Modeling the Dynamic Framing of Controversial Topics in Online Communities

Julia Mendelsohn
Stanford University
jmendels@stanford.edu

Abstract

Understanding how online communities frame controversial topics can have significant impacts in public perception, policy action, and social change. I present two neural-network based approaches for automatically detecting how a community frames a topic and how this framing changes over time. The diachronic word embedding model, trained on individual months of data from the online community Reddit, associates controversial topics with different words in different time periods. The Relationship Modeling Network (RMN) is an unsupervised neural model that learns linguistic descriptors for the community's relationship with these topics and builds a trajectory over these descriptors to characterize a topic's framing over time. While the diachronic word embeddings results are more clearly related to the entities and are more sensitive to framing shifts over time, the RMN is able to learn abstract descriptors that are more suitable for describing these relationships. It is furthermore able to detect the importance of persuasive words that are subtly used to construct arguments. This work provides promising foundations for applying neural networks to the problem of understanding framing.

1 Introduction

Humans use language as a crucial tool for persuading, promoting particular viewpoints, and shaping perceptions. One mechanism that gives language this power and is found in nearly all communication is known as **framing**, which generally refers to how a speaker defines an issue and relates that issue to other problems, people, or values (Entman, 1993). Framing has also been shown to have a significant effect on public perceptions of controversial issues and strongly impact policy changes and social movements (Benford and Snow, 2000; Chong and Druckman, 2007).

Because of the vast impact that framing has in our society, it is essential to take advantage of "computational techniques that can identify and draw attention to the language of framing" (Baumer et al., 2015). Much of the literature about framing focuses on how elites (political leaders and media outlets) shape everyday citizens' perceptions and attitudes (Chong and Druckman, 2007; Baumer et al., 2015; Tsur et al., 2015; Card et al., 2016). However, to the extent of the author's knowledge, there has been no computational study that can automatically extract a topic's framing from the language used by the everyday citizens in public discourse. Social media and online discussions play a major role in many people's lives, and the frames used in online communities is thus likely to shape the public's perceptions.

Furthermore, little research to date has acknowledged the dynamic nature of framing. The overall representation and perception of certain topics can change drastically within months. Another contribution of this paper are frameworks that can capture how a community's relationship with contentious topics changes over time, which could be of additional value to social science research.

In this paper, I discuss several techniques that can be used to understand the framing of controversial topics in the online community Reddit. I demonstrate that we can use neural models to capture differences in framing across different Reddit sub-communities, but I primarily focus on understanding how a topic's framing within a community can change over time.

I first discuss the use of diachronic word embeddings for analyzing the changing relationships between a community and various controversial topics. Each topic's nearest neighbors in the vector space were treated as a proxy for its frame. Thus, the ways in which a topic's neighboring words shift over time provides clues as to how its framing changes. We find that a topic's neighbors can be representative of a particular viewpoint, and we could leverage our own contextual knowledge to understand how a topic's representation has changed over time.

One main disadvantage of the word embedding approach is that it does not always directly capture the relationship between the community and these controversial entities, and analysis often relies on contextual knowledge. This brings us to the second approach: a Relationship Modeling Network (RMN). The RMN is an unsupervised recurrent neural network that models the changing relationship between a community and a topic, which I use as a proxy for the topic's dynamic framing. Drawing heavily upon the insights of Iyyer et al. (2016) and Wang et al. (2016), the RMN learns a set of linguistic descriptors to describe these relationship trajectories and provide insight into how a topic's framing shifts. Because they are not bound to the topics via semantic similarity, RMNs are able to capture more abstract descriptors than word embeddings, and could directly characterize the relationship.

2 Related Work

2.1 Framing

There is a substantial amount of work in political science, sociology, linguistics, and computer science that deals with framing theory. Framing is the psychological mechanism that involves "choosing a few elements of perceived reality and assembling a narrative that highlights connections among them to promote a particular interpretation" (Entman, 2007). This phenomenon is most obvious in fixed lexical entries (Schuldt et al., 2011). By altering the framing, specific word choices can drastically impact the public's perception. One striking example can be seen in the distinction between "Obamacare" and the "Affordable Care Act". These two phrases were framed so differently in political discourse that 35% of Americans did not even realize that they referred to the same entity, according to a 2017 poll [1]. Framing goes beyond specific word choices, however. Even when just considering the entry "Affordable Care Act", discussions can either frame the healthcare bill as an issue of social justice or one of economic burden (Tsur et al., 2015).

