitter by incorporating linguistic, social, and behavioral features from Twitter. My approach can be modified to handle additional politicians or issues, as well as those of other countries, by incorporating the proper domain knowledge (e.g., replacing party with voting history, using new keywords for different issues in other countries, or changing events such as Supreme Court rulings to Parliament votes). Further, because these models require minimal initial supervision, they can be applied to future political discourse as well.

Contrary to previous works, which typically focus on a single aspect of this complex microblogging behavior, I construct holistic models connecting party line biases, temporal behaviors, and issue framing into global predictive models. These models are capable of identifying fine-grained stances and agreement patterns, as well as issue-independent framing strategies and moral motivations. Despite having no direct supervision and using only intuitive local classifiers to bootstrap the global model, my approach results in a strong predictive model which helps shed light on political discourse within and across party lines. My models also serve as an interesting exploratory tool to study the evolution of trends in political discourse patterns on Twitter and their ramifications in the real world.

This dissertation has outlined the progression of my weakly supervised modeling approach for the classification and analysis of political discourse on social media microblogs. Previous works in this domain have modeled basic language features for specific prediction tasks. My results have shown that by modeling abstractions of higher-level patterns of language, social relationships, and behavioral activities on Twitter, political discourse prediction tasks can be studied at both user and tweet level, as well as across a variety of political issues and administrations. By leveraging these abstractions, rather than raw textual input alone, weakly supervised models which require minimal initialization can be designed to understand and explore the dynamics of political discourse on social media. These models can shed light on the interplay of political ideologies, morality, and policy framing techniques to help society better understand how politicians are delivering their messages to the public and the real-world ramifications of these political discourse strategies.

Bibliography

Abbott, R., Walker, M., Anand, P., Fox Tree, J. E., Bowmani, R., and King, J. (2011). How can you say such things?!?: Recognizing disagreement in informal political argument. In *Proc. of the Workshop on Language in Social Media*.

Abu-Jbara, A., King, B., Diab, M., and Radev, D. (2013). Identifying opinion subgroups in arabic online discussions. In *Proc. of ACL*.

Bach, S. H., Broecheler, M., Huang, B., and Getoor, L. (2015). Hinge-loss markov random fields and probabilistic soft logic. *arXiv preprint arXiv:1505.04406*.

Bach, S. H., Huang, B., London, B., and Getoor, L. (2013). Hinge-loss Markov random fields: Convex inference for structured prediction. In *Proc. of UAI*.

Bakliwal, A., Foster, J., van der Puil, J., O'Brien, R., Tounsi, L., and Hughes, M. (2013). Sentiment analysis of political tweets: Towards an accurate classifier. In *Proc. of ACL*.

Bamman, D. and Smith, N. A. (2015). Open extraction of fine-grained political statements. In *Proc. of EMNLP*.

Baumer, E., Elovic, E., Qin, Y., Polletta, F., and Gay, G. (2015). Testing and comparing computational approaches for identifying the language of framing in political news. In *Proc. of NAACL*.

Bermingham, A. and Smeaton, A. F. (2011). On using twitter to monitor political sentiment and predict election results. In *Proc. of the Workshop on Sentiment Analysis where AI meets Psychology*.

Boydstun, A., Card, D., Gross, J. H., Resnik, P., and Smith, N. A. (2014). Tracking the development of media frames within and across policy issues. *American Political Science Association*.

Burch, L. M., Frederick, E. L., and Pegoraro, A. (2015). Kissing in the carnage: An examination of framing on twitter during the vancouver riots. *Journal of Broadcasting & Electronic Media*, 59(3):399–415.

Card, D., Boydstun, A. E., Gross, J. H., Resnik, P., and Smith, N. A. (2015). The media frames corpus: Annotations of frames across issues. In *Proc. of ACL*.

Choi, E., Tan, C., Lee, L., Danescu-Niculescu-Mizil, C., and Spindel, J. (2012). Hedge detection as a lens on framing in the gmo debates: A position paper. In *Proc. of ACL Workshops*.

Chong, D. and Druckman, J. N. (2007). Framing theory. *Annu. Rev. Polit. Sci.*, 10:103–126.

Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A., and Menczer, F. (2011). Predicting the political alignment of twitter users. In *Proc. of PASSAT*.

Djemili, S., Longhi, J., Marinica, C., Kotzinos, D., and Sarfati, G.-E. (2014). What does twitter have to say about ideology? In *NLP 4 CMC*.

Ebrahimi, J., Dou, D., and Lowd, D. (2016). Weakly supervised tweet stance classification by relational bootstrapping. In *Proc. of EMNLP*.

Eisenstein, J. (2013). What to do about bad language on the internet. In *Proc. of NAACL*.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58.

Fulgoni, D., Carpenter, J., Ungar, L., and Preotiuc-Pietro, D. (2016). An empirical exploration of moral foundations theory in partisan news sources. In *Proc. of LREC*.

Ganitkevitch, J., Van Durme, B., and Callison-Burch, C. (2013). The paraphrase database. In *Proc. of NAACL-HLT*.

Garten, J., Boghrati, R., Hoover, J., Johnson, K. M., and Dehghani, M. (2016). Morality between the lines: Detecting moral sentiment in text. In *IJCAI workshops*.

Gerrish, S. and Blei, D. M. (2012). How they vote: Issue-adjusted models of legislative behavior. In *Advances in Neural Information Processing Systems*, pages 2753–2761.

Graham, J., Haidt, J., and Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5):1029.

Graham, J., Nosek, B. A., and Haidt, J. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PloS one*, 7(12):e50092.

Greene, S. and Resnik, P. (2009). More than words: Syntactic packaging and implicit sentiment. In *Proc. of NAACL*.

Groshek, J. and Al-Rawi, A. (2013). Public sentiment and critical framing in social media content during the 2012 u.s. presidential campaign. *Social Science Computer Review*, 31(5):563–576.

Haidt, J. and Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20(1):98–116.

Haidt, J. and Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4):55–66.

Harlow, S. and Johnson, T. (2011). The arab spring— overthrowing the protest paradigm? how the new york times, global voices and twitter covered the egyptian revolution. *International Journal of Communication*, 5(0).

Hasan, K. S. and Ng, V. (2014). Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proc. of EMNLP*.

Huang, B., Bach, S. H., Norris, E., Pujara, J., and Getoor, L. (2012). Social group modeling with probabilistic soft logic. In *NIPS Workshops*.

Iyyer, Enns, Boyd-Graber, and Resnik (2014). Political ideology detection using recursive neural networks. In *Proc. of ACL*.

Jang, S. and Hart, P. (2015). Polarized frames on "climate change" and "global warming" across countries and states: Evidence from twitter big data. *Global Environmental Change*, 32:11–17.

Johnson, K. and Goldwasser, D. (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*.

Johnson, K. and Goldwasser, D. (2018). Classification of moral foundations in microblog political discourse. In *Proc. of ACL*.

Johnson, K., Jin, D., and Goldwasser, D. (2017a). Leveraging behavioral and social information for weakly supervised collective classification of political discourse on twitter. In *Proc. of ACL*.

Johnson, K., Jin, D., and Goldwasser, D. (2017b). Modeling of political discourse framing on twitter. In *Proc. of ICWSM*.

Lee, I.-T., Goindani, M., Li, C., Jin, D., Johnson, K., Zhang, X., Pacheco, M., and Goldwasser, D. (2017). Purduenlp at semeval-2017 task 1: Predicting semantic textual similarity with paraphrase and event embeddings. In *Proc. of SemEval*.

Li, J., Ritter, A., Cardie, C., and Hovy, E. H. (2014a). Major life event extraction from twitter based on congratulations/condolences speech acts. In *Proc. of EMNLP*.

Li, J., Ritter, A., and Hovy, E. H. (2014b). Weakly supervised user profile extraction from twitter. In *Proc. of ACL*.

Lin, Y., Hoover, J., Dehghani, M., Mooijman, M., and Ji, H. (2017). Acquiring background knowledge to improve moral value prediction. *arXiv preprint arXiv:1709.05467*.

Maireder, A. and Ausserhofer, J. (2013). National politics on twitter: Structures and topics of a networked public sphere. In *Information, Communication, and Society*.

