Sentiment Analysis

What is 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.

Google Product Search

HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner

$89 online, **$100** nearby ★★★★☆ 377 reviews

September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 sh

Reviews

Summary - Based on 377 reviews

1 star | 2 | 3 | 4 stars | 5 stars

What people are saying

- ease of use — "This was very easy to setup to four computers."
- value — "Appreciate good quality at a fair price."
- setup — "Overall pretty easy setup."
- customer service — "I DO like honest tech support people."
- size — "Pretty Paper weight."
- mode — "Photos were fair on the high quality mode."
- colors — "Full color prints came out with great quality."

Bing Shopping

HP Officejet 6500A E710N Multifunction Printer

Product summary | Find best price | **Customer reviews** | Specifications | Related items

$121.53 - $242.39 (14 stores)

☐ Compare

Average rating ★★★½☆ (144)

★★★★★		(55)
★★★★☆		(54)
★★★☆☆		(10)
★★☆☆☆		(6)
★☆☆☆☆		(23)
☆☆☆☆☆		(0)

Most mentioned

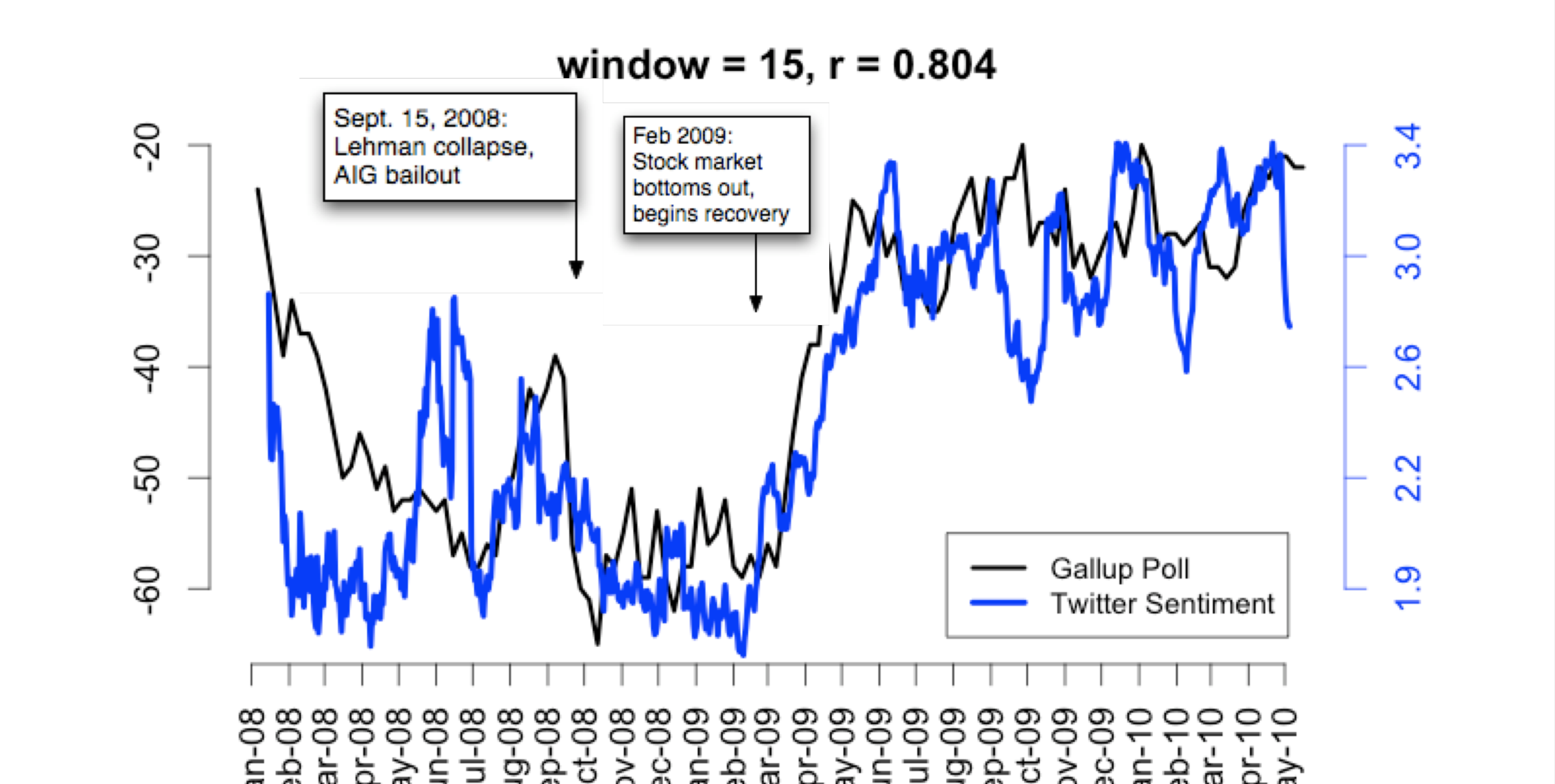
Performance		(57)
Ease of Use		(43)
Print Speed		(39)
Connectivity		(31)
More ▼		

