Sentiment analysis and context dependence

Christopher Potts

CS 244U: Natural language understanding
Feb 23
Overview

1. Sentiment as blended and continuous (Experience Project data)
2. Topic-relative sentiment (review data)
3. Sentiment as social: congressional voting data (Thomas et al. 2006)
4. Sentiment as social: Twitter users (Tan et al. 2011)
5. Sentiment as social: Experience Project users and groups
6. Sentiment and morphosyntax
Sentiment as blended and continuous

This one is for the long-suffering fans, the bittersweet memories, the hilariously embarrassing moments, . . .
Sentiment as a classification problem

- Pioneered by Pang et al. (2002), who apply Naive Bayes, MaxEnt, and SVMs to the task of classifying movie reviews as positive or negative,
- and by Turney (2002), who developed vector-based unsupervised techniques (see also Turney and Littman 2003).
- Extended to different sentiment dimensions and different categories sets (Cabral and Hortaçsu 2006; Pang and Lee 2005; Goldberg and Zhu 2006; Snyder and Barzilay 2007; Bruce and Wiebe 1999; Wiebe et al. 1999; Hatzivassiloglou and Wiebe 2000; Riloff and Wiebe 2003; Riloff et al. 2005; Pang and Lee 2004; Thomas et al. 2006; Liu et al. 2003; Alm et al. 2005; Wiebe et al. 2005; Neviarouskaya et al. 2010).
- **Fundamental assumption**: each textual unit (at whatever level of analysis) either has or does not have each sentiment label — usually it has exactly one label.
- **Fundamental assumption**: while the set of all labels might be ranked, they are not continuous.
MaxEnt for sentiment classification

**Definition (MaxEnt)**

\[
P(\text{class}|\text{text}, \lambda) = \frac{\exp \left( \sum_i \lambda_i f_i(\text{class}, \text{text}) \right)}{\sum_{\text{class}'} \exp \left( \sum_i \lambda_i f_i(\text{class}', \text{text}) \right)}
\]

Minimize:

\[
- \sum_{\text{class}, \text{text}} \log P(\text{class}|\text{text}, \lambda) + \log P(\lambda)
\]

Gradient:

\[
\text{empirical count}(f_i, c) - \text{predicted count}(f_i, \lambda)
\]

- A powerful modeling idea for sentiment — can handle features of different type and feature sets with internal statistical dependencies.
- Output is a probability distribution, but classification is typically just based on the most probable class, with little attention to the full distribution.
- Uncertainty about the underlying labels in empirical count\((f_i, c)\) is typically also suppressed/ignored.
Objections to sentiment as classification

- The expression of emotion in language is nuanced, blended, and continuous (Russell (1980); Ekman (1992); Wilson et al. (2006)).
- Human reactions are equally complex and multi-dimensional.
- Insisting on a single label doesn’t do justice to the author’s intentions, and it leads to unreliable labels.
- Few attempts to address this at present (Potts and Schwarz 2010; Potts 2011; Maas et al. 2011; Socher et al. 2011), though that will definitely change soon:
  - New datasets emerging
  - Demands from industry
  - New statistical models
Experience Project confessions: blended, continuous sentiment reactions

<table>
<thead>
<tr>
<th>Category: FRIENDS CONFESSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posted by BrokenAngelWishes on January 20th, 2010 at 12:38 PM</td>
</tr>
<tr>
<td>I really hate being shy... I just want to be able to talk to someone about anything and everything and be myself. That's all I've ever wanted.</td>
</tr>
</tbody>
</table>

14 Reactions
- you rock (1)
- teehee (2)
- I understand (10)
- sorry, hugs (1)
- wow, just wow (0)

6 Comments (add your own)
| Posted by bigbadbear on January 20th, 2010 at 12:41 PM |
| I was really shy when I was younger. I got better when I entered the work field and gained confidence. I think you will grow out of it. :) |
| like 1 dislike Flag |
Experience Project confessions: blended, continuous sentiment reactions

Confession: I really hate being shy . . . I just want to be able to talk to someone about anything and everything and be myself . . . That's all I've ever wanted.
Reactions: *hugs*: 1; *rock*: 1; *teehee*: 2; *understand*: 10; *just wow*: 0

Confession: subconsciously, I constantly narrate my own life in my head. in third person. in a british accent. Insane? Probably
Reactions: *hugs*: 0; *rock*: 7; *teehee*: 8; *understand*: 0; *just wow*: 1

Confession: I have a crush on my boss! *blush* eeek *back to work*
Reactions: *hugs*: 1; *rock*: 0; *teehee*: 4; *understand*: 1; *just wow*: 0

Confession: I bought a case of beer, now I'm watching a South Park marathon while getting drunk :P
Reactions: *hugs*: 2; *rock*: 3; *teehee*: 2, *understand*: 3, *just wow*: 0

Table: Sample Experience Project confessions with associated reaction data.
Experience Project confessions: blended, continuous sentiment reactions

<table>
<thead>
<tr>
<th></th>
<th>Texts</th>
<th>Words</th>
<th>Vocab</th>
<th>Mean words/text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confessions</td>
<td>194,372</td>
<td>21,518,718</td>
<td>143,712</td>
<td>110.71</td>
</tr>
<tr>
<td>Comments</td>
<td>405,483</td>
<td>15,109,194</td>
<td>280,768</td>
<td>37.26</td>
</tr>
</tbody>
</table>

**Table:** The overall size of the corpus.
Reaction distributions

<table>
<thead>
<tr>
<th>Category</th>
<th>Reactions</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>exclamative, positive</td>
<td>sorry, hugs</td>
<td>91,222</td>
<td>(22%)</td>
</tr>
<tr>
<td></td>
<td>you rock</td>
<td>80,798</td>
<td>(19%)</td>
</tr>
<tr>
<td>amused</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>solidarity</td>
<td>teehee</td>
<td>59,597</td>
<td>(14%)</td>
</tr>
<tr>
<td>sympathy</td>
<td>I understand</td>
<td>125,026</td>
<td>(30%)</td>
</tr>
<tr>
<td>exclamative, negative (shocked)</td>
<td>wow, just wow</td>
<td>60,952</td>
<td>(15%)</td>
</tr>
</tbody>
</table>

Total: 417,595

(a) All reactions.

