Automatic Identification of Personal Insults on Social News Sites

Sara Owsley Sood
Outline

- Sentiment Analysis
- Problem
- Data
- Methods and Evaluation
- Future Work
Outline

- Sentiment Analysis
- Problem
- Data
- Methods and Evaluation
- Future Work
How do we make decisions?

Our decisions are influenced by...

- Friends
- Media
- Experts
Why Analyze Sentiment?

Customer Reviews

2,162 Reviews

- 5 star: 1,458
- 4 star: 368
- 3 star: 122
- 2 star: 69
- 1 star: 145

Average Customer Review ★★★★★ (2,162 customer reviews)

Most Helpful Customer Reviews

4,448 of 4,653 people found the following review helpful:

★★★★★ The Lines Between iPod Touch and iPhone Have Started to Blur, September 7, 2010

By Scott Showalter "purefusion" (Ohio, USA) - See all my reviews

This review is from: Apple iPod touch 32GB (4th Generation) - Black - Current Version (Electronics)

Having had a chance to spend a little time with a review model gives me a chance to share the experience with you a bit early (before my own arrives). I'll take you hands-on with the new model, plus I'll share from my past two years of iPod touch ownership altogether, especially for those
Sentiment Analysis
(Opinion Mining)

Automatically label documents with their ‘sentiment’

- Toward a topic
- Aggregated over documents
- More fine-grained analysis
- Within specific domains
Supervised Classification Example
Supervised Classification Example
Supervised Classification Example
Supervised Classification Example
Supervised Classification Example
Supervised Classification Example

Accuracy
Supervised Classification Example

Accuracy

15/20 = 0.75
Supervised Classification Example

Accuracy
\[ \frac{15}{20} = 0.75 \]

Precision
Supervised Classification Example

Accuracy
15/20 = 0.75

Precision
7/12 = 0.58
Supervised Classification Example

Accuracy
\[ \frac{15}{20} = 0.75 \]

Precision
\[ \frac{7}{12} = 0.58 \]

Recall
Supervised Classification
Example

Accuracy
$\frac{15}{20} = 0.75$

Precision
$\frac{7}{12} = 0.58$

Recall
$\frac{7}{7} = 1.0$
Supervised Classification Example

Accuracy
15/20 = 0.75

Precision
7/12 = 0.58

Recall
7/7 = 1.0

F1 (2PR/(P+R))
Supervised Classification Example

Accuracy
\[ \frac{15}{20} = 0.75 \]

Precision
\[ \frac{7}{12} = 0.58 \]

Recall
\[ \frac{7}{7} = 1.0 \]

F1 (\[2PR/(P+R)\])
\[ = 0.73 \]
Supervised ML in practice

- Supervised learning algorithm choice
  - Support Vector Machines
  - Naïve Bayes
  - Neural Networks
  - Decision Trees
- Configuration
- Feature Selection
- Training Data
Unsupervised Clustering Example
Unsupervised Clustering Example
This is good, because I recently met A, another grad student on the job hunt with a 1st author paper that's in the weeds. Her opinion as to why she has no interviews yet is that she doesn't have that 1st author Medline credit.

I really, really hope that's not it. Even if our paper goes in Sept 1, we likely wouldn't know about publication till Oct 1... and I don't wanna wait that long for people to call me back.

Now I'm wondering if I should even send out resumes without that... Yeah, I know, I still should. But...
What are the challenges to Sentiment Analysis?

- domain specificity
- “thwarted expectations”
- sarcasm and subtle nature of sentiment
- sufficient, high quality training data
What are the challenges to Sentiment Analysis?

Cold

Small
What are the challenges to Sentiment Analysis?

- domain specificity
- “thwarted expectations”
- sarcasm and subtle nature of sentiment
- sufficient, high quality training data
What are the challenges to Sentiment Analysis?

- domain specificity
- “thwarted expectations”
- sarcasm and subtle nature of sentiment
- sufficient, high quality training data

“This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can’t hold up.” (Pang et al, 2002)
What are the challenges to Sentiment Analysis?

- domain specificity
- “thwarted expectations”
- sarcasm and subtle nature of sentiment
- sufficient, high quality training data
What are the challenges to Sentiment Analysis?

- domain specificity
- “thwarted expectations”
- sarcasm and subtle nature of sentiment
- sufficient, high quality training data
Outline

● Sentiment Analysis
● **Problem**
● Data
● Methods and Evaluation
● Future Work
User Participation
Around Content
Number One Reason Sarah Palin Is Smiling Today

NBC Philadelphia - Submitted: Oct 2, 2009

Sarah Palin must be smiling, and not just because her new book is already #1 on the bestseller list, before it's even been released.

The ex-Alaska governor onetime nemesis, David Letterman, just confessed that he's been the Hollywood King of Late Night all these years, telling his audience he is the victim of an extortion plot and admitting, I have had sex with women who worked for me on...

