#MAGA or #TheResistance: Classifying Twitter users' political affiliation without looking at their words or friends

Rachael Tatman
University of Washington
Linguistics Department
rctatman@uw.edu

Abstract

Previous NLP work attempting to classify Twitter users' political affiliation focused on lexical features and information users' social networks. However, sociolinguistic work on the expression of political identity has found that sub-lexical features also reflic political identity. This study shows that it is possible to achieve above-chance classification of Twitter users based on only three features: how much punctuation and capitalization an account uses and what proportion of their tweets are replies.

1 Introduction

There is a growing body of work in Natural Language Processing on classifying social media users' political affiliation. To date, the bulk of this work has relied on lexical and social network features. Conover et al. used uni-grams (excluding punctuation) and social networks (Conover et al., 2011), while Cohen and Ruths used a large feature set including words, stems, di- and tri-grams and hashtags (Cohen and Ruths, 2013). Sylwester and Purver, who were interested in characterizing psychological differences between Democrats and Republicans, focused on word frequency, friend-follower ratio and Linguistic Inquiry and Word Count (Pennebaker et al., 2001)–although they also excluded punctuation from their data.

Non-lexical features, however, have proven useful in other NLP tasks. Punctuation in particular is a well-established linguistic variable. It's been used in a variety of tasks, including authorship identification (Chaski, 2005; Abbasi and Chen, 2005) and predicting gender (Bamman et al., 2012) and personality (Pennebaker et al., 2015; Golbeck et al., 2011). While punctuation is less well studied in sociolinguistics, there is some evidence that patterns of punctuation use are associated with group membership (Ongonda et al., 2010). There is also evidence that variation capitalization is an important stylistic feature in informal computer-mediated communication (Ling, 2005). This fits in well with sociolinguistic investigations of how political affiliation is expressed in speech, which has also focused on non-lexical feature like phonetic variation (Hall-Lew et al., 2010; Kirkham and Moore, 2016).

Given that non-lexical features have proven informative in characterizing political affiliation in sociolinguistic research, are they also useful for automatic classification?

2 Data

In order to classify users' political affiliation, it was necessary to construct a data set which contained a politically heterogeneous mixture of users and where users' political affiliation was known. To do this, tweets were collected from users who expressed affiliation with either #MAGA [1] or #TheResistance in their bios or user names. (Given that political hashtags are often co-opted by groups opposing their original creators in order to redirect the conversatino (Booten, 2016), labelling users based on their use of hashtags in tweets is a poor sampling strategy.)

#MAGA is an abbreviation of "Make America Great Again". This was the campaign slogan of the Trump presidential campaign, and has been widely adopted by the conservative political community (Starnes, 2017). #TheResistance is a hashtag associated with membership in a political movement formed to oppose the Trump presidency (Blow, 2017), and as a result indexes a political affiliation in direct opposition to that of

[1] This did return a small number of Twitter accounts belonging to Krav Maga enthusiasts.

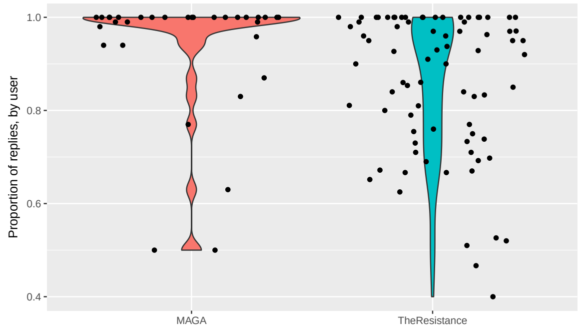

Figure 1: Proportion of user's tweets that are replies, by group. #MAGA users replied much more often #TheResistance users.

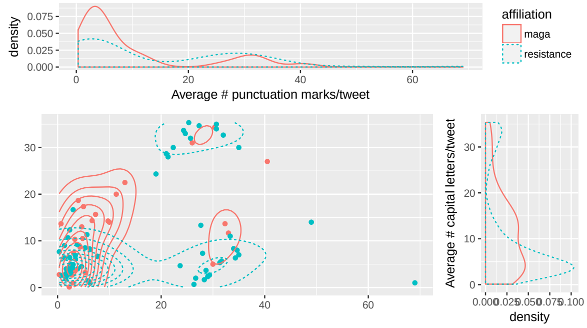

Figure 2: Distribution of users' over their average number of punctuation marks and capital letters per tweet.

#MAGA users.