Show reviews by source

Best Buy (140)
CNET (5)
Amazon.com (3)

Twitter sentiment versus Gallup Poll of Consumer Confidence

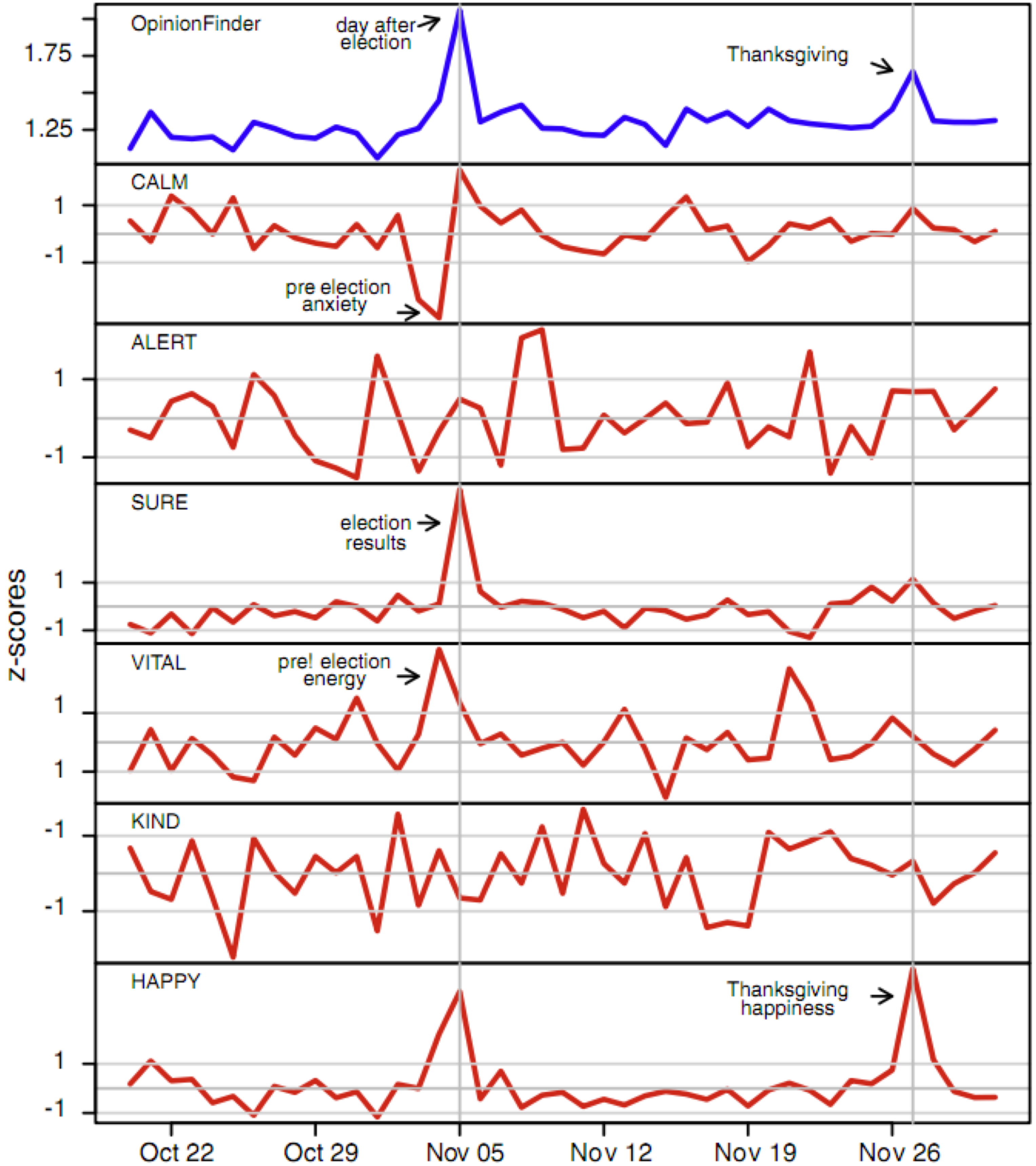
Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. 2010.
From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. In ICWSM-2010