Marchetti-Bowick, M. and Chambers, N. (2012). Learning for microblogs with distant supervision: Political forecasting with twitter. In *Proc. of EACL*.

Meraz, S. and Papacharissi, Z. (2013). Networked gatekeeping and networked framing on #egypt. *The International Journal of Press/Politics*, 18(2):138–166.

Nguyen, V.-A., Boyd-Graber, J., Resnik, P., and Miler, K. (2015). Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th congress. In *Proc. of ACL*.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *Proc. of ICWSM*.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Pla, F. and Hurtado, L. F. (2014). Political tendency identification in twitter using sentiment analysis techniques. In *Proc. of COLING*.

Recasens, M., Danescu-Niculescu-Mizil, C., and Jurafsky, D. (2013). Linguistic models for analyzing and detecting biased language. In *Proc. of ACL*.

Ritter, A., Cherry, C., and Dolan, B. (2010). Unsupervised modeling of twitter conversations. In *Proc. of NAACL*.

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

Sim, Acree, Gross, and Smith (2013). Measuring ideological proportions in political speeches. In *Proc. of EMNLP*.

Somasundaran, S. and Wiebe, J. (2009). Recognizing stances in online debates. In *Proc. of ACL*.

Somasundaran, S. and Wiebe, J. (2010). Recognizing stances in ideological on-line debates. In *Proc. of NAACL Workshops*.

Sridhar, D., Foulds, J., Huang, B., Getoor, L., and Walker, M. (2015). Joint models of disagreement and stance in online debate. In *Proc. of ACL*.

Tan, C., Lee, L., and Pang, B. (2014). The effect of wording on message propagation: Topic- and author-controlled natural experiments on twitter. In *Proc. of ACL*.

Tausczik, Y. R. and Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.

Tsur, O., Calacci, D., and Lazer, D. (2015). A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proc. of ACL*.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proc. of ICWSM*.

Vargo, C. J., Guo, L., McCombs, M., and Shaw, D. L. (2014). Network issue agendas on twitter during the 2012 u.s. presidential election. *Journal of Communication*, 64(2):296–316.

Volkova, S., Bachrach, Y., Armstrong, M., and Sharma, V. (2015). Inferring latent user properties from texts published in social media. In *Proc. of AAAI*.

Volkova, S., Coppersmith, G., and Van Durme, B. (2014). Inferring user political preferences from streaming communications. In *Proc. of ACL*.

Volkova, S., Shaffer, K., Jang, J. Y., and Hodas, N. (2017). Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, volume 2, pages 647–653.

Walker, M. A., Anand, P., Abbott, R., and Grant, R. (2012). Stance classification using dialogic properties of persuasion. In *Proc. of NAACL*.

West, R., Paskov, H. S., Leskovec, J., and Potts, C. (2014). Exploiting social network structure for person-to-person sentiment analysis. *TACL*.

Wiebe, J., Wilson, T., Bruce, R., Bell, M., and Martin, M. (2004). Learning subjective language. *Computational linguistics*.

Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S. (2005). Opinionfinder: A system for subjectivity analysis. In *In Proc. of EMNLP*.

VITA

Kristen Marie Johnson was born in Metairie, Louisiana. She received her Bachelor of Science in Psychology from the University of New Orleans (UNO) in 2007 with honors. She then received her Masters degree in Computer Science with a concentration in Bioinformatics in 2012, under the direction of Dr. Dongxiao Zhu. She joined the Department of Computer Science at Purdue University in the fall of 2012 and began her research with Dr. Dan Goldwasser in 2015. She is broadly interested in social computing - the application of machine learning and natural language processing techniques to solve computational social science problems. Her works have been published in ACL, COLING, and ICWSM. During her graduate studies she served as the computer science representative for the Purdue Women in Science Programs, a Local Sponsorships and Exhibits Chair for NAACL 2018, the Widening NLP (WiNLP) 2019 Social Media Co-chair, and as a reviewer for NLP conferences and journals. Upon graduation, she will be joining the Department of Computer Science and Engineering at Michigan State University as an Assistant Professor.