Recent computational approaches to automatic frame detection have focused on the media. Baumer et al. (2015) presented a SVM-based supervised classification approach for detecting framing in news articles published by major media outlets. Card et al. (2016) used an unsupervised model called a "Dirichlet persona model" to discover and characterize entities in a piece of newspaper text.

Tsur et al. (2015) used Bayesian methods and time-series analysis in order to understand the ways in which political parties control political discourse. They employed an LDA topic model to discover framing strategies, and then implemented further analysis to learn the relationship trajectory between topics and political agendas.

The primary distinction between previous work and this paper is that I focus on how topics are represented by a public online community instead of by the media or politicians. This paper is also more concerned with how the framing itself shifts over time, rather than its connections with later political action. Finally, I adapt more recently popularized techniques in artificial intelligence, such as word embeddings and recurrent neural models in my study.

2.2 Diachronic Word Embeddings

Previous work has been done using diachronic word embeddings, word vectors trained on data from different time periods, to study language change. Kulkarni et. al (2015) found promising results by using a distributional model to analyze language change in Amazon movie reviews and Twitter data.

Hamilton et al. (2016) used distributional methods of word embeddings to study the underlying causes of semantic change, specifically whether the change was a result of regular linguistic processes or cultural shifts.

[1]https://www.nytimes.com/2017/02/07/upshot/one-third-dont-know-obamacare-and-affordable-care-act-are-the-same.html

The authors found that their local neighborhood measure, which involved comparing a word's nearest semantic neighbors at each time period, was more sensitive to cultural shifts than their global measure, which directly compared a word's vector from different time periods. Because my goal is to detect changes in topics' framings, and this concept is more closely related to cultural shifts than linguistic processes, I opted to largely adapt Hamilton et al.'s local measure.

2.3 Relationship Modeling

Relationship modeling is very recent area of research that is closely aligned with this paper's goals. Two previous works have heavily influenced the Relationship Modeling Network (RMN) used in this project. RMNs have vast possibilities as a tool for research in the digital humanities (Iyyer et al., 2016), social science (Wang et al., 2016), and other fields. Iyyer et al. (2016) first introduced the RMN as an unsupervised model that learns how relationships between characters can develop throughout the course of a novel. Their RMN jointly learned linguistic descriptors of the relationships as well as the trajectories for specific character pairs (Iyyer et al., 2016).

Wang et al. (2016) adapted this model to investigate users' social roles in online Reddit communities. Their model was similar to that of Iyyer et al., with the exception that various users' Reddit comments were used as model inputs instead of text spans containing character mentions.

Both papers found RMNs to be more effective than traditional topic models, such as LDA. The descriptors produced by RMNs tended to be more abstract and more logical for describing relationships than those produced by traditional topic models. Iyyer et al. also included a Hidden Topic Markov Model baseline that also has a sequencing and temporal aspect. The authors found a similar trend, however, as the RMNs produced more descriptive and finer-grained results than the HTMM.

The RMN described in this paper draws significantly upon both of these models, but it is used with a different intent. Instead of investigating individuals' relationships with each other or within a community, I am looking at how a community relates to controversial topics, which can provide insights into how these topics are framed.

3 Experiment 1: Diachronic Word Embeddings

3.1 Data

The data for this experiment comes from the online forum Reddit, a community-focused, discussion-based, social media platform with 234 million users [2]. In this experiment, I focused on 3 sub-communities (known as "subreddits"): "r/politics" (a subcommunity dedicated to moderate politics), "r/Liberal", and "r/Conservative". The "'r/politics" sub-community has the most data and is thus the most suitable for diachronic studies. The data from "r/politics" contains posts and comments from 89 months: from January 2008 to May 2015. There is a total of 19.6 million posts and comments, and 1.05 billion tokens in r/politics. There is an average of 220 thousand posts and comments per month and 11.8 million tokens per month, with later years containing significantly more data. The data used from "r/Liberal", and "r/Conservative"spans January 2012 to May 2015. "r/Liberal" contains 85 thousand posts and comments, and "r/Conservative" contains about 655 thousand posts and comments. All posts and comments have already been tokenized with Spacy [3].

