Sentiment Analysis

Positive or negative movie review?

- unbelievably disappointing
- Full of zany characters and richly applied satire, and some great plot twists
- this is the greatest screwball comedy ever filmed
- It was pathetic. The worst part about it was the boxing scenes.

Twitter sentiment versus Gallup Poll of Consumer Confidence

Target Sentiment on Twitter

Sentiment analysis has many other names
- Opinion extraction
- Opinion mining
- Sentiment mining
- Subjectivity analysis

Why sentiment analysis?
- Movie: is this review positive or negative?
- Products: what do people think about the new iPhone?
- Public sentiment: how is consumer confidence? Is despair increasing?
- Politics: what do people think about this candidate or issue?
- Prediction: predict election outcomes or market trends from sentiment

Scherer Typology of Affective States
- Emotion: brief, organically synchronized reaction to the evaluation of a major event
  - angry, sad, joyful, fearful, ashamed, proud, elated
- Mood: diffuse, non-triggered low-intensity, long-duration change in subjective feeling
  - cheerful, gloomy, irritable, listless, depressed, buoyant
- Interpersonal stances: affective stance toward another person in a specific interaction
  - friendly, flirtatious, distant, cold, warm, supportive, contemptuous
- Attitudes: enduring, affectively colored beliefs, dispositions towards objects or persons
  - liking, loving, hating, valuing, desiring
- Personality traits: stable personality dispositions and typical behavior tendencies
  - anxious, reckless, morose, hostile, jealous

Sentiment Analysis
- Sentiment analysis is the detection of attitudes
  - enduring, affectively colored beliefs, dispositions towards objects or persons
  - Emily told Charlie that the new movie is disappointing.

  1. Holder (source) of attitude
  2. Target (aspect) of attitude
  3. Type of attitude
    - From a set of types
      - sad, love, hate, value, desire, etc.
      - Or (more commonly) simple weighted polarity:
        - positive, negative, neutral, together with strength
  4. Text containing the attitude
    - Sentence or entire document

  Simplest task:
  - Is the attitude of this text positive or negative?
  - More complex:
    - Rank the attitude of this text from 1 to 5
  - Advanced:
    - Detect the target, source, or complex attitude types
Sentiment Analysis

- Simplest task:
  - Is the attitude of this text positive or negative?
- More complex:
  - Rank the attitude of this text from 1 to 5
- Advanced:
  - Detect the target, source, or complex attitude types

Sentiment Classification in Movie Reviews


- Polarity detection:
  - Is an IMDB movie review positive or negative?
- Data: Polarity Data 2.0:
  - http://www.cs.cornell.edu/people/pabo/movie-review-data

Baseline Algorithm (adapted from Pang and Lee)

- Tokenization
- Feature Extraction
- Classification using different classifiers
  - Naïve Bayes
  - MaxEnt
  - SVM

IMDB data in the Pang and Lee database

when _star wars_ came out some twenty years ago , the image of traveling throughout the stars has become a commonplace image . [...] when han solo goes light speed , the stars change to bright lines , going towards the viewer in lines that converge at an invisible point .
cool . _october sky_ offers a much simpler image – that of a single white dot , traveling horizontally across the night sky . [...]

“snake eyes” is the most aggravating kind of movie : the kind that shows so much potential then becomes unbelievably disappointing . It’s not just because this is a brian de palma film , and since he’s a great director and one who’s films are always greeted with at least some fanfare . and it’s not even because this was a film starring nicolas cage and since he gives a brauvara performance , this film is hardly worth his talents

Sentiment Tokenization Issues

- Deal with HTML and XML markup
- Twitter mark-up (names, hash tags)
- Capitalization (preserve for words in all caps)
- Phone numbers, dates
- Emoticons
- Usefull code:
  - Christopher Potts sentiment tokenizer
  - Brendan O’Connor twitter tokenizer