Twitter sentiment:

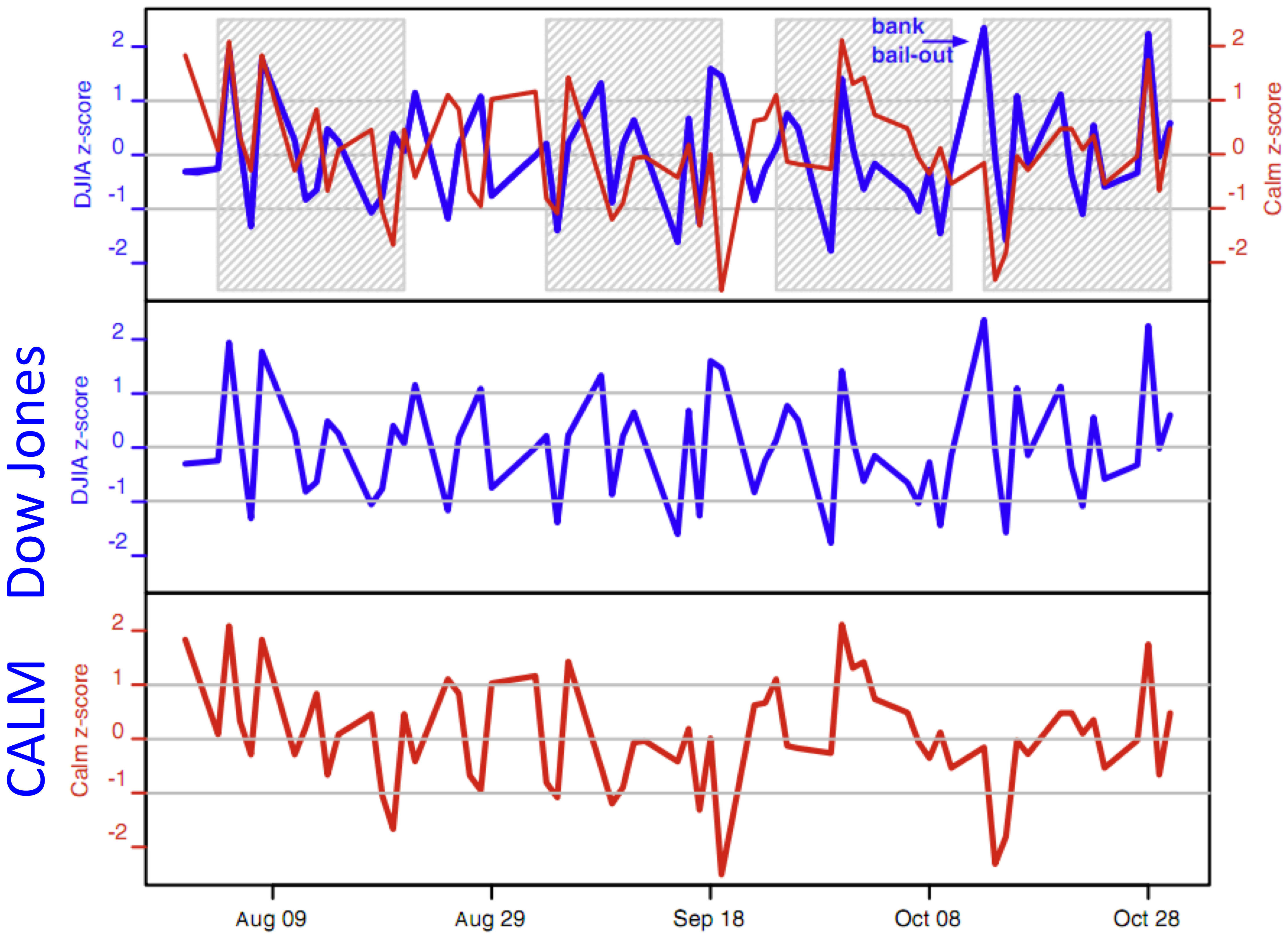
Johan Bollen, Huina Mao, Xiaojun Zeng. 2011.
Twitter mood predicts the stock market,

Journal of Computational Science 2:1, 1-8. 10.1016/j.jocs.2010.12.007.



