Political Speech/Text

Detecting bias and partisanship in political speeches and blogs

Julie Medero

UW-EE

March 1, 2011
Today’s Papers

Feature Selection and Evaluation for Identifying the Content of Political Conflict
- *Burt Monroe, Michael Colaresi, Kevin Quinn (political scientists)*
- Survey methods of lexical feature selection and evaluation
- Qualitative comparison of output features

Shedding (a Thousand Points of) Light on Biased Language
- *Tae Yano, Philip Resnik, Noah Smith (computer scientists)*
- Use feature selection to target sentences for human annotation
- Quantitative analysis of annotation distributions
Outline

1. Feature Selection and Evaluation Methods

2. Applications
   - Toy Examples
   - Analyzing Bias in Blogs

3. Discussion
Outline

1 Feature Selection and Evaluation Methods

2 Applications
   • Toy Examples
   • Analyzing Bias in Blogs

3 Discussion
### Feature Selection and Evaluation

<table>
<thead>
<tr>
<th>Difference Between Selection and Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feature Selection</strong></td>
</tr>
<tr>
<td><strong>Feature Evaluation</strong></td>
</tr>
</tbody>
</table>
Feature Selection and Evaluation

<table>
<thead>
<tr>
<th>Difference Between Selection and Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Selection</td>
</tr>
<tr>
<td>Feature Evaluation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying partisan words in Congressional speeches on a topic</td>
</tr>
</tbody>
</table>
Feature Selection and Evaluation

**Difference Between Selection and Evaluation**

Feature Selection  Binary, in-or-out Decision  
Feature Evaluation  Quantify usefulness of features (e.g. weighting)

**Task Description**

Identifying partisan words in Congressional speeches on a topic

**Three Ways to Frame the Problem**

- Classification Task
- Non-Model Approaches
- Model Approaches
### Approach Summary

- Use standard ML tools (SVM, AdaBoost, Random Forests)
- Find words that predict partisanship: $c : w \rightarrow p$

### Problems

- Partisanship is not a function of word choice – it’s (if anything) the other way around
- Treating partisanship as an unknown doesn’t make sense – it’s observable
## Non-Model Approaches

### Methods

- (Normalized) Frequency Difference
- (Log) Odds Ratio
- TF-IDF
- WordScore

### Problems

- Over-emphasis on high-frequency words
- Stop lists can take out useful words, leave high-frequency non-function words (e.g. Senate)
- Over-emphasis on low-frequency words
  - Frequency threshold just pushes problem off to words just over the threshold
- Smoothing
  - Need to handle case of zero count in one set
Non-Model Approaches

Methods

- (Normalized) Frequency Difference
- (Log) Odds Ratio
- TF-IDF
- WordScore

Problems

- Over-emphasis on high-frequency words
  Stop lists can take out useful words, leave high-frequency non-function words (e.g. Senate)
Non-Model Approaches

Methods

- (Normalized) Frequency Difference \( H \)
- (Log) Odds Ratio \( L \)
- TF-IDF
- WordScore \( L \), \( H \) (normalized)

Problems

- \( H \) Over-emphasis on high-frequency words
  Stop lists can take out useful words, leave high-frequency non-function words (e.g. Senate)
- \( L \) Over-emphasis on low-frequency words
  Frequency threshold just pushes problem off to words just over the threshold
# Non-Model Approaches

## Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>H</th>
<th>L</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Normalized) Frequency Difference</td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td>(Log) Odds Ratio</td>
<td>L</td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>TF-IDF</td>
<td></td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>WordScore (normalized)</td>
<td>L</td>
<td>H</td>
<td></td>
</tr>
</tbody>
</table>

## Problems

<table>
<thead>
<tr>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H</strong> Over-emphasis on high-frequency words</td>
</tr>
<tr>
<td>Over-emphasis on high-frequency non-function words (e.g. Senate)</td>
</tr>
<tr>
<td><strong>L</strong> Over-emphasis on low-frequency words</td>
</tr>
<tr>
<td>Frequency threshold just pushes problem off to words just over the threshold</td>
</tr>
<tr>
<td><strong>S</strong> Smoothing</td>
</tr>
<tr>
<td>Need to handle case of zero count in one set</td>
</tr>
</tbody>
</table>

Julie Medero (UW-EE)  
Political Speech/Text  
March 1, 2011
Model-Based Approach

Generative model of word choice, $P(w|p)$

- Assume a multinomial distribution
- ML Estimate of odds ratio is the smoothed log odds-ratio
- Bonus: generative model means we can estimate (and scale by) sample variance
Model-Based Approach

Generative model of word choice, $P(w|p)$
- Assume a multinomial distribution
- ML Estimate of odds ratio is the smoothed log odds-ratio
- Bonus: generative model means we can estimate (and scale by) sample variance

Choosing a Prior
- An informed prior can help keep function words from being weighted too heavily
- Dirichlet prior works well in practice and yields closed-form solutions
- Laplace priors force more parameters to zero, but makes computation messy.
Outline

1. Feature Selection and Evaluation Methods

2. Applications
   - Toy Examples
   - Analyzing Bias in Blogs

3. Discussion
Gender Differences in Senate Speeches

Gender–Predicted Words, Abortion, 106th Congress

- Female Democrats
  - women
  - militari
- Male Democrats
  - pregnanc
  - mother

Frequency of Word within Topic
Partisanship of "vote", Judicial Nominations, 106th Congress

- Word is Democratic
  - Democrats demand votes on Clinton nominees

- Word is Republican
  - Republicans demand votes on Bush nominees

Polarization of Issues Over Time

Polarization, 106th Congress

Year

1997 1999 2001 2003 2005

Budget
Education
Defense
Sports
No child left behind
Invasion of Iraq
Post-9/11

Var_W(\hat{\beta}_{RW})
What is Bias?