At least he apologized; Palin accepted his apology so I'll follow her lead. She was right about Letterman, which we all knew—and hopefully America realizes she has been right about OH so much more.

Sarah Palin : "Todd ya so !"

GO SARAH !

You go, girl, we're behind you 100%. Yes, that's right, blue dogs are Palin Democrats!

The number 1 reason Sarah is smiling today, is because she just got a fat check for $7,000,000 bucks. Who wouldn't be smiling? Notice too, it's listed under the title, "ENTERTAINMENT". That's the most truthful thing about this story. Obviously, she coughed up a substantial amount of the millions to get Todd to keep quiet about their impending divorce, by buying him a new boat and sending off into the Bering Sea to do some fishing. What he doesn't know yet is, the hull of the boat is made of bamboo and paper machete.
C1: Man Palin is a has been stale news, good for small town hicks selling news to those like them. To them a pig farting is big news.

Community managers

.... content review .... moderation .... behavior management ....
Outline

- Sentiment Analysis
- Problem
- Data
- Methods and Evaluation
- Future Work
Data Set

- 168,973 distinct threads (each based on a news story)
- 1,655,131 comments (from March to May 2010)
Data Set

Distribution of Comments Among Users

Log of Number of Users

Number of Comments per User

0 500 1000 1500 2000
Crowdsourcing

Please choose the option which best describe the tone of this message:

- Negative  - Somewhat Negative  - Neutral  - Somewhat Positive  - Positive

Does this message contain any content you would describe as "profanity"?

- No  - Yes

Thinking about the intent of the comment's author, does the message contain what you would describe as an "insult"?

- No  - Yes

In your opinion, is the insult directed at the author of a previous comment?

Note: If the message contains more than one insult, please answer with respect to the portion you feel is most insulting.

- No  - Yes  - Unsure

Submit
Crowdsourcing

Mechanical Turk is a marketplace for work. We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.

205,351 HITs available. View them now.

Make Money by working on HITs

HITs - Human Intelligence Tasks - are individual tasks that you work on. Find HITs now.

As a Mechanical Turk Worker you:
- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task Work Earn money

or learn more about being a Worker

Get Results from Mechanical Turk Workers

Ask workers to complete HITs - Human Intelligence Tasks - and get results using Mechanical Turk. Register Now

As a Mechanical Turk Requester you:
- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account Load your tasks Get results

Get Started
Crowdsourcing

- Faster and more economical than using dedicated raters (Tetreault, Filatova & Chodorow, 2010)
- Reliable and high quality results (Sheng, Provost & Ipeirotis, 2008)
- Quality assurance via gold data and consensus thresholds
- Useful for
  - conducting user studies (Kittur, Chi & Suh, 2008)
  - translation task (Callison-Burch, 2009)
Crowdsourcing

- 221 turkers
- 6500 comments (11 or 12 turkers per comment)
- Agreement of 0.80 or higher on
  - 91% of comments for presence of profanity
  - 72% of comments for presence of insults
  - 34% of comments (that contain insults) for the object of the insult
Outline

- Sentiment Analysis
- Problem
- Data
- **Methods and Evaluation**
- Future Work
Systems

- Detecting Insults
- Classifying the Insult Object
- Detecting Personal Insults
- Detecting Profanity
Detecting Personal Insults – Relevance Analysis