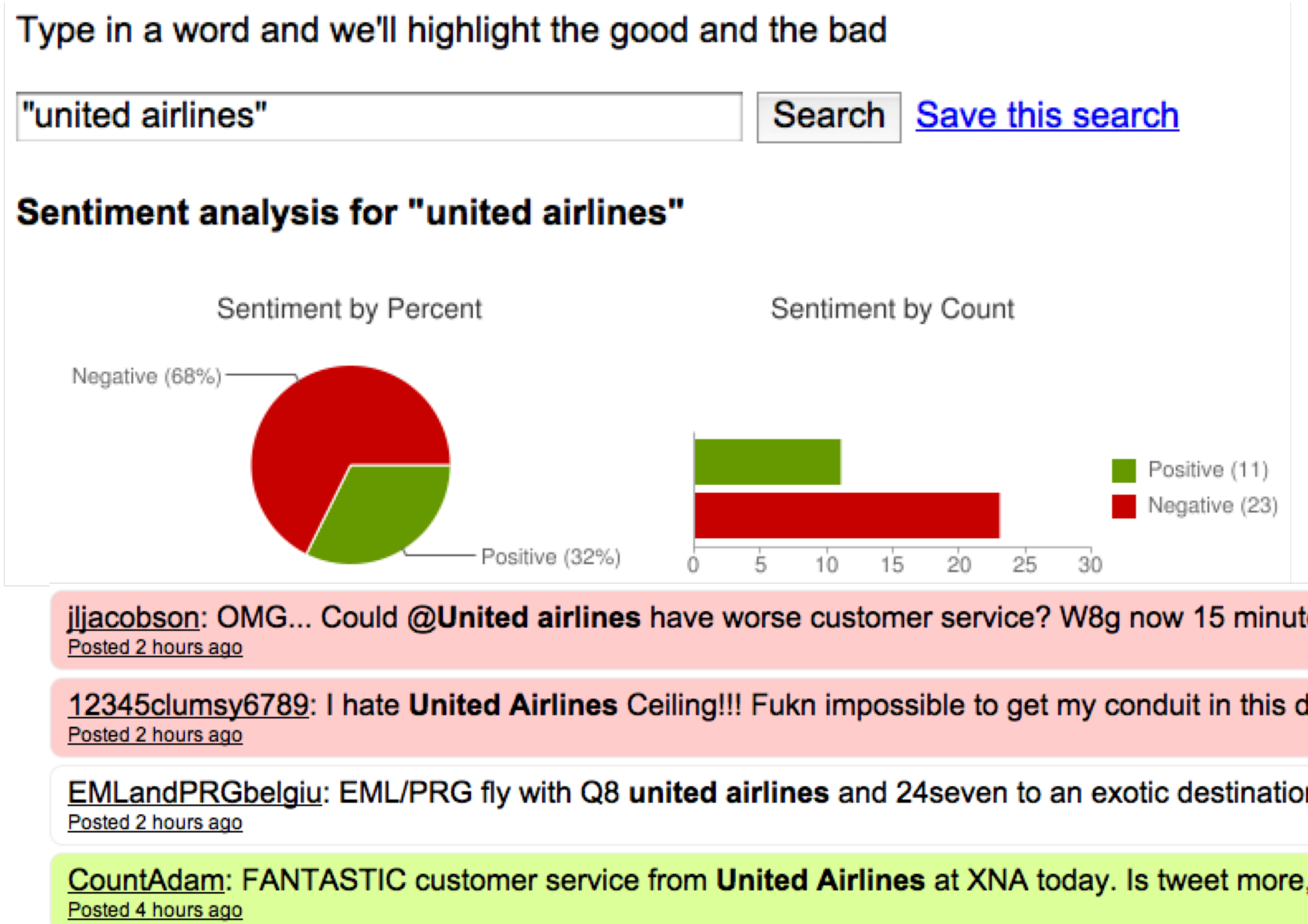
Bollen et al. (2011)

- CALM predicts DJIA 3 days later
- At least one current hedge fund uses this algorithm



Target Sentiment on Twitter

- Twitter Sentiment App
- Alec Go, Richa Bhayani, Lei Huang. 2009. Twitter Sentiment Classification using Distant Supervision



