Honors Project Proposal

Kenneth Sanders

Semesters: 2017 Summer (Sessions 1 and 2)

Advisors: Anna Rumshisky, Fred Martin

Courses: Honors Project 1, 2 Directed Studies

### **Abstract**

Social media has given rise to a fast and impulsive platform to voice personal opinions. Twitter offers a unique platform for people to voice their opinions and listen to others. However, since tweets are so short, the message must be distilled to its simplest form. This makes it ideal for language processing in this aspect, since opinions are shared as bluntly and directly with very little implications or sarcasm. This project aims to use these opinions to construct a way of measuring the amount of conflict between two opposing groups of hashtags. This will be done by vectorising words, and using them to produce hashtag vectors which can be used to represent the similarity between two hashtags, which in turn will allow for an analysis of the polarization of the conflict.

### **Overview**

The goal of this project is to develop a generalized procedure and program to measure the magnitude of conflicts using data from Twitter, specifically oppositional hashtags. The procedure will be developed on topics which have a notable division of beliefs in society, such as political elections and disputed topics, and then tested on murkier subjects in which divisions of beliefs are not as well defined. This will allow us to gauge the effectiveness of the procedure in relation to intuitive views of conflict. Additionally, since studies have already been done to produce similar procedures, results from our procedure can be compared to others for analysis of effectiveness.

The first step of the project will aim to create a measurement between two hashtags to express their similarity. This will be done in a base series of steps, which are as follows:

1. Carry out word2vec in order to get vectors on which cosine similarity could be used

2. For each hashtag, sum all word vectors each time a word appears in the same tweet as a hashtag.

3. Normalize the new vectors created for the hashtag

Word2vec is an algorithm consisting of a simple neural network with noise contrastive emulation, which was formulated by Mikolov et al. This allows us to create vector representations of each word incorporated into our vocabulary. Intuitively, much of the meaning of hashtags can be ascribed to the words that they are used with. So, it makes sense that the vector representations of the words that are used with hashtags could potentially be summed in order to gain an at least approximate representation of the relationships between the hashtags. This in turn can be used to generate a graph of relations between the entire dataset of hashtags, which can then be clustered and analyzed.

Once this hashtag graph is created, and cluster analysis has been carried out, it can also be compared to a graph composed from the word vectors themselves in order to measure the accuracy of hashtags in relation to the word vectors they are derived from. The hashtag vectors could also be created within certain time periods in order to chart the lexical drift, specificity, and volatility, as well as to gain more information about the communities of hashtags in relation to each other.

### **Deliverable**

The product of this project will be in the format of an IEEE conference manuscript, as well as a more detailed paper on the topic that describes each step taken in detail, and references the code created. The code base for the project will be established to aid in computations theorized in the manuscript, including ‘hashtag2vec’ and functions to analyze the resulting vectors. The program to actually analyze the data will be written in Python, and will most likely implement the TensorFlow, NLTK, and NumPy libraries in particular. The data will be input as a PSQL database due to the large size of the dataset. The intermediate variables, such as word vectors, dictionary, and formatted tweets will also be cached in PSQL tables due to their somewhat long creation time. This will also ease debugging and allow improvements to be made without rebuilding the intermediate variables that would be unaffected. Assuming there is enough time, a program to aid in visualizing the hashtag vectors will also be created.

### **Schedule**

|  |  |
| --- | --- |
| Week | Task |
| 5/15/17 | Literature analysis |
| 5/22/17 | Plan procedure and code outline |
| 5/29/17 | Build code structure – objects and main |
| 6/5/17 | Finalize code |
| 6/12/17 | Apply code to topics and analyze accuracy |
| 6/19/17 | Revisit program and procedure for improvements |
| 6/26/17 | Finalize procedure and program |
| 7/3/17 | Draft paper & manuscript |
| 7/10/17 | Finish paper & manuscript |

### **Working Bibliography**

Mikolov, Tomas and Sutskever, Ilya and Chen, Kai and Corrado, Greg S and Dean, Jeff. “Distributed Representations of Words and Phrases and their Compositionality”, *Advances in Neural Information Processing Systems 26*. Eds. C. J. C. Burges and L. Bottou and M. Welling and Z. Ghahramani and K. Q. Weinberger Curran Associates, Inc., 2013, 3111--3119. Web.<http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>

This paper details the word2vec algorithm which is critical to the project, since the hashtag vectors are based off of the word vectors. The word2vec algorithm is based on the skip-gram model of words being input as training into a simple neural net, from which a vector and associated vector space can be extracted.

Conover, M. D. et al. "Political Polarization On Twitter". Fifth International AAAI Conference On Weblogs And Social Media. Bloomington: Indiana University. Web. 5 May 2017. <<http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2847/3275.pdf>>

This paper discusses interactions on Twitter and how to extract a measurement of polarization by analyzing social network structure, especially in a political environment. Since polarization of Twitter networks is integral to the analysis of the effectiveness of the algorithm I hope to produce.

Sarita Yardi and Danah Boyd, Dynamic debates: An analysis of group polarization over time on twitter, Bulletin of Science, Technology & Society 30 (2010), no. 5, 316–327. Web. 5 May 2017. <<http://journals.sagepub.com.libproxy.uml.edu/doi/pdf/10.1177/0270467610380011>>

This is an analysis of debates on twitter and the group dynamics that arise from them. This will help in the hashtag vector analysis phase during which clustering and interpretation of the results may be aided by an understanding of group theory.