3.2 Controversial Entities

Twenty-five distinct controversial entities were studied in both the diachronic word embeddings and RMN experiments. A full list can be found in the appendix. These entities include politically divisive issues, such as "abortion" and "immigration", and politicians, such as "Obama" and "Putin". Because these sub-communities focus primarily on U.S. Politics, I also included several countries that tend to cause friction in discussions about American politics. These countries are not controversial topics in and of themselves, but are framed in similar ways.

3.3 Model

The distributional model provides a means to indirectly characterize how a the changing relationship between a community and a controversial entity by investigating how the entity's vector's semantic neighbors change

[2]https://about.reddit.com/advertise/
[3]https://spacy.io/

over time. I trained word vectors for each individul month of data from "r/politics". I then manually observed compared differences and observed trends in how these associated words shifted in the semantic space contructed by the communities themselves. Because "r/Conservative" and "r/Liberal" are considerably smaller, I trained word vectors on the entire data set and did not investigate diachrony in these two communities.

I trained word embeddings using the word2vec skip-gram negative sampling technique as described by Mikolov et al. (2013), and implemented the model with the Gensim toolkit [4].

With word2vec skip-gram technique, each word $word_i$ is represented by two vectors: v_i, which is the vector that represents when the word is in the center of a fixed window, and u_i, which represents the word when it is in another word's context. With the skip-gram technique, word embeddings are optimized in order to predict a window's context words $u_{i-j}, ..., u_{i-1}, u_{i+1}, ..., u_{i+j}$ given a center word v_i and a window of size $2j + 1$. The model is a 2-layer feedforward neural network that is optimized via Stochastic Gradient Descent. The standard loss, J, includes calculation of a softmax function in order to estimate the probabilities of a context word occurring given the center word:

$$J = -log \prod_{j=0,j\neq m}^{2m} P(u_{c-m+j}|v_c) = -log \prod_{j=0,j\neq m}^{2m} \frac{exp(u_{c-m+j}^T v_c)}{\sum_{k=1}^{|V|} exp(u_k^T v_c)}$$

Because this loss necessitates a summation over the entire vocabulary, negative sampling is used for computational efficiency. The negative sampling loss randomly selects a small set of negative (incorrect) context words for each center word and ensures that they have a low probability of occuring. Specifically, the loss is modified to:

$$J = -log(\sigma(u_{c-m+j}^T v_c)) - \sum_{k=1}^{K} log(\sigma(-\tilde{u_k}^T v_c))$$

where $-\tilde{u_k}$ is the negative context vector and $\sigma(x) = \frac{1}{1+e^{-x}}$.

Word vectors were trained on all of the data from "r/Liberal" and "r/Conservative", and individual months of data from "r/politics". While training on years instead of months would have provided more reliable vectors as a result of more data, years are too large to reflect the rapidly changing nature of current events and public opinion, which is vital for studying framing. Only words that occurred at least three times within any given month were included in training. Word vectors for controversial topics that contained multiple words were simply represented by the sum of their individual vectors, which empirically produced reliable results.

Once I obtained embeddings for each entity, I found the k nearest-neighbors for each controversial entity ent_i. The set of nearest neighbors for ent_i, $N(ent_i)$, is composed of the k words from the vocabulary, $\{w_1, ..., w_k\}$, that maximize the cosine-similarity $cossim(w_j, ent_i)$.

4 Experiment 2: Relationship Modeling Network

The data and "controversial entities" used in this experiment are the same as those used in the diachronic word embeddings approach. While the RMN could easily be extended to model framing trajectories in multiple communities, this paper focuses only on its use with the subreddit "r/politics".

4.1 Preprocessing

Posts and comments that discussed a controversial entity were extracted via keyword matching. For each topic, 1000 posts and comments were randomly sampled with the chronological order intact, giving us a corpus of 25,000 posts/comments. All punctuation, articles, and conjunctions were removed.Words that were among the top 10,000 most frequent words were kept; Wang et al. (2016) reported more reliable results by including this preprocessing step.