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 … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood**: diffuse non-caused 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
  - *nervous, anxious, reckless, morose, hostile, jealous*

Scherer Typology of Affective States

- **Emotion**: brief organically synchronized … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood**: diffuse non-caused 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
  - *nervous, anxious, reckless, morose, hostile, jealous*

Sentiment Analysis

- Sentiment analysis is the detection of **attitudes**

  "enduring, affectively colored beliefs, dispositions towards objects or persons"

  1. **Holder (source)** of attitude
  2. **Target (aspect)** of attitude
  3. **Type** of attitude
     - From a set of types
       - *Like, 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

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 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 Analysis

What is Sentiment Analysis?

Sentiment Analysis

A Baseline Algorithm

Sentiment Classification in Movie Reviews

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.
Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. ACL, 271-278

- 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

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 depalma 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 .

Baseline Algorithm (adapted from Pang and Lee)

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

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
- Useful code:
  - [Christopher Potts sentiment tokenizer]
  - [Brendan O'Connor twitter tokenizer]

Potts emoticons

```
[<>]?                          # optional hat/brow
[:;=8]                         # eyes
[\-o\*\']?                     # optional nose
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
|                              #### reverse orientation
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
[\-o\*\']?                     # optional nose
[:;=8]                         # eyes
[<>]?                          # optional hat/brow
```

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 turns out to work better, at least on this data

Negation

Das, Sanjiv and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific Finance Association Annual Conference (APFA).
Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.