Liu, Can et al. “IUCL at SemEval-2016 Task 6: An Ensemble Model for Stance Detection in Twitter.” SemEval@NAACL-HLT (2016). Web. 19 May 2017. <<http://www.aclweb.org/anthology/S/S16/S16-1064.pdf>>

This paper on stance detection on Twitter is similar to my project in the central idea of detecting on which side of an argument people stand. However, my project focuses on the big picture of hashtags as opposed to the individual user in this paper. It still may be very relevant in the way that it analyzes a stance, and such an algorithm may be modified for use with hashtags to aid in my project.

W. B. Claster, M. Cooper and P. Sallis, "Thailand -- Tourism and Conflict: Modeling Sentiment from Twitter Tweets Using Naïve Bayes and Unsupervised Artificial Neural Nets," *2010 Second International Conference on Computational Intelligence, Modelling and Simulation*, Bali, 2010, pp. 89-94. Web. 19 May 2017. <<http://ieeexplore.ieee.org.libproxy.uml.edu/stamp/stamp.jsp?tp=&arnumber=5701826&isnumber=5701812&tag=1>>

This paper describes methods of analysis of sentiment in tweets, which I may find a way to include in the weighting of the networks to be constructed. The analysis of sentiments towards a subject certainly seems like it could be applicable in the summation of word vectors into hashtag vectors.

TensorFlow. Google Inc., n.d. Web. 26 Apr. 2017 <<https://www.tensorflow.org/>>

This is a very comprehensive mathematical library which includes many of the sub-algorithms which word2vec will require. Additionally, the library will probably supply other functions that will be useful in the analysis portion of my project.

"NumPy." NumPy Developers, n.d. Web. 19 May 2017. <<http://www.numpy.org/>>.

Another library which is heavily relied upon by the TensorFlow library, this provides many of the base classes used for data processing in python, such as vectors and multidimensional arrays. Classes from this library will also be used to store the data from the database in memory, and allows for efficient linear algebra operations on them.

"Natural Language Toolkit." NLTK Project, n.d. Web. 19 May 2017. <<http://www.nltk.org/>>.

The Natural Language Toolkit is a fairly extensive python library that provides many tools such as twitter tokenization and stop words. This will be integral to data pre-processing to deal with the inconsistency of language on Twitter, so that word2vec can be effective.

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. "Quantifying Controversy in Social Media." *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining - WSDM 16*(2016): n. pag. Web. 5 May 2017. <<https://arxiv.org/pdf/1507.05224.pdf>>

This paper details methods of controversy detection on Twitter through network analysis. This is very important to my analysis of the hashtag vectors and the weights they allow me to add to the social networks I will be analyzing

Thoughts:

1. Format and tokenize tweets - lexical analysis

a. Good starting tokenizer:<http://www.nltk.org/api/nltk.tokenize.html>

b. Integrate other syntax filters and stop word filters (when counting words only)

c. Tokens converted to integers for word2vec

2. Use word2vec to create a vectors representing relations of words in our tweets

3. Graph hashtags based on their simultaneous use in a tweet or a user using both hashtags in the same tweets (other ways?)

4. Create hashtag vector spaces by summing over the vector spaces of the word vectors and normalizing

a. This might be improved by taking into account specificity and volatility of words, which we could then obtain distinctiveness and dynamicity from for each hashtag ( another weight in the graph?) see<https://cs.stanford.edu/people/jure/pubs/identity-icwsm17.pdf>

5. Find cosine of vector spaces between two hashtags to find the weight of an edge in graph from 3.)

6. analyze the now weighted graph(s) of hashtags and look for clusters and signs of conflict, possibly through<http://www.mapequation.org/publications.html#Rosvall-Axelsson-Bergstrom-2009-Map-equation>

7. Compare this to conventional graphing by user interest: ie, looking at vocabulary of users and finding a relationship between them with word2vec profiles

Other things:

Any way to incorporate analysis of

Similar papers

http://emnlp2014.org/papers/pdf/EMNLP2014194.pdf

Methodology

Neural networking – specifically TensorFlow - will be used to analyze the probability of the sentiment of a tweet in relation to a topic. This network will be supervised while being trained on tweet data already gathered on a subject, and allowed to work freely on new data in order to produce measurable results. The neural network will be structured such that the meaning of a tweet is determined by the relationship of keywords.

Once a network is produced and the data is interpreted by the neural network, cluster analysis be used on the networked tweets, hashtags and accounts. This will allow comparison to results from other studies, as well as general knowledge analysis and other measures of conflict, such as polls and expert predictions.

Project Rationale and Justification:

Twitter creates a fantastic platform for impulsive commentary and reactions. This in turn implies that there is a blunt truthfulness to conflicting tweets. The aim of this project in developing a process to process opposing twitter groups by analyzing their hashtags and tweets and generating of graphs relationships between accounts, hashtags, and the keywords contained in the tweets that use hashtags. Further understanding the trends of debated topics in the world allow us to predict and use the popularity of these topics in many industries such as advertising, investments, and lobbying.