<table>
<thead>
<tr>
<th>Comment</th>
<th>TFIDF Relevance Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The Palin family sucks, always have and always will. There are a few sex scandals in Sarah &amp; Todd’s past. Not to mention ethics problems etc. David is probably a nicer person than Sarah &amp; Todd put together. She can smile at Dave’s misfortune while conveniently forgetting her own sins.”</td>
<td>5.239</td>
</tr>
<tr>
<td>“Say what you will. Palin rooked her naysayers. You can call her book whatever you can to dilute her success but, WHO’S LAUGHING NOW! The Clowns and Comics are presently in conference crying and consoling each other with their big towels wiping off their makeup and drying their big old shoes. Waaaaahhhhhhh!”</td>
<td>3.754</td>
</tr>
<tr>
<td>“At least he apologized; Palin accepted his apology so I’ll follow her lead. She was right about Letterman, which we all knew--and hopefully America realizes she has been right about OH so much more.”</td>
<td>3.717</td>
</tr>
<tr>
<td>ClassyMingle.com is the best and largest online personals site dedicated to men and women seeking a higher caliber online dating experience.</td>
<td>0.359</td>
</tr>
<tr>
<td>Whyyyyyyyyyyyyyyyyy so many people are interested in an ageless relationship. Young girls want to have fun with 40+ men and young guys want to have fun with 40+ women. There are many sites focusing on this kind of relationships such as <a href="http://www.Seekingsugar.com">http://www.Seekingsugar.com</a></td>
<td>0.696</td>
</tr>
<tr>
<td>Who cares!!! My boyfriend thinks the same with me. He is eight years older than me, lol. We met online at Agelessmatch.com a nice and free place for Younger Women and Older Men, or Older Women and Younger Men, to interact with each other. Maybe you wanna check out or tell your friends.</td>
<td>0.351</td>
</tr>
<tr>
<td>Task</td>
<td>Method</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>Insult Detection</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sood</td>
</tr>
<tr>
<td></td>
<td>Random baseline</td>
</tr>
<tr>
<td><strong>Insult Object Classification</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sood</td>
</tr>
<tr>
<td></td>
<td>Random baseline</td>
</tr>
<tr>
<td><strong>Personal Insult Detection</strong></td>
<td></td>
</tr>
<tr>
<td>Sood (2 SVMs + relevance)</td>
<td></td>
</tr>
<tr>
<td>Sood (2 SVMs)</td>
<td></td>
</tr>
<tr>
<td>Yin, et al. – on Myspace</td>
<td></td>
</tr>
<tr>
<td>Yin et al. – on Slashdot</td>
<td></td>
</tr>
<tr>
<td>Random baseline</td>
<td></td>
</tr>
<tr>
<td><strong>Profanity Detection</strong></td>
<td></td>
</tr>
<tr>
<td>Sood</td>
<td></td>
</tr>
<tr>
<td>List-based baseline</td>
<td></td>
</tr>
<tr>
<td>Random baseline</td>
<td></td>
</tr>
</tbody>
</table>
Outline

- Sentiment Analysis
- Problem
- Data
- Method
- Evaluation
- Future Work
Future Work

- How will this integrate into a Community Management setting?
- Does our approach generalize to other sites/domains?
- If not, can we adapt the system to a new domain with unlabeled data from the target domain? Or do we need labeled data from each site?
- What impact do insults have on the comments that follow?
Differences Between Domains

<table>
<thead>
<tr>
<th>Category</th>
<th>N²</th>
<th>Profanity Occurrence (%)</th>
<th>Profanity ( \chi^2 )</th>
<th>Insult Occurrence (%)</th>
<th>Insult ( \chi^2 )</th>
<th>Directed Insult Occurrence (%)</th>
<th>Directed Insult ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>5905</td>
<td>9.40</td>
<td>--</td>
<td>20.60</td>
<td>--</td>
<td>11.00</td>
<td>--</td>
</tr>
<tr>
<td>Politics</td>
<td>1677</td>
<td>10.70</td>
<td>7.18†</td>
<td>26.80</td>
<td>72.92**</td>
<td>14.30</td>
<td>32.73***</td>
</tr>
<tr>
<td>News</td>
<td>1824</td>
<td>9.90</td>
<td>1.35</td>
<td>21.60</td>
<td>1.85</td>
<td>11.40</td>
<td>0.46</td>
</tr>
<tr>
<td>Business</td>
<td>476</td>
<td>9.70</td>
<td>0.059</td>
<td>16.70</td>
<td>4.96</td>
<td>9.50</td>
<td>1.25</td>
</tr>
<tr>
<td>Entertainment</td>
<td>686</td>
<td>9.30</td>
<td>0.0012</td>
<td>18.70</td>
<td>1.92</td>
<td>9.10</td>
<td>2.98</td>
</tr>
<tr>
<td>Health</td>
<td>177</td>
<td>9.00</td>
<td>0.023</td>
<td>14.10</td>
<td>4.62</td>
<td>4.80</td>
<td>6.66†</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>126</td>
<td>7.90</td>
<td>0.31</td>
<td>10.70</td>
<td>7.65*</td>
<td>1.70</td>
<td>10.83**</td>
</tr>
<tr>
<td>World</td>
<td>467</td>
<td>7.70</td>
<td>1.73</td>
<td>19.00</td>
<td>0.86</td>
<td>9.10</td>
<td>1.79</td>
</tr>
<tr>
<td>Science</td>
<td>225</td>
<td>6.70</td>
<td>2.05</td>
<td>14.60</td>
<td>4.91</td>
<td>9.90</td>
<td>0.29</td>
</tr>
<tr>
<td>Travel</td>
<td>18</td>
<td>5.60</td>
<td>0.31</td>
<td>18.80</td>
<td>0.034</td>
<td>6.70</td>
<td>0.29</td>
</tr>
<tr>
<td>Sports</td>
<td>229</td>
<td>5.20</td>
<td>4.87</td>
<td>14.70</td>
<td>5.21</td>
<td>3.80</td>
<td>12.04**</td>
</tr>
</tbody>
</table>

*** \( p \leq .001 \), ** \( p \leq .01 \), * \( p \leq .05 \), † \( p \leq .1 \)
Questions?