[4]https://radimrehurek.com/gensim/apiref.html

4.2 Model

My model closely follows that of Iyyer et al. (2016) and Want et al. (2016) but is adapted to the task of framing. The model essentially consists of a recurrent neural network that attempts to reconstruct vector representations of posts via a small set of embeddings of linguistic descriptors. Because the goal of this paper is to study a contentious entity's framing within a community, the RMN learns embeddings for both the communities and the entities. Learning these embeddings can account for "idiosyncrasies or superficial variation" in the language, and the model can thus use detect the variation that truly corresponds to different framings (Wang et al., 2016). Intuitively, we need to represent both the topic and the community in the RMN if we want to learn the community's relationship with the topic.

My dataset consists of posts that contain at least one mention of a controversial topic. More formally, for each controversial entity, e_i, and for each community, c_j, we have a chronologically ordered sequence of posts $P_{ij} = [p_{ij}^{(1)}, p_{ij}^{(2)}, ..., p_{ij}^{(n)}, ...]$. Because both posts and comments are treated the same by the model, these sequences of "posts" actually include the comments as well. After preprocessing, each post is a bag-of-words of variable length: $p_{ij}^{(n)} = [w_1, w_2, ..., w_l]$. Each word in a post is represented by its 300-dimensional GloVe vector, trained on Common Crawl data (CITE GLOVE). Posts are then represented as the average of each word's vector. For a post of length l, where $E_{word}[w_i]$ represents the GloVe embedding for word w_i:

$$v_{post} = \frac{1}{l}\sum_{i=1}^{l} E_{word}[w_i] \qquad v_{post} \in \mathbb{R}^{d_{word}}$$

During training, the model learns K relationship descriptors. These descriptors r_k for $k \in \{1, ..., K\}$ are also represented as d_{word}-dimensional vectors so that they could be interpreted by looking at the words with neighboring GloVe vectors.

For each post, we obtain vectors not only for the text, but also for the community, v_{comm}, and the controversial entity mentioned in the text, v_{ent}, via embedding matrices E_{comm} and E_{ent}. $E_{comm} \in \mathbb{R}^{N \times d_{comm}}$ and $E_{ent} \in \mathbb{R}^{M \times d_{ent}}$, where N is the number of communities and M is the number of entities ($M = 25$ here). Note that in this paper, $N = 1$ in order to allow for more in-depth analysis. These embedding matrices, and all weight matrices, are randomly initialized using the Xavier initializer, and are updated during training.
v_{post},v_{ent},and v_{comm} are then concatenated and passed through a linear layer followed by a ReLU nonlinearity. The hidden layer is thus defined as:

$$h_t = ReLU(W_h \cdot [v_{post}; v_{ent}; v_{comm}]) \qquad W_h \in \mathbb{R}^{(d_{word}+d_{ent}+d_{comm}) \times d_{hidden}}$$

A softmax is then computed over the hidden layer in order to score each of the possible descriptors for this post. However, in order to take previous posts about the same entity into account, I add a recurrent connection to previous score distributions. In particular, following Iyyer et al. (2016), I include a linear interpolation between the previous score distribution and the current hidden layer:

$$d_t = \alpha * softmax(W_d \cdot [h_t; d_{t-1}]) + (1 - \alpha) * d_{t-1} \qquad W_d \in \mathbb{R}^{(d_{hidden}+K) \times K}$$

For this project, I fixed α =0.5, but it could be learned by the model itself in future work. d_t represents the relationship state at time t. Then, the RMN learns the embeddings for the relationship descriptors, which are stored in a matrix R, where $R \in \mathbb{R}^{d_{word} \times K}$. At time t, the predicted descriptor is simply the row of R that corresponds to the argmax of d_t. The descriptor matrix is also initialized with the Xavier initializer. In order to train the RMN, I use R to create a reconstruction vector: $r_t = R^T d_t$. Thus, r_t is esentially a weighted average over all of the descriptor vectors. Because our goal is to have this weighted average, r_t, approximate the original post vector v_{post} (intuitively, capturing the same meaning), I incorporate the distance between r_t and v_{post} in the loss function.