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
```

Reminder: Naïve Bayes

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(w_i \mid c_j)$$

$$\hat{P}(w \mid c) = \frac{count(w,c)+1}{count(c)+|V|}$$

Binarized (Boolean feature) Multinomial Naïve Bayes

- 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

Boolean Multinomial Naïve Bayes: Learning

- From training corpus, extract *Vocabulary*

- Calculate $P(c_j)$ terms
  - For each c_j in C do

    $docs_j \leftarrow$ all docs with class $=c_j$

$$P(c_j) \leftarrow \frac{|docs_j|}{|\text{total \# documents}|}$$

- Calculate $P(w_k \mid c_j)$ terms
  - Remove duplicates in each doc:
    - For each word type w in doc_j
      - Retain only a single instance of w
  - $Text_j \leftarrow$ single doc containing all $docs_j$
  - For each word w_k in *Vocabulary*

    $n_k \leftarrow$ # of occurrences of w_k in $Text_j$

$$P(w_k \mid c_j) \leftarrow \frac{n_k + \alpha}{n + \alpha |Vocabulary|}$$

Boolean Multinomial Naïve Bayes on a test document *d*

- First remove all duplicate words from *d*
- Then compute NB using the same equation:

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(w_i \mid c_j)$$

Normal vs. Boolean Multinomial NB

Normal	Doc	Words	Class
Training	1	Chinese Beijing Chinese	c
	2	Chinese Chinese Shanghai	c
	3	Chinese Macao	c
	4	Tokyo Japan Chinese	j
Test	5	Chinese Chinese Chinese Tokyo Japan	?

Boolean	Doc	Words	Class
Training	1	Chinese Beijing	c
	2	Chinese Shanghai	c
	3	Chinese Macao	c
	4	Tokyo Japan Chinese	j
Test	5	Chinese Tokyo Japan	?

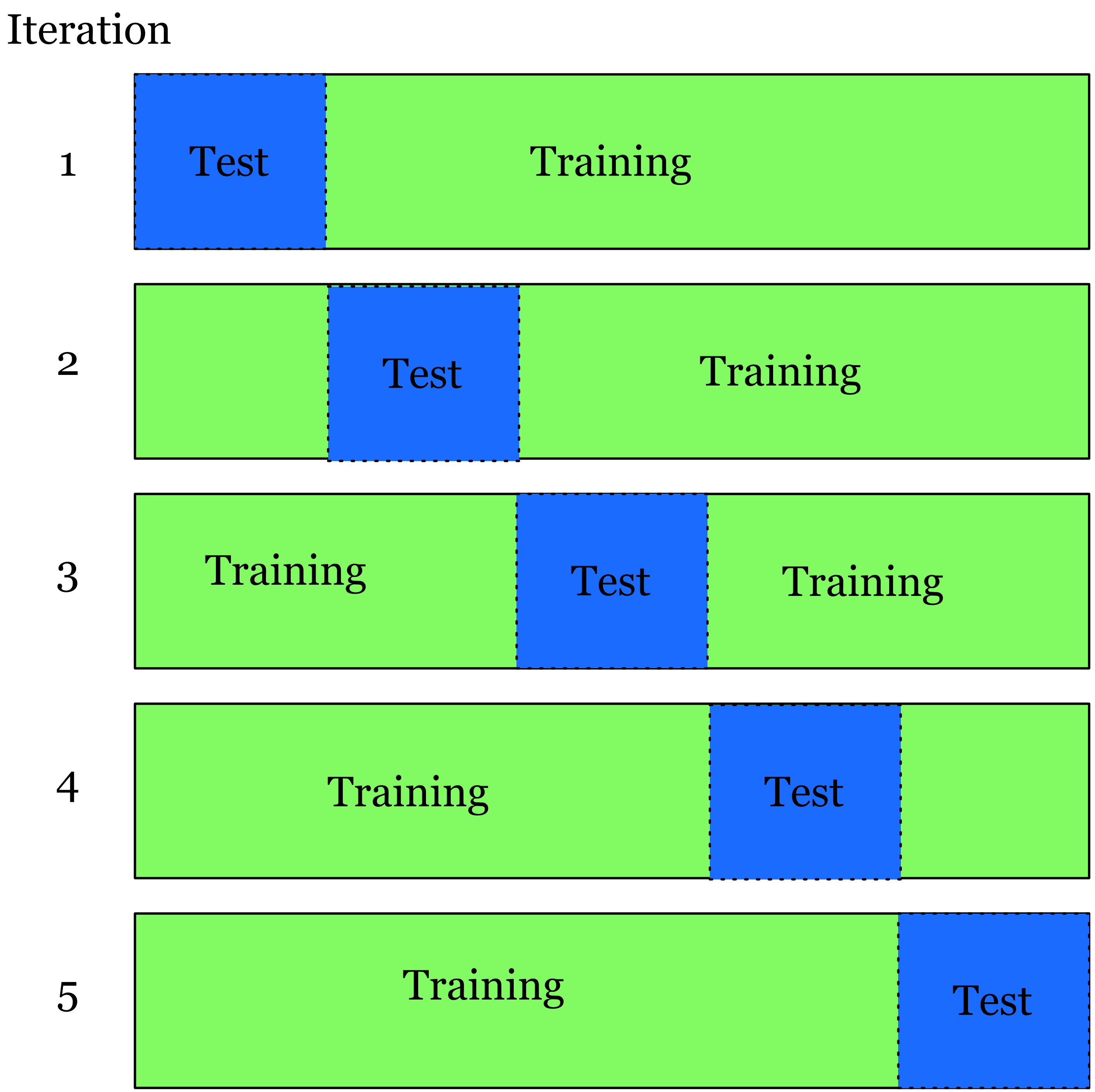
Binarized (Boolean feature) Multinomial Naïve Bayes

B. Pang, L. Lee, and S. Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.