<table>
<thead>
<tr>
<th>Texts</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 1</td>
<td>140,467</td>
</tr>
<tr>
<td>≥ 2</td>
<td>92,880</td>
</tr>
<tr>
<td>≥ 3</td>
<td>60,880</td>
</tr>
<tr>
<td>≥ 4</td>
<td>39,342</td>
</tr>
<tr>
<td>≥ 5</td>
<td>25,434</td>
</tr>
</tbody>
</table>

(b) Per text.

Table: In general, reader reactions are sympathetic and supportive.
Reaction distributions

Figure: The entropy of the reaction distributions.
## Counting and visualizing: Experience Project

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat.</td>
<td>Count</td>
<td>Total</td>
<td>Pr(_{EP}(w</td>
<td>c))</td>
</tr>
<tr>
<td>hugs</td>
<td>108</td>
<td>2,153,134</td>
<td>0.00005</td>
<td>0.25</td>
</tr>
<tr>
<td>rock</td>
<td>34</td>
<td>1,330,084</td>
<td>0.00002</td>
<td>0.13</td>
</tr>
<tr>
<td>teehee</td>
<td>25</td>
<td>845,397</td>
<td>0.00003</td>
<td>0.15</td>
</tr>
<tr>
<td>understand</td>
<td>197</td>
<td>3,447,377</td>
<td>0.00006</td>
<td>0.29</td>
</tr>
<tr>
<td>just wow</td>
<td>29</td>
<td>838,059</td>
<td>0.00004</td>
<td>0.18</td>
</tr>
</tbody>
</table>

**Pr\(_{EP}(w|c)\) \(\overset{def}{=} \) Count\((w, r)\)/Total\((r)\)**

**Pr\(_{EP}(c|w)\) \(\overset{def}{=} \) \(\frac{Pr\(_{EP}(w|c)\)}{\sum_{x \in \text{Categories}} Pr\(_{EP}(w|x)\)}\)**

![disappoint(ed/ing) (145 tokens) chart]
Word-level sentiment examples

(a) Words eliciting predominantly ‘You rock’ reactions. The data reveal other dimensions as well, including mixes of light-heartedness, negative exclamativity.

(b) Words eliciting sympathetic (‘sorry, hugs’, ‘I understand’) reactions. Other categories rise to prominence as well, depending on the lexical semantics and pragmatics of the word.

Figure: Word–category associations in the EP data.
A model for sentiment distributions

Definition (MaxEnt with distributional labels)

\[ P(\text{class}|\text{text}, \lambda) = \frac{\exp \left( \sum_i \lambda_i f_i(\text{class}, \text{text}) \right)}{\sum_{\text{class}'} \exp \left( \sum_i \lambda_i f_i(\text{class}', \text{text}) \right)} \]

Minimize the KL divergence of the predicted distribution from the empirical one:

\[ \sum_{\text{class}, \text{text}} \text{empiricalProb}(\text{class}|\text{text}) \log \left( \frac{\text{empiricalProb}(\text{class}|\text{text})}{P(\text{class}|\text{text}, \lambda)} \right) \]

Gradient:

\[ \sum_{\text{text}} \text{empiricalProb}(\text{class}|\text{text}) - P(\text{class}|\text{text}, \lambda) \]
## Some results

<table>
<thead>
<tr>
<th>Features</th>
<th>$\geq$ 5 reactions</th>
<th>$\geq$ 1 reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KL</td>
<td>Max Acc.</td>
</tr>
<tr>
<td>Uniform Reactions</td>
<td>0.861</td>
<td>20.2</td>
</tr>
<tr>
<td>Mean Training Reactions</td>
<td>0.763</td>
<td>43.0</td>
</tr>
<tr>
<td>Bag of Words (All unigrams)</td>
<td>0.637</td>
<td>56.0</td>
</tr>
<tr>
<td>Bag of Words (Top 5000 unigrams)</td>
<td>0.640</td>
<td>54.9</td>
</tr>
<tr>
<td>LSA</td>
<td>0.667</td>
<td>51.8</td>
</tr>
<tr>
<td>Our Method Laplacian Prior</td>
<td>0.621</td>
<td>55.7</td>
</tr>
<tr>
<td>Our Method Gaussian Prior</td>
<td>0.620</td>
<td>55.2</td>
</tr>
</tbody>
</table>

**Table:** Results from Maas et al. 2011. The first two are simple baselines. The ‘Bag of words’ models are MaxEnt/softmax. LSA and ‘Our method’ uses word vectors for predictions, by training on the average score in the vector. ‘Our method’ is distinguished primarily by combining an unsupervised VSM with a supervised component using star-ratings.
Topic-relative sentiment

- Sentiment is often topic relative
  (“We loved the food but hated the waiter.”)

- Sentiment vocabulary is topic dependent
  (tasty, beautiful, melodious, plush, . . .)

- Sentiment feature values can vary dramatically by topic
  (“The movie {Scream/Love Story} was totally gross!”)
Attribute-relative sentiment (Liu et al. 2005)

Figure 1: Visual comparison of consumer opinions on two products.

Associated datasets:
http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
OpenTable: attribute-level ratings

Figure: OpenTable rating distributions. Positive reviews dominate in all categories. ‘Noise’ is fundamentally different, since it doesn’t have a standard preference ordering.

(a) Comparisons with ‘Overall’. In each panel, the overall rating value is subtracted from the other rating value. Thus, a value of 0 indicates agreement between the two ratings for the review in question.

(b) Correlations.