Pre-processing Social Media Text

- Social Media Text is noisy
  - Informal e.g., slang
  - Misspellings e.g., covfefe
  - Elongated words e.g., can’t waittt
  - Hashtags
  - Emoticons
  - Lits
  - Random capitalization e.g., NOT COOL!
Pre-processing: Hashtags

- Hashtagged words are good labels of sentiments and emotions
- Some jerk just stole my photo on tumblr! #rgranger
- Hashtag Sentiment Lexicon
  - created from a large collection of hashtagged tweets
  - New hashtags are being generated every minute
  - Breaking long hashtags into smaller instances
    - #killthebill -> kill the bill

Extracting Features for Sentiment Classification

- How to handle negation
  - I didn’t like this movie vs
  - I really like this movie
- Which words to use?
  - Only adjectives
  - All words
  - All words often turn out to work better

Negation

Add NOT_ to every word between negation and following punctuation:

didn’t like this movie, but I

didn’t NOT_like NOT_this NOT_movie but I

Remainder: Naïve Bayes

\[ c_{NB} = \arg \max_{c \in C} P(c) \prod_{i \in \text{positions}} P(w_i | c) \]

\[ \hat{P}(w | c) = \frac{\text{count}(w, c) + 1}{\text{count}(c) + |V|} \]

Binarized (Boolean feature)

- Intuition:
  - For sentiment (and probably for other text classification domains)
  - Word occurrence may matter more than word frequency
    - The occurrence of the word 'fantastic' tells us a lot
    - The fact that it occurs 5 times may not tell us much more.
  - Boolean Multinomial Naïve Bayes
    - Clips all the word counts in each document at 1

Normal vs. Boolean Multinomial NB

<table>
<thead>
<tr>
<th></th>
<th>Doc</th>
<th>Words</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>1</td>
<td>Chinese Beijing</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Chinese Shanghai</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Chinese Macao</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Tokyo Japan-Chinese</td>
<td>j</td>
</tr>
<tr>
<td>Test</td>
<td>5</td>
<td>Chinese Chinese Tokyo Japan</td>
<td>?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Doc</th>
<th>Words</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>1</td>
<td>Chinese Beijing</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Chinese Shanghai</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Chinese Macao</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Tokyo Japan-Chinese</td>
<td>j</td>
</tr>
<tr>
<td>Test</td>
<td>5</td>
<td>Chinese Tokyo Japan</td>
<td>?</td>
</tr>
</tbody>
</table>
Binarized (Boolean feature) Multinomial Naive Bayes

• Binary seems to work better than full word counts
• Other possibility: $\log(\text{freq}(w))$

Problems: What makes reviews hard to classify?

• Subtlety:
  • Perfume review in Perfumes: the Guide:
    • "If you are reading this because it is your darling fragrance, please wear it at home exclusively, and tape the windows shut."
    • Dorothy Parker on Katherine Hepburn
    • "She runs the gamut of emotions from A to B"

Thwarted Expectations and Ordering Effects

• "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."
• Well as usual Keanu Reeves is nothing special, but surprisingly, the very talented Laurence Fishbourne is not so good either, I was surprised.

Sentiment Lexicons

The General Inquirer

• Home page: http://www.wjh.harvard.edu/~inquirer
• List of Categories: http://www.wjh.harvard.edu/~inquirer/inquirerbasic.xls
• Categories:
  • Positive (1915 words) and Negative (2291 words)
  • Strong vs Weak, Active vs Passive, Overstated versus Understated
  • Pleasure, Pain, Virtue, Vice, Motivation, Cognitive-Orientatio, etc
• Free for Research Use

LIWC (Linguistic Inquiry and Word Count)