The RMN parameters are updated via a contrastive max-margin loss function similar to the one that Iyyer et al. present. For a specific entity i and community j, the loss is computed over all posts in the sequence P_{ij}.

$$J_{ij}(\theta) = \sum_{t=0}^{|P_{ij}|} \sum_{n \in S} max(0, 1 - r_t v_{post} + r_t v_n)$$

In this formulation, S is a set of posts that are randomly sampled subset of posts from the entire dataset and v_n is the vector for negatively sampled post. This loss encourages parameter updates in order to maximize the inner product between the true post's vector and the reconstructed vector while simultaneously minimizing the inner product between the reconstructed vector and the negative samples. Note that all vectors are normalized before the loss is calculated.

In order to discourage learning descriptors that are too similar to each other, we add a penalty term to the loss function (Iyyer et al., 2016):

$$X(\theta) = ||RR^T - I||$$

where I is the identity matrix. Our final loss is thus: $L(\theta) = J(\theta) + \lambda X(\theta)$, where λ is a hyperparameter.

Each descriptor $r_k \in R$ has the same dimensions as the original word embeddings. Thus, descriptor vectors can be interpreted by the words whose vectors are their nearest semantic neighbors. These are determined by measuring cosine similarity in a manner similar to that discussed in Experiment 1. In order to make the linguistic descriptors even more interpretable, I added a constraint that the words must be within the 5000 most common English words [5]. Each descriptor vector was labeled with ten words. The full list of descriptors can be found in the appendix. For the visualizations, I usually label the descriptor with the first nearest word (except for a couple exceptions, where I use the second or third nearest word as a label).

Below is a diagram of the RMN architecture.

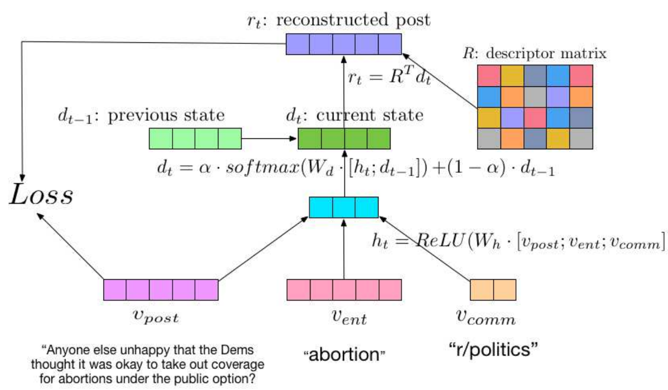

Figure 1: RMN architecture shown for one time step. For each post, the vectors for the post, entity, and community are concatenated and passed through a linear layer and softmax to get h_t. This and the previous time step are used to compute the vector d_t, a distribution over all descriptors that describes the relationship state at time t. The model is trained by maximizing the similarity between the post vector and the reconstruction vector, which is a linear combination of descriptors from the matrix R.

The hyperparameters used in this study are in the Appendix. Note that the number of epochs was varied from 15 to 50. While 50 epochs empirically yielded better descriptors, 15 epochs yielded more sensible trajectories.

5 Experiment 1 Results and Evaluation

The results for the diachronic word embedding can be divided into **five** underlying themes. Because of the highly qualitative and subjective nature of evaluating this experiment, it makes sense to analyze these themes within the context of specific examples. Although we look at just five examples, these trends hold throughout the results.

1. Concrete entities are associated with other concrete entities.

It is natural for concrete entities, such as people, to be associated with other people because word vectors tend to capture semantic similarity. In Figure 410 in the appendix, "clinton" is associated with other politicians and family members. This unfortunately means that neighboring words often do not provide **direct** insight into

[5] http://www.wordfrequency.info/free.asp

how a topic is framed within a community, since concrete entities such as names are not the optimal descriptors for these relationships.

2. Changes over time can simultaneously reflect current events and changes in framing.

Figure 510 in appendix shows how "refugees" was more closely associated with words related to the Palestinian refugee crisis. By May 2015, the associated words shifted dramatically, largely as a response to the dramatic rise in the number of Syrian refugees in 2014 and 2015. While the neighboring words do not explicitly tell us how a topic is framed, they provide crucial insights into framing shifts. Note that some of the words associated with "refugees" in May 2015 tend to be more malicious. This is likely to reflect a rising fear and suspicion of refugees in this sub-community.