---

| Overall, Food | 0.82 |
| Overall, Service | 0.77 |
| Overall, Ambiance | 0.70 |
| Food, Service | 0.57 |
| Food, Ambiance | 0.56 |
| Ambiance, Service | 0.54 |

Figure: OpenTable rating category comparisons. ‘Overall’ and ‘Food’ are highly correlated.
Counting and visualizing: IMDB

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Count</td>
<td>Total</td>
<td>Pr(w</td>
<td>r)</td>
</tr>
<tr>
<td>-4.5</td>
<td>2983</td>
<td>28,962,201</td>
<td>0.00010</td>
<td>0.17</td>
</tr>
<tr>
<td>-3.5</td>
<td>1056</td>
<td>13,436,851</td>
<td>0.00008</td>
<td>0.13</td>
</tr>
<tr>
<td>-2.5</td>
<td>1041</td>
<td>15,987,151</td>
<td>0.00007</td>
<td>0.11</td>
</tr>
<tr>
<td>-1.5</td>
<td>819</td>
<td>17,095,212</td>
<td>0.00005</td>
<td>0.08</td>
</tr>
<tr>
<td>-0.5</td>
<td>848</td>
<td>23,293,790</td>
<td>0.00004</td>
<td>0.06</td>
</tr>
<tr>
<td>+0.5</td>
<td>975</td>
<td>31,317,918</td>
<td>0.00003</td>
<td>0.05</td>
</tr>
<tr>
<td>+1.5</td>
<td>1407</td>
<td>45,913,948</td>
<td>0.00003</td>
<td>0.05</td>
</tr>
<tr>
<td>+2.5</td>
<td>2326</td>
<td>55,634,817</td>
<td>0.00004</td>
<td>0.07</td>
</tr>
<tr>
<td>+3.5</td>
<td>2940</td>
<td>45,941,763</td>
<td>0.00006</td>
<td>0.11</td>
</tr>
<tr>
<td>+4.5</td>
<td>7915</td>
<td>84,294,625</td>
<td>0.00009</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Pr(w|r) \(\overset{\text{def}}{=} \frac{\text{Count}(w, r)}{\text{Total}(r)}\)

Pr(r|w) \(\overset{\text{def}}{=} \frac{\text{Pr}(w|r)}{\sum_{x \in \text{Rating}} \text{Pr}(w|x)}\)
Counting and visualizing: IMDB

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Count</td>
<td>Total</td>
<td>Pr(w</td>
<td>r)</td>
</tr>
<tr>
<td>-4.5</td>
<td>2983</td>
<td>28,962,201</td>
<td>0.00010</td>
<td>0.17</td>
</tr>
<tr>
<td>-3.5</td>
<td>1056</td>
<td>13,436,851</td>
<td>0.00008</td>
<td>0.13</td>
</tr>
<tr>
<td>-2.5</td>
<td>1041</td>
<td>15,987,151</td>
<td>0.00007</td>
<td>0.11</td>
</tr>
<tr>
<td>-1.5</td>
<td>819</td>
<td>17,095,212</td>
<td>0.00005</td>
<td>0.08</td>
</tr>
<tr>
<td>-0.5</td>
<td>848</td>
<td>23,293,790</td>
<td>0.00005</td>
<td>0.06</td>
</tr>
<tr>
<td>+0.5</td>
<td>975</td>
<td>31,317,918</td>
<td>0.00004</td>
<td>0.05</td>
</tr>
<tr>
<td>+1.5</td>
<td>1407</td>
<td>45,913,948</td>
<td>0.00003</td>
<td>0.05</td>
</tr>
<tr>
<td>+2.5</td>
<td>2326</td>
<td>55,634,817</td>
<td>0.00004</td>
<td>0.07</td>
</tr>
<tr>
<td>+3.5</td>
<td>2940</td>
<td>45,941,763</td>
<td>0.00006</td>
<td>0.11</td>
</tr>
<tr>
<td>+4.5</td>
<td>7915</td>
<td>84,294,625</td>
<td>0.00009</td>
<td>0.16</td>
</tr>
</tbody>
</table>

\[
\text{Pr}(w|r) \overset{\text{def}}{=} \frac{\text{Count}(w, r)}{\text{Total}(r)}
\]

\[
\text{Pr}(r|w) \overset{\text{def}}{=} \frac{\text{Pr}(w|r)}{\sum_{x \in \text{Rating}} \text{Pr}(w|x)}
\]

Pr(wow) = \logit^{-1}(\text{intercept } + \text{rating})

Rating coef. = 0.01 (p = 0.875)
### Counting and visualizing: IMDB

| Rating | Count | Total | Pr(w|r) | Pr(r|w) |
|--------|-------|-------|---------|---------|
| −4.5   | 2983  | 28,962,201 | 0.00010 | 0.17    |
| −3.5   | 1056  | 13,436,851 | 0.00008 | 0.13    |
| −2.5   | 1041  | 15,987,151 | 0.00007 | 0.11    |
| −1.5   | 819   | 17,095,212 | 0.00005 | 0.08    |
| −0.5   | 848   | 23,293,790 | 0.00004 | 0.06    |
| +0.5   | 975   | 31,317,918 | 0.00003 | 0.05    |
| +1.5   | 1407  | 45,913,948 | 0.00003 | 0.05    |
| +2.5   | 2326  | 55,634,817 | 0.00004 | 0.07    |
| +3.5   | 2940  | 45,941,763 | 0.00006 | 0.11    |
| +4.5   | 7915  | 84,294,625 | 0.00009 | 0.16    |

\[
Pr(w|r) \overset{\text{def}}{=} \frac{\text{Count}(w, r)}{\text{Total}(r)}
\]

\[
Pr(r|w) \overset{\text{def}}{=} \frac{Pr(w|r)}{\sum_{x \in \text{Rating}} Pr(w|x)}
\]

\[
\text{logit}^{-1}\left(\text{intercept } + \frac{\text{rating}}{\text{rating}^2}\right)
\]