This sampling returned 110 accounts, 32 associated with #MAGA and 78 associated with #TheResisitance. Up to 100 of the most recent from each account were collected using the Twitter public streaming API on April 11, 2017. This resulted in a collection of 6992 total tweets, 1925 from #MAGA accounts, and 5067 from #TheResistance.

For each account, the average number of punctuation marks and capitalized letters per tweet and the proportion of tweets which were re-tweets and replies was calculated.

3 Modelling and Classification

In order to determine whether these features were associated with users' political affiliation, a logistic regression was conducted with affiliation (#MAGA or #TheResistance) as the dependent variable, and capitalization, punctuation, re-tweets and replies as independent variables. While the proportion of a users' tweets which were re-tweets was not significantly associated with affiliation, the other three variables were. Proportion of replies, average number of punctuation characters per tweet and average number of capital letters per tweet were all significantly ($P < 0.05$) associated with political affiliation. #MAGA users replied more often, used less punctuation and used more capital letters than #TheResistance users, as can be seen in Figures 1 and 2.

But are these three robust features enough to be used for classification? In order to determine this, a KNN classifier was built using the R 'class' package (Venables and Ripley, 2002). Training and testing was done using 80/20 cross-validation on a balanced subset of an equal number of #MAGA and #TheResistance users (32 each, randomly sampled without replacement in the case of #TheResistance users).

While the model presented here did not beat the state-of-the-art (see Table 1), it did correctly classify 78% of the test data using only three features.

Table 1: Though this model did not beat the state-of-the art, it achieved classification well above chance using only three non-lexical features.

Study	**Accuracy**
Conover	87%
Cohen (politically active accounts)	84%
KNN classifier (this study)	78%

4 Conclusion

This study shows that it is possible to classify Twitter users' political affiliation well above chance without using lexical or social-network features. It provides clear evidence that researchers looking at political affiliation should reconsider stripping punctuation from tweets, as they contain useful information on community norms.

Further work is necessary to determine whether the features discussed here (how often a user replies and how much capitalization and punctuation they use) vary with any other social factors such as age, socioeconomic status, level of education or gender.

It should be noted that these accounts explicitly express political affiliation, as so are are likely to be politically active. Previous work has found that politically active accounts are the easiest to classify (Cohen and Ruths, 2013); these features may not be as discriminative for less political accounts.

References

Ahmed Abbasi and Hsinchun Chen. 2005. Applying authorship analysis to extremist-group web forum messages. *IEEE Intelligent Systems* 20(5):67–75.

David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2012. Gender in twitter: Styles, stances, and social networks. *CoRR abs/1210.4567* .

Charles M. Blow. 2017. The opinion pages: Resilience of the resistance. *The New York Times* https://www.nytimes.com/2017/04/24/opinion/resilience-of-the-resistance-donald-trump.html?$_r = 0$.

Kyle Booten. 2016. Hashtag drift: Tracing the evolving uses of political hashtags over time. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, pages 2401–2405.

Carole E Chaski. 2005. Whos at the keyboard? authorship attribution in digital evidence investigations. *International journal of digital evidence* 4(1):1–13.

Raviv Cohen and Derek Ruths. 2013. Classifying political orientation on twitter: It's not easy! In *ICWSM*.

Michael D. Conover, Bruno Gonçalves, Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. 2011. Predicting the political alignment of twitter users. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*. IEEE, pages 192–199.

Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting personality from twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*. IEEE, pages 149–156.

Lauren Hall-Lew, Elizabeth Coppock, and Rebecca L Starr. 2010. Indexing political persuasion: variation in the iraq vowels. *American Speech* 85(1):91–102.

Sam Kirkham and Emma Moore. 2016. Constructing social meaning in political discourse: Phonetic variation and verb processes in ed miliband's speeches. *Language in Society* 45(01):87–111.

Rich Ling. 2005. The sociolinguistics of sms: An analysis of sms use by a random sample of norwegians. In *Mobile communications*, Springer, pages 335–349.

Nancy Anashia Ongonda, Peter Maina Matu, and Omondi Oketch. 2010. Punctuation as a sociolinguistic variable in text messages. *Sino-US English Teaching* 7(9):42–47.

James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of liwc2015. Technical report.

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

Todd Starnes. 2017. *The Deplorables' Guide to Making America Great Again*. Charisma Media.

Karolina Sylwester and Matthew Purver. 2015. Twitter language use reflects psychological differences between democrats and republicans. *PloS one* 10(9):e0137422.

William N. Venables and Brian D. Ripley. 2002. *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0. http://www.stats.ox.ac.uk/pub/MASS4.