V. Metsis, I. Androutsopoulos, G. Paliouras. 2006. Spam Filtering with Naive Bayes – Which Naive Bayes? CEAS 2006 - Third Conference on Email and Anti-Spam.
K.-M. Schneider. 2004. On word frequency information and negative evidence in Naive Bayes text classification. ICANLP, 474-485.
JD Rennie, L Shih, J Teevan. 2003. Tackling the poor assumptions of naive bayes text classifiers. ICML 2003

- Binary seems to work better than full word counts
  - This is **not** the same as Multivariate Bernoulli Naïve Bayes
    - MBNB doesn't work well for sentiment or other text tasks
- Other possibility: log(freq(*w*))

Cross-Validation

- Break up data into 10 folds
  - (Equal positive and negative inside each fold?)
- For each fold
  - Choose the fold as a temporary test set
  - Train on 9 folds, compute performance on the test fold
- Report average performance of the 10 runs



Other issues in Classification

- MaxEnt and SVM tend to do better than Naïve Bayes

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 Analysis

A Baseline Algorithm

Sentiment Analysis

Sentiment Lexicons

The General Inquirer

Philip J. Stone, Dexter C Dunphy, Marshall S. Smith, Daniel M. Ogilvie. 1966. The General Inquirer: A Computer Approach to Content Analysis. MIT Press

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

LIWC (Linguistic Inquiry and Word Count)

Pennebaker, J.W., Booth, R.J., & Francis, M.E. (2007). Linguistic Inquiry and Word Count: LIWC 2007. Austin, TX

- 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 (*maybe, perhaps, guess*), Inhibition (*block, constraint*)
- **Pronouns, Negation** (*no, never*), **Quantifiers** (*few, many*)
- $30 or $90 fee

MPQA Subjectivity Cues Lexicon

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proc. of HLT-EMNLP-2005.

Riloff and Wiebe (2003). Learning extraction patterns for subjective expressions. EMNLP-2003.