Rating coef. = -0.02 (p = 0.105)
Rating^2 coef. = 0.05 (p < 0.001)
Counting and visualizing: IMDB

| Rating | Count | Total | Pr(w|r) | Pr(r|w) |
|--------|-------|-------|---------|---------|
| −4.5   | 2983  | 28,962,201 | 0.000010 | 0.17 |
| −3.5   | 1056  | 13,436,851 | 0.000008 | 0.13 |
| −2.5   | 1041  | 15,987,151 | 0.000007 | 0.11 |
| −1.5   | 819   | 17,095,212 | 0.000005 | 0.08 |
| −0.5   | 848   | 23,293,790 | 0.000004 | 0.06 |
| +0.5   | 975   | 31,317,918 | 0.000003 | 0.05 |
| +1.5   | 1407  | 45,913,948 | 0.000003 | 0.05 |
| +2.5   | 2326  | 55,634,817 | 0.000004 | 0.07 |
| +3.5   | 2940  | 45,941,763 | 0.000006 | 0.11 |
| +4.5   | 7915  | 84,294,625 | 0.000009 | 0.16 |

\[ \Pr(w|r) \overset{\text{def}}{=} \frac{\text{Count}(w, r)}{\text{Total}(r)} \]

\[ \Pr(r|w) \overset{\text{def}}{=} \frac{\Pr(w|r)}{\sum_{x \in \text{Rating}} \Pr(w|x)} \]
**IMDb movie reviews: word-level distributional profiles**

Positive and negative scalar terms

- **good** – 732963 tokens
  - Rating coef. = 0.01 (p = 0.152)
  - Rating^2 coef. = -0.02 (p < 0.001)

- **excellent** – 136404 tokens
  - Rating coef. = 0.22 (p < 0.001)

- **awesome** – 43134 tokens
  - Rating coef. = 0.19 (p < 0.001)
  - Rating^2 coef. = 0.03 (p < 0.001)

- **amazing** – 93569 tokens
  - Rating coef. = 0.19 (p < 0.001)
  - Rating^2 coef. = 0.04 (p < 0.001)

- **disappointing** – 20342 tokens
  - Rating coef. = -0.21 (p < 0.001)
  - Rating^2 coef. = -0.06 (p < 0.001)

- **annoying** – 33536 tokens
  - Rating coef. = -0.21 (p < 0.001)
  - Rating^2 coef. = -0.02 (p < 0.001)

- **bad** – 254146 tokens
  - Rating coef. = -0.20 (p < 0.001)
  - Rating^2 coef. = 0.01 (p < 0.001)

- **terrible** – 45470 tokens
  - Rating coef. = -0.28 (p < 0.001)
  - Rating^2 coef. = 0.02 (p < 0.001)
IMDB movie reviews: word-level distributional profiles

Emphasizing and attenuating terms

- **ever** – 247634 tokens
  - Rating coef. = -0.01 (p = 0.124)
  - Rating^2 coef. = 0.06 (p < 0.001)

- **totally** – 57637 tokens
  - Rating coef. = -0.06 (p < 0.001)
  - Rating^2 coef. = 0.02 (p = 0.005)

- **utterly** – 16442 tokens
  - Rating coef. = -0.11 (p < 0.001)
  - Rating^2 coef. = 0.02 (p = 0.005)

- **damn** – 14252 tokens
  - Rating coef. = -0.03 (p = 0.001)
  - Rating^2 coef. = 0.02 (p < 0.001)

- **but** – 1818566 tokens
  - Rating coef. = -0.02 (p = 0.004)
  - Rating^2 coef. = -0.02 (p < 0.001)

- **somewhat** – 49434 tokens
  - Rating coef. = 0.04 (p = 0.003)
  - Rating^2 coef. = -0.05 (p < 0.001)

- **okay** – 36099 tokens
  - Rating coef. = -0.13 (p < 0.001)
  - Rating^2 coef. = -0.03 (p = 0.002)

- **bland** – 9494 tokens
  - Rating coef. = -0.25 (p < 0.001)
  - Rating^2 coef. = -0.06 (p < 0.001)
IMDB movie reviews: variation by genre

**depressing**

- **Action** – 1264 tokens
  - Rating distribution
- **Comedy** – 1600 tokens
  - Rating distribution
- **Drama** – 3759 tokens
  - Rating distribution
- **Documentary** – 213 tokens
  - Rating distribution

**scary**

- **Horror** – 14498 tokens
  - Rating distribution
- **Drama** – 7943 tokens
  - Rating distribution
- **Action** – 5814 tokens
  - Rating distribution
- **Comedy** – 6416 tokens
  - Rating distribution

**Sandler**

- **Action** – 628 tokens
  - Rating distribution
- **Animation** – 170 tokens
  - Rating distribution
- **Comedy** – 5840 tokens
  - Rating distribution
- **Drama** – 414 tokens
  - Rating distribution
Other examples of topic/aspect relative sentiment

- User-level variation in both author and reader reactions
  - age
  - formality
  - in-group/public

- Dialect features and regional differences in emphasis
Sentiment as social: Convote (Thomas et al. 2006)

- Using text and social ties to predict congressional voting.
- Adapts the hierarchical model of Pang and Lee (2004), where subjectivity scores are used to focus a subsequent polarity classifier.
- A pioneering attempt to treat sentiment (here, support/opposition) as a social phenomenon.
The Convote corpus

<table>
<thead>
<tr>
<th>Bill</th>
<th>052</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>400011</td>
</tr>
<tr>
<td>Party</td>
<td>Democrat</td>
</tr>
<tr>
<td>Vote</td>
<td>No</td>
</tr>
<tr>
<td>Sample</td>
<td>the question is, what happens during those 45 days? we will need to support elections. there is not a single member of this house who has not supported some form of general election, a special election, to replace the members at some point. but during that 45 days, what happens?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bill</th>
<th>052</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>400077</td>
</tr>
<tr>
<td>Party</td>
<td>Republican</td>
</tr>
<tr>
<td>Vote</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample</td>
<td>i believe this is a fair rule that allows for a full discussion of the relevant points pertaining to the legislation before us. mr. speaker, h.r. 841 is an important step forward in addressing what are critical shortcomings in america 's plan for the continuity of this house in the event of an unexpected disaster or attack.</td>
</tr>
</tbody>
</table>
The Convote corpus