• Home page: http://www.liwc.net/
• 2300 words, >70 classes
• Affective Processes
  • negative emotion (bad, weird, hate, problem, tough)
  • positive emotion (love, nice, sweet)
• Cognitive Processes
  • Tentative (may be, perhaps, guess), Inhibition (block, constraint)
• Pronouns, Negation (no, never), Quantifiers (few, many)
• Not free though!
SentiWordNet
Stefano Baccianelli, Andrea Esuli, and Fabrizio Sebastiani. 2010 SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC 2010
• Home page: http://sentiwordnet.isti.cnr.it/
• 6885 words from 8221 lemmas
• 2718 positive
• 4912 negative
• Each word annotated for intensity (strong, weak)
• GNU GPL

Bing Liu Opinion Lexicon
• Bing Liu’s Page on Opinion Mining
  http://www.cs.uic.edu/~liub/FBS/opinion-lexicon
  http://sentiwordnet.isti.cnr.it/
• 6786 words
• 2006 positive
• 4783 negative

Disagreements between polarity lexicons
Christopher Potts, Sentiment Tutorial, 2011

<table>
<thead>
<tr>
<th></th>
<th>Opinion Lexicon</th>
<th>General Inquirer</th>
<th>SentiWordNet</th>
<th>LIWC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPQA</td>
<td>32/2402 (0.6%)</td>
<td>32/2411 (1%)</td>
<td>32/2411 (1%)</td>
<td>9/403 (2%)</td>
</tr>
<tr>
<td>Opinion Lexicon</td>
<td>32/2402 (0.6%)</td>
<td>32/2411 (1%)</td>
<td>32/2411 (1%)</td>
<td>9/403 (2%)</td>
</tr>
<tr>
<td>General Inquirer</td>
<td>520/2306 (23%)</td>
<td>520/2306 (23%)</td>
<td>520/2306 (23%)</td>
<td>1/204 (0.5%)</td>
</tr>
<tr>
<td>SentiWordNet</td>
<td></td>
<td></td>
<td>1004/3994 (27%)</td>
<td></td>
</tr>
<tr>
<td>LIWC</td>
<td></td>
<td></td>
<td>1004/3994 (27%)</td>
<td>174/694 (25%)</td>
</tr>
</tbody>
</table>

Analyzing the polarity of each word in IMDB
• How likely is each word to appear in each sentiment class?
• Count("bad") in 1-star, 2-star, 3-star, etc.
• But can’t use raw counts.
• Instead, likelihood:
  \[ P(w|c) \propto \sum_{w} f(w,c) \]
• Make them comparable between words:
  • Scaled likelihood:
    \[ \frac{P(w|c)}{P(w)} \]
Other sentiment feature: Logical negation


- Is logical negation (no, not) associated with negative sentiment?
- Potts experiment:
  - Count negation (not, n't, no, never) in online reviews
  - Regress against the review rating

Potts 2011 Results: More negation in negative sentiment

![Graph showing negation counts across review ratings](image)

Learning Sentiment Lexicons

Semi-supervised learning of lexicons

- Use a small amount of information
  - A few labeled examples
  - A few hand-built patterns
  - To bootstrap a lexicon

Hatzivassiloglou and McKeown intuition for identifying word polarity


- Adjectives conjoined by “and” have same polarity
  - *fair and legitimate, corrupt and brutal*
- Adjectives conjoined by “but” do not have the same polarity
  - *fair but brutal*

Hatzivassiloglou & McKeown 1997

Step 1

- Label seed set of 1336 adjectives (all >20 in 21 million word WSJ corpus)
  - 657 positive
    - adequate, central, clever, famous, intelligent, remarkable, reputed, sensitive, slender, thriving...
  - 679 negative
    - contagious, drunken, ignorant, lanky, listless, primitive, strident, troublesome, unresolved, unsuspecting...
Hatzivassiloglou & McKeown 1997

Step 2
• Expand seed set to conjoined adjectives

Step 3
• Supervised classifier assigns “polarity similarity” to each word pair, resulting in graph:

Hatzivassiloglou & McKeown 1997

Step 4
• Clustering for partitioning the graph into two

Output polarity lexicon
• Positive
  • bold decisive disturbing generous good honest important large mature patient peaceful positive proud sound stimulating straightforward strange talented vigorous witty...
  • Negative
  • ambiguous cautious cynical evasive harmful hypocritical inefficient insecure irrational irresponsible minor outspoken pleasant reckless risky selfish tedious unsupported vulnerable wasteful...