3. Word embeddings reveal powerful framing shifts when combined with contextual knowledge.

Even though names tend to be associated with other names, the changes in these neighboring names can still reveal powerful framing shifts when combined with outside knowledge. For example, in Figure 6 in the appendix10, "putin" is shown to have become associated with infamous individuals. With contextual knowledge of who these people were, we can determine that "putin" is framed considerably more negatively in later data. Perhaps this change reflects that Putin has increasingly been framed as an authoritarian figure by this community.

4. Changes over time can reveal signs of social change and growing awareness.

Many social issues have been become much more widely discussed in public discourse in the past few years. For example, for each month in 2008 and 2009, the word "transgender" was either not in the vocabulary or only outputted nonsense neighbors (Figure 7 in the appendix10). In 2015, however, the nearest neighbors have been able to much more accurately (though not perfectly) capture this word. The increased use of "transgender" in political conversations indicates a growing awareness of transgender people and the social issues that they face.

5. Differences between sub-reddits in the entity's neighbors indicate that framing is dependent on the community.

There are clear differences when comparing how words are associated in a liberal sub-community and a conservative sub-community. In Figure 810 in the appendix, we can see that the two sub-communities share several neighbors for "obama". However, we also see insulting nicknames, which are a clear indication for framing. Most interesting here is the distinction between "administration" and "president". Liberals tended to describe their head of state as President Obama, while Conservatives seem to prefer phrases such as "the Obama Administration". It is encouraging that the Diachronic Word Embedding model is able to pick up on these rather subtle distinctions.

6 Experiment 2: Results and Evaluation

In the appendix is a sample of linguistic descriptors learned by the model. The descriptors tend to be a reasonable mixture of concrete political words, casual and emotional words, and adverbs used in constructing persuasive arguments ("substantially" etc.). All of these categories are relevant for my problem; this last category is especially promising because it indicates that the RMN is able to pick up on subtle aspects of how arguments are framed.

The trajectories below show the output descriptor for three "controversial topics" for each post in 1000 randomly sampled and chronologically ordered posts per topic. Each color represents a different descriptor, and each band represents one or more consecutive posts in the sequence. In some cases, the learned descriptors are surprisingly revealing of how a community relates to a controversial topic. The primary descriptor for "Clinton", namely "lobby", was especially striking. However, such clear results did not hold for all entities. In many cases, the model outputted primarily adverbs. While many adverbs play an important role in creating arguments, the RMN is not yet able to interpret these to make deeper claims about framing; this is one of the first steps in future work. Furthermore, the model appears to not yet have the ability to capture changes over time, as the ouput descriptors for each entity seem to be relatively constant throughout the entire sequence.

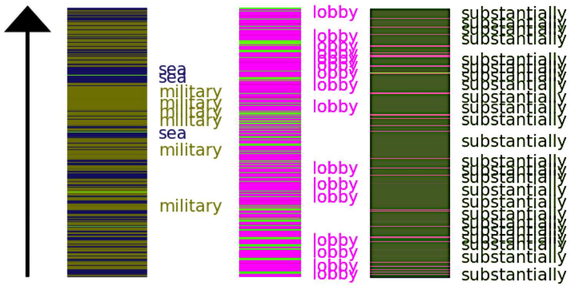

Figure 2: Trajectories of relationship between “r/politics” and “Syria”, “Clinton”, and “abortion”

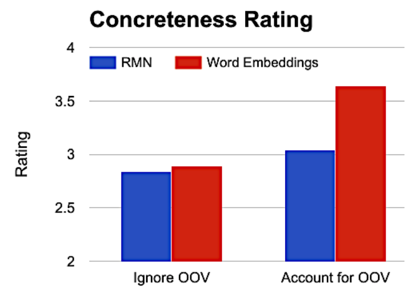

Figure 3: Concreteness Rating for RMN and Diachronic Word Embeddings.

One key issue with diachronic word embeddings was that they could not directly capture the relationship between an entity and a community because they often yielded words that are very concrete. Ideally, words that describe this relationship would be more abstract and thus more directly interepretable. I thus compared these two approaches by comparing their output words’ average concreteness rating by using the Concreteness Lexicon created by Brysbaert et al. (2014). The most abstract words have a concreteness rating close to 1, and the most concrete words (including people and places) has a concrete rating close to five.

