Beyond Binary Labels: Political Ideology Prediction of Twitter Users

Daniel Preoțiuc-Pietro

Joint work with Ye Liu (NUS), Daniel J Hopkins (Political Science), Lyle Ungar (CS)

2 August 2017
User attribute prediction from text is successful:

- **Age** (Rao et al. 2010 ACL)
- **Gender** (Burger et al. 2011 EMNLP)
- **Location** (Eisenstein et al. 2010 EMNLP)
- **Personality** (Schwartz et al. 2013 PLoS One)
- **Impact** (Lampos et al. 2014 EACL)
- **Political Orientation** (Volkova et al. 2014 ACL)
- **Mental Illness** (Coppersmith et al. 2014 ACL)
- **Occupation** (Preotiu-Pietro et al. 2015 ACL)
- **Income** (Preotiu-Pietro et al. 2015 PLoS One)

... and useful in many applications.
Hypothesis:
Political ideology of a user is disclosed through language use

- partisan political mentions or issues
  
  @realDonaldTrump your program last night was top notch! You Sir, are a class act! God bless our Vets #MakeAmericaGreatAgain #Trump2016

- cultural differences
  Disappointed today. Either I trust God to "have this" or I don't. I truly do, but still disappointed.
Previous CS/NLP research used data sets with user labels identified through:

1. User descriptions

H1 Users are far more likely to be politically engaged
2. Partisan Hashtags

H2 The prediction problem was so far over-simplified
3. Lists of Conservative/Liberal users

H3 Neutral users
4. Followers of partisan accounts

H4 Differences in language use exist between moderate and extreme users
Data

- Political ideology
  - specific of country and culture
  - our use case is US politics (similar to all previous work)
  - the major US ideology spectrum is Conservative – Liberal
  - seven point scale
Data

We collect a new data set:

- 3.938 users (4.8M tweets)
- public Twitter handle with >100 posts

Political ideology is reported through an online survey

- only way to obtain unbiased ground truth labels (Flekova et al. 2016 ACL, Carpenter et al. 2016 SPPS)
- additionally reported age, gender and other demographics
Data

- Data available at preotiuc.ro
  - full data for research purposes
  - aggregate for replicability

- Twitter Developer Agreement & Policy VII.A4
  
  "Twitter Content, and information derived from Twitter Content, may not be used by, or knowingly displayed, distributed, or otherwise made available to any entity to target, segment, or profile individuals based on [...] political affiliation or beliefs"

- Study approved by the Internal Review Board (IRB) of the University of Pennsylvania
Class Distribution

- V.Con. (1): 195
- Conservative (2): 401
- Mod.Con. (3): 453
- Neutral (4): 696
- Mod.Lib. (5): 501
- Liberal (6): 692
- V.Lib. (7): 594
Data

For comparison to previous work, we collect a data set:

- 13.651 users (25.5M tweets)
- follow liberal/conservative politicians on Twitter
H1 Previous studies used users far more likely to be politically engaged

H2 The prediction problem was so far over-simplified

H3 Neutral users can be identified

H4 Differences in language use exist between moderate and extreme users
H1 Previous studies used users far more likely to be politically engaged

Manually coded:

- Political words (234)
- Political NEs: mentions of politician proper names (39)
- Media NEs: mentions of political media sources and pundints (20)
## Engagement

Data set obtained using previous methods

<table>
<thead>
<tr>
<th>Political word usage across user groups</th>
<th>Media/Pundit Names</th>
<th>Politician Names</th>
<th>Political Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con. Follow</td>
<td>2.64</td>
<td>0.73</td>
<td>0.11</td>
</tr>
<tr>
<td>Lib. Follow</td>
<td>2.95</td>
<td>0.79</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Average percentage of political word usage
Engagement

Our data set

<table>
<thead>
<tr>
<th></th>
<th>Media/Pundit Names</th>
<th>Politician Names</th>
<th>Political Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.11</td>
<td>0.73</td>
<td>2.64</td>
</tr>
<tr>
<td></td>
<td>0.24</td>
<td>0.14</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.07</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.09</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>0.12</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>0.19</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>0.79</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Political word usage across user groups

Average percentage of political word usage
## Engagement

Our data set

### Political word usage across user groups

- **Media/Pundit Names**
- **Politician Names**
- **Political Words**

<table>
<thead>
<tr>
<th>User Group</th>
<th>Con. Follow</th>
<th>V.Con.(1)</th>
<th>Con.(2)</th>
<th>M.Con.(3)</th>
<th>Mod.(4)</th>
<th>M.Lib.(5)</th>
<th>Lib.(6)</th>
<th>V.Lib.(7)</th>
<th>Lib. Follow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.64</td>
<td>0.76</td>
<td>0.55</td>
<td>0.42</td>
<td>0.36</td>
<td>0.46</td>
<td>0.76</td>
<td>2.95</td>
<td>0.79</td>
</tr>
</tbody>
</table>

**Average percentage of political word usage**
Engagement

Take aways:

- 3x more political terms for automatically identified users compared to the highest survey-based scores
- Almost perfectly symmetrical U-shape across all three types of political terms
- The difference between 1-2/6-7 is larger than 2-3/5-6
Hypotheses

H1 Previous studies used users far more likely to be politically engaged

H2 The prediction problem was so far over-simplified

H3 Neutral users can be identified

H4 Differences in language use exist between moderate and extreme users
H2 The prediction problem was so far over-simplified

ROC AUC, Logistic Regression, 10-fold cross-validation
H2 The prediction problem was so far over-simplified
H2 The prediction problem was so far over-simplified

ROC AUC, Logistic Regression, 10-fold cross-validation
**H2** The prediction problem was so far over-simplified

ROC AUC, Logistic Regression, 10 fold-cross validation
Over-simplification

Predicting continuous political leaning (1 – 7)

Pearson R between predictions and true labels, Linear Regression, 10-fold cross-validation
Over-simplification

Seven-class classification

Accuracy, 10-fold cross-validation

GR – Logistic regression with Group Lasso regularisation
Hypotheses

H1 Previous studies used users far more likely to be politically engaged

H2 The prediction problem was so far over-simplified

H3 Neutral users can be identified

H4 Differences in language use exist between moderate and extreme users
Neutral Users

H3 Neutral users can be identified

Words associated with either extreme conservative or liberal

Words associated with neutral users

Correlations are age and gender controlled. Extreme groups are combined using matched age and gender distributions.
H3a There is a separate dimension of political engagement

Combine the classes into a scale: 4 – 3 & 5 – 2 & 6 – 1 & 7

Pearson R between predictions and true labels, Linear Regression, 10 fold-cross validation
Hypotheses

H1 Previous studies used users far more likely to be politically engaged

H2 The prediction problem was so far over-simplified

H3 Neutral users can be identified

H4 Differences in language use exist between moderate and extreme users
Moderate Users

H4 Differences between moderate and extreme users

Words associated with moderate liberals (5 and 6).

Words associated with extreme liberals (7).

Correlations are age and gender controlled.
User-level trait acquisition methodologies can generate non-representative samples

Political ideology:
- Goes beyond binary classes
- The problem was to date over-simplified
- New data set available for research
- New model to identify political leaning and engagement
Questions?

www.preotiuc.ro

wwbp.org