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>train</th>
<th>test</th>
<th>development</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech segments</td>
<td>3857</td>
<td>2740</td>
<td>860</td>
<td>257</td>
</tr>
<tr>
<td>debates</td>
<td>53</td>
<td>38</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>average number of speech</td>
<td>72.8</td>
<td>72.1</td>
<td>86.0</td>
<td>51.4</td>
</tr>
<tr>
<td>segments per debate</td>
<td>32.1</td>
<td>30.9</td>
<td>41.1</td>
<td>22.6</td>
</tr>
</tbody>
</table>

Table 1: Corpus statistics.

Hierarchy of texts:

Debates (collections of speeches by different speakers)

↑

Speeches (collections of segments by the same speaker)

↑

Speech segments (documents in the corpus)
Basic classification with same-speech links

1. SVM classifier with unigram-presence features predicting, for each speech-segment, how the speaker voted (Y or N).

2. For each document $s$ belonging to speech $S$, the SVM score for $s$ is divided by the standard deviation for all $s' \in S$.

3. Debate-graph construction with minimal cuts:

$$\text{score}(s) \leq -2 \Rightarrow \begin{cases} \text{source} & \rightarrow & s \\ s & \rightarrow & \text{sink} \end{cases}$$

$$\text{score}(s) \geq +2 \Rightarrow \begin{cases} \text{source} & \rightarrow & s \\ s & \rightarrow & \text{sink} \end{cases}$$

else $\Rightarrow \begin{cases} \text{source} & \rightarrow & s \\ s & \rightarrow & \text{sink} \end{cases}$
Graph construction and minimal cuts

(source = No; sink = Yes)
Graph construction and minimal cuts

(source = No; sink = Yes)
Graph construction and minimal cuts

Cost: 7500+2500

(source = No; sink = Yes)
Graph construction and minimal cuts

Cost:
7500 + 10000 + 7500

(source = No; sink = Yes)
Graph construction and minimal cuts

Cost: 10000+2500+2500

(source = No; sink = Yes)
Graph construction and minimal cuts

Cost: 2500+2500

(source = No; sink = Yes)
### Speaker references

<table>
<thead>
<tr>
<th>Bill</th>
<th>006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>400115</td>
</tr>
<tr>
<td>Party</td>
<td>Republican</td>
</tr>
<tr>
<td>Vote</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample</td>
<td>mr. speaker, i am very happy to yield 3 minutes to the gentleman from new york (mr. boehlert) xz4000350, the very distinguished chairman of the committee on science.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bill</th>
<th>006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>400035</td>
</tr>
<tr>
<td>Party</td>
<td>Republican</td>
</tr>
<tr>
<td>Vote</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample</td>
<td>mr. speaker, i rise in strong support of this balanced rules package. i want to speak particularly to the provisions regarding homeland security. [. . .]</td>
</tr>
</tbody>
</table>
Speaker reference classifier

1. Label a reference as Agree if the speaker and the Referent voted the same way, else Disagree.

2. Features: 30 unigrams before, the name, and 30 unigrams after

3. Normalized SVM scores from this classifier are then added to the debate graphs, at the level of speech segments. (Where a speaker has multiple speech segments, one is chosen at random; the infinite-weight links ensure that this information propagates to the others.)
Inter-text and inter-speaker links

\[
\begin{align*}
A &\quad s1 & -2 \\
B &\quad s2 & +1 \\
C &\quad s3 & +2 \\
C &\quad s4 & -1
\end{align*}
\]

\[
\begin{align*}
&\quad 20 \\
&\quad 2000 \\
&\quad \infty
\end{align*}
\]

(source to sink) 

\[
\begin{align*}
10,000 &\quad (1+2)^{2500}=7500 \\
0 &\quad (-1+2)^{2500}=2500 \\
10,000-2500 &\quad 10,000-7500
\end{align*}
\]

(green = spk-ref links)
Results

<table>
<thead>
<tr>
<th>Support/oppose classifier</th>
<th>Devel. set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>(“speech segment⇒yea?”)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>majority baseline</td>
<td>54.09</td>
<td>58.37</td>
</tr>
<tr>
<td>#(“support”) − #(“oppos”)</td>
<td>59.14</td>
<td>62.67</td>
</tr>
<tr>
<td>SVM [speech segment]</td>
<td>70.04</td>
<td>66.05</td>
</tr>
<tr>
<td>SVM + same-speaker links</td>
<td>79.77</td>
<td>67.21</td>
</tr>
<tr>
<td>SVM + same-speaker links . . .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ agreement links, $\theta_{agr} = 0$</td>
<td>89.11</td>
<td>70.81</td>
</tr>
<tr>
<td>+ agreement links, $\theta_{agr} = \mu$</td>
<td>87.94</td>
<td>71.16</td>
</tr>
</tbody>
</table>

Table 4: Segment-based speech-segment classification accuracy, in percent.

$\theta_{agr}$ is a free-parameter in the scaling function for speaker agreement scores. The development results suggest that 0 is the better value than $\mu$ (a mean of all the debate’s scores), but $\mu$ performs better in testing.
Sentiment as social: Twitter users (Tan et al. 2011)

**Goal**
Given a topic $q$ and a user $v$, predict whether $v$ is positive or negative wrt topic $q$.

**Guiding idea (builds on Thomas et al. 2006)**
Users in the same social network will tend to share sentiment, so bringing in these social ties will improve sentiment predictions.