There are a significant number of words that are outside of the lexicon. If these words are ignored, the RMN’s outputs are slightly more abstract on average than the word embeddings’ outputs. However, nearly all of the tokens that are outside of the lexicon (labeled OOV) are concrete entities, such as names and places. Thus, if I account for OOV words with a concreteness score of 4 (since estimates from a random sample reveal that slightly more than 80% of these entities are names and places), a more significant difference between the two approaches appears. The fact that the RMN is able to capture more abstract descriptors further suggests the value of this type of approach.

7 Conclusion and Future Work

Both approaches presented in this paper have their own advantages and drawbacks.

Diachronic word embeddings have several advantages: its descriptors are more clearly relevant to the controversial entity, and differences in an entity’s neighbors over time reveal dynamic framing. However, the main disadvantage is that the outputs tend to be concrete and cannot directly and fully describe a relationship; analysis with diachronic word embeddings requires substantial contextual knowledge.

The Relationship Modeling Network, on the other hand, does not require any context and is able to produce more abstract descriptors. It can also capture subtleties important for framing. However, the current model is not able to capture change over time, and several descriptors, though relevant to the task, tend to be difficult to interpret for social science analysis.

Future work on this project includes improving both models. Other word embedding techniques, such as WordRank, may work better on smaller data sets and produce better results. The RMN could benefit from a more sophisticated architecture and more robust hyperparameter tuning. Using word vectors trained on all Reddit data instead of GloVe may be helpful as well. Eventually, I hope to somehow combine these two approaches in order to draw upon both models’ strengths.

8 Acknowledgements

I thank my mentors, Ignacio Cases and Will Hamilton, for their guidance throughout this project, and Dan Jurafsky for pointing me to related work about framing. I would also like to thank Mohit Iyyer, whose starter code for the RMN was helpful. Finally, I would like to thank the CS 224N professors and TAs for their dedication to this class.

9 References

[1] Eric P. S. Baumer, Elisha Elovic, Ying Crystal Qin, Francesca Polletta, and Geri K. Gay. Testing and Comparing Computational Approaches for Identifying the Language of Framing in Political News. *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*, pages 1472–1482, 2015.

[2] Robert D Benford and David a Snow. FRAMING PROCESSES AND SOCIAL MOVEMENTS : An Overview and Assessment. *Annual Review Sociologic*, 26(1974):611–639, 2000.

[3] Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. Concreteness ratings for 40 thousand generally known English word lemmas. 2014.

[4] Dallas Card, Justin H Gross, Amber E Boydstun, and Noah A Smith. Analyzing Framing through the Casts of Characters in the News. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP-16)*, pages 1410–1420, 2016.

[5] Dennis Chong and James N. Druckman. Framing Theory. *Annual Review of Political Science*, 10(1):103–126, 2007.

[6] Robert M. Entman. Framing bias: Media in the distribution of power. *Journal of Communication*, 57(1):163–173, 2007.

[7] William L. Hamilton, Kevin Clark, Jure Leskovec, and Dan Jurafsky. Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora. page 15, 2016.

[8] William L Hamilton, Jure Leskovec, and Dan Jurafsky. Computational Measures of Semantic Change. 2016.

[9] William L Hamilton, Jure Leskovec, and Dan Jurafsky. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. 2016.

[10] Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-graber, and Hal Daum. Feuding Families and Former Friends : Unsupervised Learning for Dynamic Fictional Relationships. *North American Association for Computational Linguistics*, pages 1534–1544, 2016.

[11] Vivek Kulkarni, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. Statistically significant detection of linguistic change. *Proceedings of the 24th international conference on World Wide Web (WWW '15)*, pages 625–635, 2015.

[12] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. *Nips*, pages 1–9, 2013.

[13] Vincent Price, Lilach Nir, and Joseph N. Cappella. Framing public discussion of gay civil unions. *Public Opinion Quarterly*, 69(2):179–212, 2005.

[14] Jonathon P. Schuldt, Sara H. Konrath, and Norbert Schwarz. "Global warming" or "climate change"? *Public Opinion Quarterly*, 75(1):115–124, 2011.

[15] Oren Tsur, Dan Calacci, and David Lazer. A Frame of Mind: Using Statistical Models for Detection of Framing and Agenda Setting Campaigns. *Acl*, pages 1629–1638, 2015.

