Developing a Flexible Sentiment Classification Technique for Multiple Domains

Nathan Agrin

School of Information
University of California, Berkeley

December 6th, 2006
Implications

- Find opinions on current events, products or specific interests
- Determine what people like about specific services, or products
- Can allow for more specific retrieval of opinionated content, and better mapping of a global sentiment, localized sentiment, as well as a specific user’s opinions on a given subject
Project Goals

• When searching, sentiment can be calculated at runtime or determined prior to a query and used to formulate the results

• Create a classification method with which to determine if text contains a positive or negative sentiment, then store this data in a format for assisting search
Data Set

- **Movie Review Set (Primary)**
  Created by Bo Pang and Lillian Lee at Cornell
  Contains 2,000 positive & negative movie reviews

- **Product Review Set (Secondary)**
  Created by Minqing Hu and Bing Liu
  Contains 110 negative and 185 positive product reviews

- **General Inquirer**
  Used as seed list and filter for affective words
Identified Problems

- Sentiment content often contains many discrete opinions about different aspects of a larger topic, or quotations of other text.
- Sentiment may use made-up words, or sarcasm:
  
  "Hmmmm, well, the main actor, Justin Chambers, is basically an uncharismatic version of Chris O'Donnell but with less range (think about that!), and Mena Suvari, is just plain off."

- Sentiment is often based on syntactical structure, implying negation:
  
  "I feel like I should have had a grand time with "Detroit Rock City."
  It's the sort of movie I wish I could've had a lot of fun with, but I didn't."
Approach

- **Statistical Classification**
  SVM and Complimentary Naive Bayes
  Tested across domains

- **Rule Based Classification**
  Used General Inquirer data as seed list
  Tested term expansion using Wordnet
### Statistical Classification

#### Baseline

- **SMO**
  - Same Domain: 76
  - Cross Domain: 65
  - Accuracy: 76

- **CNB**
  - Same Domain: 80
  - Cross Domain: 68
  - Accuracy: 71

#### Only Affective Words

- **SMO**
  - Same Domain: 71
  - Cross Domain: 60
  - Accuracy: 65

- **CNB**
  - Same Domain: 79
  - Cross Domain: 65
  - Accuracy: 65
Rule Based Classifier

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Wordnet</th>
<th>Wordnet &amp; Negation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pos</td>
<td>893</td>
<td>862</td>
<td>890</td>
</tr>
<tr>
<td></td>
<td>726</td>
<td>695</td>
<td>685</td>
</tr>
<tr>
<td>neg</td>
<td>107</td>
<td>102</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>274</td>
<td>245</td>
<td>236</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Documents consistently scored as positive
- 1638 / 2012: positive / negative words from GI
- Could not determine the cause of accuracy issue...
Rule Based Classifier

- Many documents incorrectly classified as positive had a very small positive rating
- Increasing positive from $> 0$ to $> 5$ helped

<table>
<thead>
<tr>
<th>Wordnet &amp; Negation</th>
<th>p</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>pos</td>
<td>890</td>
<td>685</td>
</tr>
<tr>
<td>neg</td>
<td>81</td>
<td>236</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wordnet &amp; Negation (Positive $&gt; 5$)</th>
<th>p</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>pos</td>
<td>623</td>
<td>290</td>
</tr>
<tr>
<td>neg</td>
<td>377</td>
<td>710</td>
</tr>
</tbody>
</table>
**Discussion**

- Statistical Classifier preformed best with little extra data, and across domains
- Rule based classifier may be able to compete with statistical classifier in diverse domains
- Docs tended to contain many positive affective words indicating their POS is misinterpreted, or they appear more frequently, even in negative text
- Use POS tagging and chunking to train a classifier
Questions?