Turney Algorithm
1. Extract a phrasal lexicon from reviews
2. Learn polarity of each phrase
3. Rate a review by the average polarity of its phrases
Extract two-word phrases with adjectives

<table>
<thead>
<tr>
<th>First Word</th>
<th>Second Word</th>
<th>Third Word (not extracted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JJ</td>
<td>NN or NNS</td>
<td>anything</td>
</tr>
<tr>
<td>RB, RBR, RBS</td>
<td>JJ</td>
<td>Not NN nor NNS</td>
</tr>
<tr>
<td>JJ</td>
<td>JJ</td>
<td>Not NN or NNS</td>
</tr>
<tr>
<td>NN or NNS</td>
<td>JJ</td>
<td>Not NN nor NNS</td>
</tr>
<tr>
<td>RB, RBR, RBS</td>
<td>VB, VBD, VBN, VBG</td>
<td>anything</td>
</tr>
</tbody>
</table>

How to measure polarity of a phrase?

- Positive phrases co-occur more with “excellent”
- Negative phrases co-occur more with “poor”
- But how to measure co-occurrence?

Pointwise Mutual Information

- **Pointwise mutual information:**
  - How much more do events x and y co-occur than if they were independent?
  \[ \text{PMI}(X,Y) = \log_2 \frac{P(x,y)}{P(x)P(y)} \]
- **PMI between two words:**
  - How much more do two words co-occur than if they were independent?
  \[ \text{PMI}(\text{word}_1, \text{word}_2) = \log_2 \frac{P(\text{word}_1, \text{word}_2)}{P(\text{word}_1)P(\text{word}_2)} \]

How to Estimate Pointwise Mutual Information

- Query search engine
  - \[ P(\text{word}) \text{ estimated by } \frac{\text{hits(\text{word})}}{N} \]
  - \[ P(\text{word}_1, \text{word}_2) \text{ by } \frac{\text{hits(\text{word}_1 \text{ NEAR word}_2)}}{N} \]
  \[ \text{PMI}(\text{word}_1, \text{word}_2) = \log_2 \frac{\frac{\text{hits(\text{word}_1 \text{ NEAR word}_2)}}{N}}{\frac{\text{hits(\text{word}_1)}}{N} \times \frac{\text{hits(\text{word}_2)}}{N}} \]

Does phrase appear more with “poor” or “excellent”? 

\[ \text{Polarity(phrase)} = \text{PMI}(\text{phrase}, \text{"excellent"}) - \text{PMI}(\text{phrase}, \text{"poor"}) \]

Phrases from a thumbs-up review

<table>
<thead>
<tr>
<th>Phrase</th>
<th>POS tags</th>
<th>Polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>online service</td>
<td>JJ NN</td>
<td>2.8</td>
</tr>
<tr>
<td>online experience</td>
<td>JJ NN</td>
<td>2.3</td>
</tr>
<tr>
<td>direct deposit</td>
<td>JJ NN</td>
<td>1.3</td>
</tr>
<tr>
<td>local branch</td>
<td>JJ NN</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low fees</td>
<td>JJ NNS</td>
<td>0.33</td>
</tr>
<tr>
<td>true service</td>
<td>JJ NN</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>other bank</td>
<td>JJ NN</td>
<td>-0.85</td>
</tr>
<tr>
<td>inconveniently located</td>
<td>JJ NNS</td>
<td>-1.5</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.32</td>
</tr>
</tbody>
</table>
Phrases from a thumbs-down review