[16] Alex Wang, William L Hamilton, and Jure Leskovec. Learning Linguistic Descriptors of User Roles in Online Communities. pages 76–85, 2016.

10 Appendix

Table 1: "Controversial" topics

clinton	racism	obama	islam	pakistan
iran	immigration	russia	abortion	lgbt
obamacare	climate_change	feminist	putin	marijuana
israel_palestine	marriage_equality	death_penalty	gender	mexico
refugee	syria	prostitution	vaccine	transgender

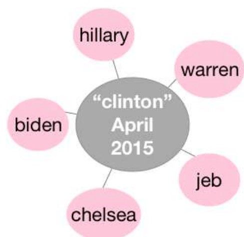

Figure 4: Neighbors of "clinton" in April 2015

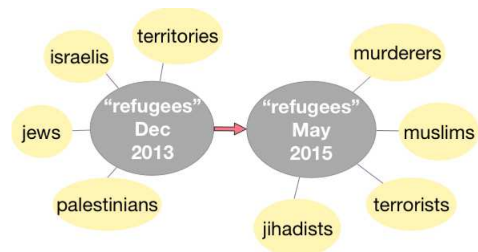

Figure 5: Neighbors of "refugees" in December 2013 and May 2015

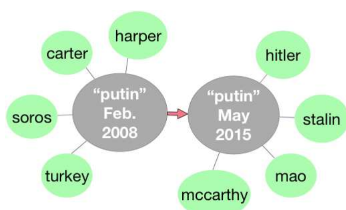

Figure 6: Neighbors of "putin" in February 2008 and May 2015

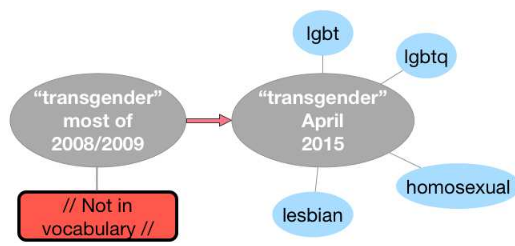

Figure 7: Neighbors of "transgender" in 2008/2009 and April 2015

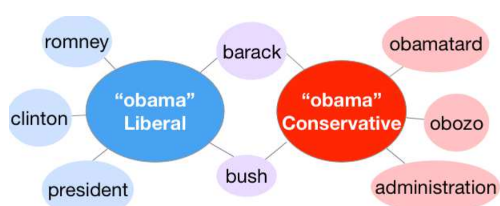

Figure 8: Neighbors of "obama" in "r/Liberal" and "r/Conservative"

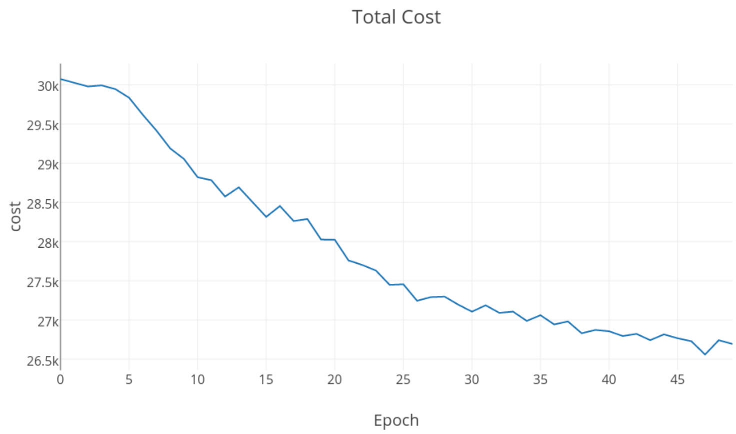

Figure 9: Total training cost per epoch

Table 2: Sample of Descriptor Outputs from RMN

Desc. ID				
03	Republican	senator	Democrat	Legislator
07	wow	cute	shit	huh
08	substantially	significantly	normally	considerably
16	perceive	irony	disturbing	fault

Table 3: RMN Hyperparameters

Parameter	Value
d_post	300
d_ent	300
d_comm	50
d_hidden	50
# descriptors	20
# neg. samples	30
# epochs	15-50
learning rate	1e-3
λ	1e-3
α	0.5