**Data**
Topically-clustered tweets, with social network determined by the following relation or the connection user $a$ makes with user $b$ by tweeting “@b . . .”
Dataset (Tan et al. 2011)

### Table 1: Statistics for our main datasets.

<table>
<thead>
<tr>
<th>Topic</th>
<th># users</th>
<th># t-follow edges</th>
<th># @ edges</th>
<th># on-topic tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>dir.</td>
<td>mutual</td>
<td>dir.</td>
</tr>
<tr>
<td>Obama</td>
<td>889</td>
<td>7,838</td>
<td>2,949</td>
<td>2,358</td>
</tr>
<tr>
<td>Sarah Palin</td>
<td>310</td>
<td>1,003</td>
<td>264</td>
<td>449</td>
</tr>
<tr>
<td>Glenn Beck</td>
<td>313</td>
<td>486</td>
<td>159</td>
<td>148</td>
</tr>
<tr>
<td>Lakers</td>
<td>640</td>
<td>2,297</td>
<td>353</td>
<td>1,167</td>
</tr>
<tr>
<td>Fox News</td>
<td>231</td>
<td>130</td>
<td>32</td>
<td>37</td>
</tr>
</tbody>
</table>

- Set of topics chosen by hand, explicitly favoring polarizing topics so that the classes could be balanced.
- For the following relations, ‘dir’ means that the following or @-link goes in at least one direction, whereas mutual means that it goes in both directions.
- User-level polarity was determined by inspecting biographies and in some cases their tweets and using that information to assign a label by hand.
- The dataset is only partially labeled.
In keeping with the guiding intuition,

- connected users tend to share the same sentiment (left); and
- users who share sentiment are more likely to be connected.

Figure 1: Shared sentiment conditioned on type of connection. Y-axis: probability of two users $v_i$ and $v_j$ having the same sentiment label, conditioned on relationship type. The left plot is for the t-follow graph, while the right one is for the @ graph. “random”: pairs formed by randomly choosing users. “directed”: at least one user in the pair links to the other. “mutual”: both users in the pair link to each other. Note that the very last bar (a value of 1 for “Fox News”, mutual @-graph) is based on only 5 edges (datapoints).

Figure 2: Connectedness conditioned on labels. Y-axis: probability that two users are connected, conditioned on whether or not the users have the same sentiment.
Graphical structure and model (Tan et al. 2011)

Heterogeneous graph

\[ k = \text{user sentiment label} \in \{0,1\} \]
\[ l = \text{tweets sentiment label} \in \{0,1\} \]
\[ v_i \text{ sentiment label} \in \{0,1\} \]
\[ \text{unknown label} \in \{0,1\} \text{ for } v_i \text{ tweets} \]

\[ \log P(Y) = \left( \sum_{v_i \in V} \left[ \sum_{t \in \text{tweets}_{v_i}, k, \ell} \mu_{k,\ell} f_{k,\ell}(y_{v_i}, \hat{y}_t) \right. \right. \]
\[ \left. \left. + \sum_{v_j \in \text{Neighbors}_{v_i}, k, \ell} \lambda_{k,\ell} h_{k,\ell}(y_{v_i}, y_{v_j}) \right] \right) - \log Z, \]

vector of user labels for the topic in question

User–tweet factor

\[ f_{k,\ell}(y_{v_i}, \hat{y}_t) = \begin{cases} \frac{w_{\text{labeled}}}{|\text{tweets}_{v_i}|}, & y_{v_i} = k, \hat{y}_t = \ell, v_i \text{ labeled} \\ \frac{w_{\text{unlabeled}}}{|\text{tweets}_{v_i}|}, & y_{v_i} = k, \hat{y}_t = \ell, v_i \text{ unlabeled} \\ 0, & \text{otherwise} \end{cases} \]

= 1 in experiments
= 0.125 in experiments

weights for impact of the feature functions; set by counting labels or SampleRank

User–user factor

\[ h_{k,\ell}(y_{v_i}, y_{v_j}) = \begin{cases} \frac{w_{\text{relation}}}{|\text{Neighbors}_{v_i}|}, & y_{v_i} = k, y_{v_j} = \ell \\ 0, & \text{otherwise} \end{cases} \]

= 0.6 in experiments

Overview
Blended sentiment
Topic-relative
Social
Users
Context
Morphosyntax
Conclusion
Case study highlighting the value of social information

Figure 4: Case study: Portion of the t-follow graph for the topic “Obama”, where derived labels on users are indicated by green (positive) and red (negative), respectively. Each node is a user, and the center one is “BarackObama”. The numbers in the nodes are indices into the table below. (a): Ground truth (human annotation). (b) SVM Vote (baseline). (c) HGM-Learning in the directed t-follow graph. Filled nodes indicate cases where the indicated algorithm was right and the other algorithm was wrong; for instance, only our algorithm was correct on node 4.
By-topic results

Figure 6: Performance Analysis in Different Topics. The x-axes are the same as in Figure 5. Bars summarize performance results for our “10-run” experiments: the bottom and top of a bar indicate the 25th and 75th percentiles, respectively. Dots indicate median results; in pairs connected by lines, the left is “NoLearning”, while the right is “Learning”. Green: SVM vote, our baseline. Red: network-based approaches applied to the t-follow graphs. Blue: results for the @ graphs. Stars (⋆) indicate performance that is significantly better than the baseline, according to the paired t-test.
Possible extensions of the Convote/Twitter-graph approach

- Newspapers as users, soft constraints on shared sentiment, network structure given by opinion and corp structure
- Scientific network + shared views on controversial topics
- Convote 2012—stronger social feats due to increased polarization?
- Sentiment in hyperlink structures
Sentiment as social: Experience Project

Confession: I really hate being shy . . . I just want to be able to talk to someone about anything and everything and be myself... That’s all I’ve ever wanted.
Reactions: hugs: 1; rock: 1; teehee: 2; understand: 10; just wow: 0;

Confession: I bought a case of beer, now I’m watching a South Park marathon while getting drunk :P
Reactions: hugs: 2; rock: 3; teehee: 2, understand: 3, just wow: 0;

Table: Sample Experience Project confessions with associated reaction data, author demographics, and text groups.
Sentiment as social: Experience Project

Confession: I really hate being shy . . . I just want to be able to talk to someone about anything and everything and be myself... That’s all I’ve ever wanted.