<table>
<thead>
<tr>
<th>Phrase</th>
<th>POS tags</th>
<th>Polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>direct deposits</td>
<td>JJ NNS</td>
<td>5.8</td>
</tr>
<tr>
<td>online web</td>
<td>JJ NN</td>
<td>1.9</td>
</tr>
<tr>
<td>very handy</td>
<td>RB JJ</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>virtual monopoly</td>
<td>JJ NNS</td>
<td>-2.0</td>
</tr>
<tr>
<td>lesser evil</td>
<td>RBR JJ</td>
<td>-2.3</td>
</tr>
<tr>
<td>other problems</td>
<td>JJ NNS</td>
<td>-2.8</td>
</tr>
<tr>
<td>low funds</td>
<td>JJ NNS</td>
<td>-6.8</td>
</tr>
<tr>
<td>unethical practices</td>
<td>JJ NNS</td>
<td>-8.5</td>
</tr>
<tr>
<td>Average</td>
<td>JJ NNS</td>
<td>-1.2</td>
</tr>
</tbody>
</table>

Results of Turney algorithm

- 410 reviews from Epinions
- 170 (41%) negative
- 240 (59%) positive
- Majority class baseline: 59%
- Turney algorithm: 74%
- Phrases rather than words
- Learns domain-specific information

Using WordNet to learn polarity

- WordNet: online thesaurus (covered in later lecture).
- Create positive ("good") and negative seed-words ("terrible")
- Find Synonyms and Antonyms
  - Positive Set: Add synonyms of positive words ("well") and antonyms of negative words ("awful")
  - Negative Set: Add synonyms of negative words ("awful") and antonyms of positive words ("well")
- Repeat, following chains of synonyms
- Filter

Other Sentiment Tasks

- Important for finding aspects or attributes
- Target of sentiment
- The food was great but the service was awful

Finding aspect/attribute/target of sentiment

- Frequent phrases + rules
  - Find all highly frequent phrases across reviews ("fish tacos")
  - Filter by rules like "occurs right after sentiment word"
  - ".great fish tacos" means fish tacos a likely aspect

<table>
<thead>
<tr>
<th>Casino</th>
<th>casino, buffet, pool, resort, beds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children's Barber</td>
<td>haircut, job, experience, kids</td>
</tr>
<tr>
<td>Greek Restaurant</td>
<td>food, wine, service, appetizer, lamb</td>
</tr>
<tr>
<td>Department Store</td>
<td>selection, department, sales, shop, clothing</td>
</tr>
</tbody>
</table>
Finding aspect/attribute/target of sentiment

- The aspect name may not be in the sentence
- For restaurants/hotels, aspects are well-understood
- Supervised classification
  - Hand-label a small corpus of restaurant review sentences with aspect
    - food, décor, service, value, NONE
  - Train a classifier to assign an aspect to a sentence
    - "Given this sentence, is the aspect food, décor, service, value, or NONE?"

Results of Blair-Goldensohn et al. method

- **Rooms** (3/5 stars, 41 comments)
  - (+) the room was clean and everything worked fine – even the water pressure ...
  - (+) We went because of the free room and was pleasantly pleased ...
  - (-) the worst hotel I had ever stayed at ...

- **Service** (3/5 stars, 31 comments)
  - (+) upon checking out another couple was checking early due to a problem ...
  - (+) every single hotel staff member treated us great and answered every ...
  - (-) the food is cold and the service gives new meaning to SLOW

- **Dining** (3/5 stars, 18 comments)
  - (+) our favorite place to stay in biloxi. the food is great also the service ...
  - (+) offer of free buffet for joining the Play

Summary on Sentiment

- Generally modeled as classification or regression task
  - predict a binary or ordinal label
- Features:
  - Negation is important
  - Using all words (in naïve bayes) works well for some tasks
  - Finding subsets of words may help in other tasks
    - Hand-built polarity lexicons
    - Use seeds and semi-supervised learning to induce lexicons