- Tendency or preference towards a particular perspective, ideology or result
- Different from subjectivity
Analyzing Bias in Blogs

What is Bias?
- Tendency or preference towards a particular perspective, ideology or result
- Different from subjectivity

Task Description
- Look at distribution of biased sentences in liberal and conservative blogs
- Base labels on human annotations
- Use word feature selection to pick sentences likely to have bias
Sources

Extracted 260k sentences from 13k blog posts:
- Liberal: Digby, Think Progress, Talking Points Memo
- Conservative: American Thinker, Hot Air, Michelle Malkin
Data Set

Sources
Extracted 260k sentences from 13k blog posts:
- Liberal: Digby, Think Progress, Talking Points Memo
- Conservative: American Thinker, Hot Air, Michelle Malkin

Sentence Selection
- Limit to 8-40 tokens long
- Choose sentences from 3 categories:
  - Partisan bigrams selected using log-likelihood ratio + stop word list
  - Pennebaker emotion words (Negative Emotion, Positive Emotion, Causation, Anger)
  - Kill verbs (kill, slaughter, assassinate, shoot, poison, strangle, smother, choke, drown, suffocate, starve)
Annotators

Amazon Mechanical Turk

- U.S. citizens with 90% task approval rating
- Avg. pairwise $\kappa = .55, .50$ for 50 most frequent workers
- One annotator discarded for low agreement
- Paid $0.02$-$0.04$ per sentence (total cost $212$)
Annotators

Amazon Mechanical Turk

- U.S. citizens with 90% task approval rating
- Avg. pairwise $\kappa = .55, .50$ for 50 most frequent workers
- One annotator discarded for low agreement
- Paid $0.02$-$0.04$ per sentence (total cost $212$)

User Information Collected

- Who voted for in 2008 presidential election?
- Liberal or conservative for social issues?
- Liberal or conservative for fiscal issues?
Label political leanings

Please read the following sentence.

Jason Horowitz sums up the problem succinctly: Obama wanted nothing to do with the Clintons.

Do you think this sentence shows the author's preference towards a particular position (bias) in politics? If so, to what extent does this sentence display political bias?

- I don't see any bias
- Somewhat biased
- Very much biased

If there is a bias, to what extent is it liberal or conservative? Please check 'No obvious bias' if the sentence has no obvious bias.

- Very liberal
- Moderately liberal
- I don't see any bias
- Moderately conservative
- Very conservative
- I think it is biased, but am not sure which side

☐ 1: Jason
☐ 2: Horowitz
☐ 3: sums
☐ 4: up
☐ 5: the
☐ 6: problem
☐ 7: succinctly
☐ 8: :
☐ 9: Obama
☐ 10: wanted
☐ 11: nothing
☐ 12: to
☐ 13: do
☐ 14: with
☐ 15: the
☐ 16: Clintons.
# Results

## Overall distribution of Labels

<table>
<thead>
<tr>
<th>Bias type</th>
<th>VL</th>
<th>ML</th>
<th>NB</th>
<th>MC</th>
<th>VC</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>% judged</td>
<td>4.0</td>
<td>8.5</td>
<td>54.8</td>
<td>8.2</td>
<td>6.7</td>
<td>17.9</td>
</tr>
</tbody>
</table>

## Distribution of Labels by Source

<table>
<thead>
<tr>
<th></th>
<th>at</th>
<th>ha</th>
<th>mm</th>
<th>db</th>
<th>tp</th>
<th>tpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>22.3</td>
<td>23.6</td>
<td>7.0</td>
<td>16.2</td>
<td>15.9</td>
<td>14.9</td>
</tr>
<tr>
<td>NB</td>
<td>23.7</td>
<td>22.3</td>
<td>6.1</td>
<td>15.7</td>
<td>17.0</td>
<td>15.3</td>
</tr>
<tr>
<td>VC</td>
<td>24.8</td>
<td>32.3</td>
<td>19.3</td>
<td>6.9</td>
<td>7.5</td>
<td>9.2</td>
</tr>
<tr>
<td>MC</td>
<td>24.4</td>
<td>33.6</td>
<td>8.0</td>
<td>8.2</td>
<td>13.6</td>
<td>12.2</td>
</tr>
<tr>
<td>ML</td>
<td>16.6</td>
<td>15.2</td>
<td>3.4</td>
<td>21.1</td>
<td>22.9</td>
<td>20.9</td>
</tr>
<tr>
<td>VL</td>
<td>16.7</td>
<td>9.0</td>
<td>4.3</td>
<td>31.0</td>
<td>22.4</td>
<td>16.7</td>
</tr>
<tr>
<td>B</td>
<td>20.1</td>
<td>25.4</td>
<td>7.2</td>
<td>19.5</td>
<td>12.3</td>
<td>13.7</td>
</tr>
</tbody>
</table>
Outline

1. Feature Selection and Evaluation Methods

2. Applications
   - Toy Examples
   - Analyzing Bias in Blogs

3. Discussion
Discussion

Things that Surprised Me

- How new the application of NLP to political texts is – the 2008 article puts “text-as-data” in quotes!

Questions

- Is the Bias paper measuring the content of the blogs, what their sentence-selection features choose, or some combination?
- What sort of consistency did the Bias study get on the word evidence question? How did that correlate with their sentence-selection features?
- Aren’t you really just sticking your threshold into your prior parameters with the Laplace prior?