Reactions: hugs: 1; rock: 1; teehee: 2; understand: 10; just wow: 0;

Author age 21
Author gender female
Text group friends

Confession: I bought a case of beer, now I’m watching a South Park marathon while getting drunk :P

Reactions: hugs: 2; rock: 3; teehee: 2, understand: 3, just wow: 0

Author age 25
Author gender male
Text group health

Table: Sample Experience Project confessions with associated reaction data, author demographics, and text groups.
Table: Contextual metadata. The EP’s demographics seem to be skewed towards young women writing about issues concerning their interpersonal relationships.
The influences of context

Figure: Text groups show the most variability. Age and gender are more stable by comparison, though the relationships remain interesting.
The influences of context

(a) Text groups.
The influences of context

Figure: Text groups show the most variability. Age and gender are more stable by comparison, though the relationships remain interesting.
The influences of context

(c) Gender.
The influences of text groups

**Figure:** Words eliciting predominantly ‘You rock’ reactions. The data reveal other dimensions as well, including mixes of light-heartedness, negative exclamativitiy.

**Figure:** The bimodal distribution of *survive* seems to derive from an underlying distinction in text group.
The influences of age

Figure: Age is a source of variation in responses to *drunk*.
Modeling ideas

- Demographic and text-group features can be treated on par with linguistic features.
- They could also be brought in as hierarchical effects in a multi-level generalized linear model (Gelman and Hill 2007; Baayen 2008).
- In ongoing work with Andrew Maas, Peter Pham, and Andrew Ng, we have been using Conditional Random Fields (Lafferty et al. 2001; Sutton and McCallum 2010) to define context-relative feature functions to directly model the distribution $P(\text{class}|\text{text}, \text{context}, \lambda)$. 
Sentiment and morphosyntax

I’ve so far concentrated on general features of the context of use. Sentiment is also profoundly influenced by the immediate linguistic context.

1. That was fun :)
2. That was miserable :(  
3. I stubbed my damn toe  
4. What’s with these friggin QR codes?  
5. It was wonderful.  
6. He knows it is wonderful.  
7. It was not wonderful.  
8. No one found it to be wonderful.  
9. They said it would be wonderful, but they were wrong: it was awful!  
10. This “wonderful” movie turned out to be boring.
Degree modification

The intensifiers *really* and *very* enhance sentiment:

**good** – 1154386 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.08</th>
<th>0.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>2.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Rating coef. = -0.01 (p = 0.092)
Rating*2 coef. = -0.02 (p < 0.001)

**bad** – 499177 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.11</th>
<th>0.24</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Rating coef. = -0.22 (p < 0.001)
Rating*2 coef. = 0.01 (p = 0.089)

**wonderful** – 102679 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.04</th>
<th>0.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Rating coef. = 0.25 (p < 0.001)
Rating*2 coef. = 0.01 (p = 0.051)

**very good** – 80212 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.16</th>
<th>0.27</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Rating coef. = 0.13 (p < 0.001)
Rating*2 coef. = -0.03 (p = 0.001)

**very bad** – 7197 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.03</th>
<th>0.24</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Rating coef. = -0.28 (p < 0.001)
Rating*2 coef. = 0.01 (p = 0.051)

**really wonderful** – 423 tokens

<table>
<thead>
<tr>
<th>Rating</th>
<th>0.03</th>
<th>0.24</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Rating coef. = 0.26 (p < 0.001)
Exclamatives

Exclamatives (e.g., *what a view!*) both create and enhance sentiment:

- **wonderful – 94238 tokens**
  - Rating coef. = 0.21 (p < 0.001)
  - Rating^2 coef. = 0.02 (p = 0.007)

- **terrible – 45470 tokens**
  - Rating coef. = -0.28 (p < 0.001)
  - Rating^2 coef. = 0.02 (p < 0.001)

- **movie – 2261241 tokens**
  - Rating coef. = -0.02 (p < 0.001)
  - Rating^2 coef. = 0.03 (p = 0.052)

- **what a wonderful – 1005 tokens**
  - Rating coef. = 0.26 (p < 0.001)
  - Rating^2 coef. = 0.04 (p = 0.007)

- **what a terrible – 414 tokens**
  - Rating coef. = -0.32 (p < 0.001)
  - Rating^2 coef. = 0.03 (p = 0.052)

- **what a movie – 962 tokens**
  - Rating coef. = 0.15 (p < 0.001)
  - Rating^2 coef. = 0.05 (p < 0.001)
Negation

Negating mid-scalar terms leads to polarity reversal. Negating high-scalar terms (positive or negative) leads to mere attenuation.

**good** – 732963 tokens

Rating coef. = 0.23 (p < 0.001)
Rating*2 coef. = -0.02 (p < 0.001)

**bad** – 254146 tokens

Rating coef. = -0.33 (p < 0.001)
Rating*2 coef. = 0.01 (p < 0.001)

**delighted** – 2528 tokens

Rating coef. = 0.12 (p = 0.215)
Rating*2 coef. = 0 (p = 0.568)

**miserable** – 4245 tokens

Rating coef. = -0.33 (p < 0.001)
Rating*2 coef. = 0.02 (p < 0.001)

**good**_NEG_ – 169772 tokens

Rating coef. = 0.1 (p = 0.007)
Rating*2 coef. = -0.01 (p < 0.001)

**bad**_NEG_ – 113865 tokens

Rating coef. = 0.12 (p = 0.164)
Rating*2 coef. = -0.02 (p = 0.011)

**delighted**_NEG_ – 277 tokens

Rating coef. = 0.22 (p = 0.206)
Rating*2 coef. = 0 (p = 0.815)

**miserable**_NEG_ – 724 tokens

Rating coef. = -0.4 (p = 0.002)
Rating*2 coef. = 0.02 (p = 0.015)

not good  $\approx$  bad  
not bad   $\approx$  good

not delighted  $\not\approx$  miserable
not miserable  $\not\approx$  delighted
**Attenuators**

Adverbials like *pretty* weaken/attenuate sentiment:

<table>
<thead>
<tr>
<th>Good</th>
<th>732963 tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretty good</td>
<td>32276 tokens</td>
</tr>
<tr>
<td>Amazing</td>
<td>93569 tokens</td>
</tr>
<tr>
<td>Pretty amazing</td>
<td>567 tokens</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bad</th>
<th>254146 tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretty bad</td>
<td>4191 tokens</td>
</tr>
<tr>
<td>Bad</td>
<td>254146 tokens</td>
</tr>
<tr>
<td>Pretty amazing</td>
<td>567 tokens</td>
</tr>
</tbody>
</table>

**Examples:**

- **good**: Rating coef. = 0.23 (p < 0.001)
  Rating^2 coef. = -0.02 (p < 0.001)

- **bad**: Rating coef. = -0.33 (p < 0.001)
  Rating^2 coef. = 0.01 (p < 0.001)

- **amazing**: Rating coef. = -0.2 (p < 0.001)
  Rating^2 coef. = 0.04 (p < 0.001)

- **pretty good**: Rating coef. = 0.76 (p < 0.001)
  Rating^2 coef. = -0.07 (p < 0.001)

- **pretty bad**: Rating coef. = 0.17 (p = 0.174)
  Rating^2 coef. = -0.04 (p = 0.007)

- **pretty amazing**: Rating coef. = 0.44 (p < 0.001)
  Rating^2 coef. = -0.03 (p < 0.001)
Intensification: the weak overtake the strong

Low-scalar modifiers are likely to be intensified, which can confuse models into thinking that they are stronger than their high-scalar counterparts:

- **bad – 499,177 tokens**
  - Cat = -0.22 (p < 0.001)
  - Cat^2 = 0.01 (p = 0.089)

- **incredibly/r bad/a – 738 tokens**
  - Cat = -0.53 (p = 0.001)
  - Cat^2 = 0.01 (p = 0.31)

- **awful – 50,274 tokens**
  - Cat = -0.33 (p < 0.001)
  - Cat^2 = 0.02 (p < 0.001)
Attitude predictions and thwarted expectations

I had been looking forward to this film since I heard about it early last year, when Matthew Perry had just signed on. I'm a big fan of Perry's subtle sense of humor, and in addition, I think Chris Farley's on-edge, extreme acting was a riot. So naturally, when the trailer for "Almost Heroes" hit theaters, I almost jumped up and down. A soda in hand, the lights dimming, I was ready to be blown away by Farley's final starring role and what was supposed to be Matthew Perry's big breakthrough. I was ready to be just amazed; for this to be among Farley's best, in spite of David Spade's absence. I was ready to be laughing my head off the minute the credits ran. Sadly, none of this came to pass. The humor is spotty at best, with good moments and laughable one-liners few and far between. Perry and Farley have no chemistry; the role that Perry was cast in seems obviously written for Spade, for it's his type of humor, and not at all what Perry is associated with. And the movie tries to be smart, a subject best left alone when it's a Farley flick. The movie is a major disappointment, with only a few scenes worth a first look, let alone a second. Perry delivers not one humorous line the whole movie, and not surprisingly; the only reason the movie made the top ten grossing list opening week was because it was advertised with Farley. And Farley's classic humor is widespread, too. Almost Heroes almost works, but misses the wagon-train by quite a longshot. Guys, let's leave the exploring to Lewis and Clark, huh? Stick to "Tommy Boy", and we'll all be "friends".

Table: An example of thwarted expectations. This is a negative review. Inquirer positive terms are in blue, and Inquirer negative terms are red. There are 20 positive terms and six negative ones, for a Pos:Neg ratio of 3.33.
Attitude predictions and thwarted expectations

Figure: Inquirer Pos:Neg ratios obtained by counting the terms in the review that are classified as Positiv or Negativ in the Harvard Inquirer (Stone et al. 1966).

Proposed feature: the Pos:Neg ratio if that ratio is below 1 (lower quartile for the whole Pang & Lee data set) or above 1.76 (upper quartile), else 1.31 (the median). The goal is to single out ‘imbalanced’ reviews as potentially untrustworthy. (For a similar idea, see Pang et al. 2002.)
Looking ahead to Richard Socher’s lecture

1 Sentiment-relevant semantic influences can come from
   • negation
   • adverbs and other modifiers
   • attitude predications, including modals and hedges
   • and combinations of all of the above.

2 This is just to say that all aspects of semantic composition are relevant.

3 Thus, rather than treating it as series of isolated and separate problems, we should approach it as part of a theory of semantic composition.

4 This is precisely what Richard Socher is seeking to do (Socher et al. 2011). Lots more about that on Tuesday!
Conclusion

Central insights

- Sentiment is blended and continuous.
- Sentiment is social and context-dependent.
- Sentiment is as hard as semantic composition.

Opportunities

- Increasingly, we have the rights dataset and models to honor the above insights.
- Careful, flexible sentiment analysis systems are in high demand.
- Extensions:
  - How does sentiment flow in a social network?
  - How does it affect the flow of other information?
  - What does sentiment reveal about social ties, media bias, polarization, ...
References I


References II

Maas, Andrew; Andrew Ng; and Christopher Potts. 2011. Multi-dimensional sentiment analysis with learned representations. Ms., Stanford University.


References III


Thomas, Matt; Bo Pang; and Lillian Lee. 2006. Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In Proceedings of EMNLP, 327–335.


References IV