- Home page: http://www.cs.pitt.edu/mpqa/subj_lexicon.html
- 6885 words from 8221 lemmas
  - 2718 positive
  - 4912 negative
- Each word annotated for intensity (strong, weak)
- GNU GPL

Bing Liu Opinion Lexicon

Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. ACM SIGKDD-2004.

- Bing Liu's Page on Opinion Mining
- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

- 6786 words
  - 2006 positive
  - 4783 negative

SentiWordNet

Stefano Baccianella, 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/
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
- [estimable(J,3)] "may be computed or estimated"

  `Pos   0    Neg 0    Obj 1`

- [estimable(J,1)] "deserving of respect or high regard"

  `Pos .75  Neg 0    Obj .25`

Disagreements between polarity lexicons

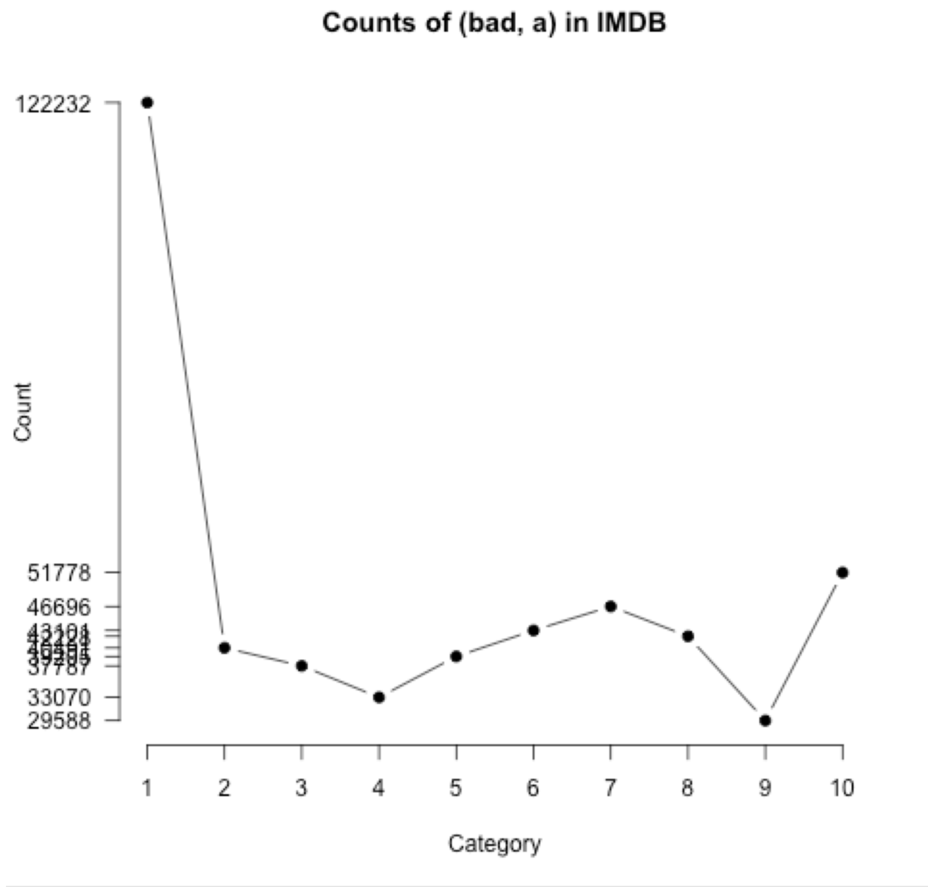
Christopher Potts, Sentiment Tutorial, 2011

	Opinion Lexicon	General Inquirer	SentiWordNet	LIWC
MPQA	33/5402 **(0.6%)**	49/2867 **(2%)**	1127/4214 **(27%)**	12/363 **(3%)**
Opinion Lexicon		32/2411 **(1%)**	1004/3994 **(25%)**	9/403 **(2%)**
General Inquirer			520/2306 **(23%)**	1/204 **(0.5%)**
SentiWordNet				174/694 **(25%)**
LIWC				

Analyzing the polarity of each word in IMDB

Potts, Christopher. 2011. On the negativity of negation. SALT 20, 636-659.

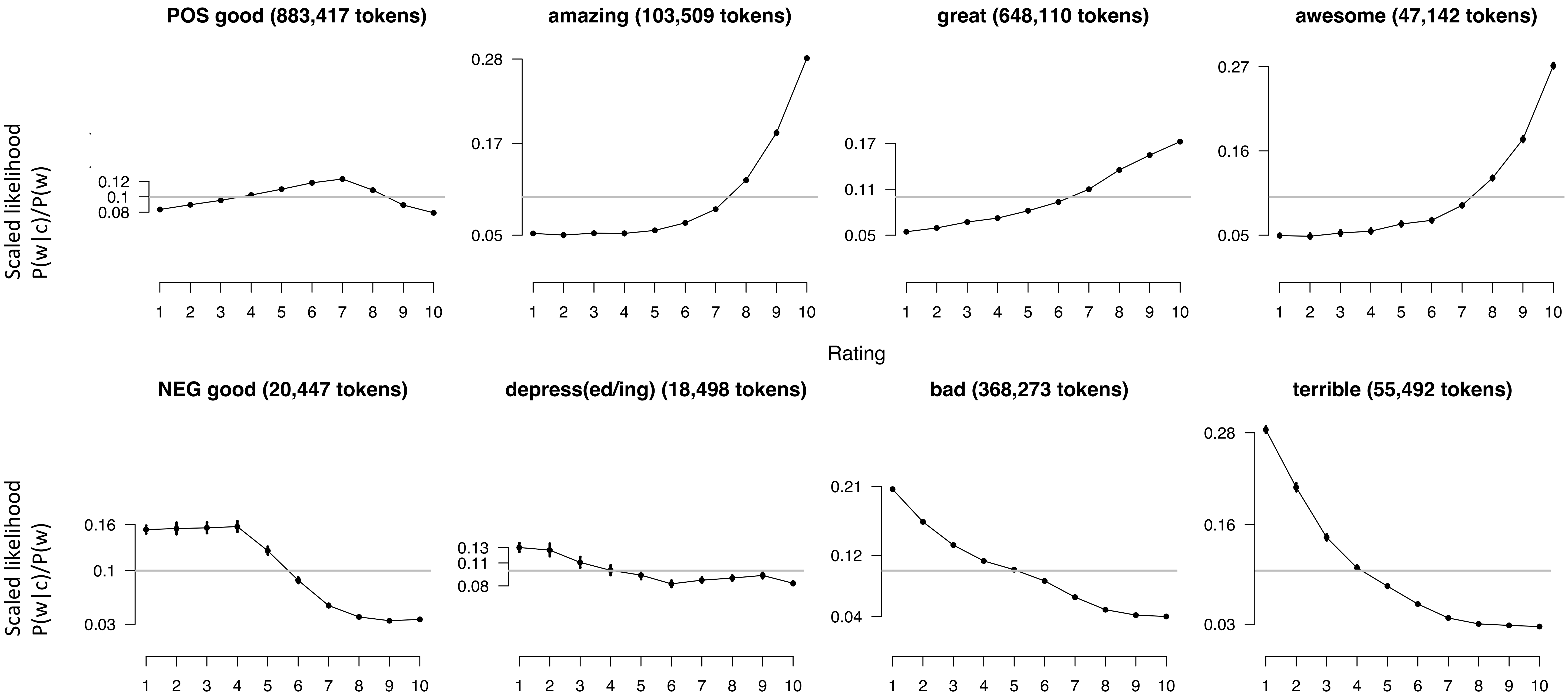
- How likely is each word to appear in each sentiment class?
- Count("bad") in 1-star, 2-star, 3-star, etc.
- But can't use raw counts:
- Instead, **likelihood:** $$P(w|c) = \frac{f(w,c)}{\sum_{w \in c} f(w,c)}$$
- Make them comparable between words
  - **Scaled likelihood:** $$\frac{P(w|c)}{P(w)}$$



Dan Jurafsky

Analyzing the polarity of each word in IMDB

Potts, Christopher. 2011. On the negativity of negation. SALT 20, 636-659.

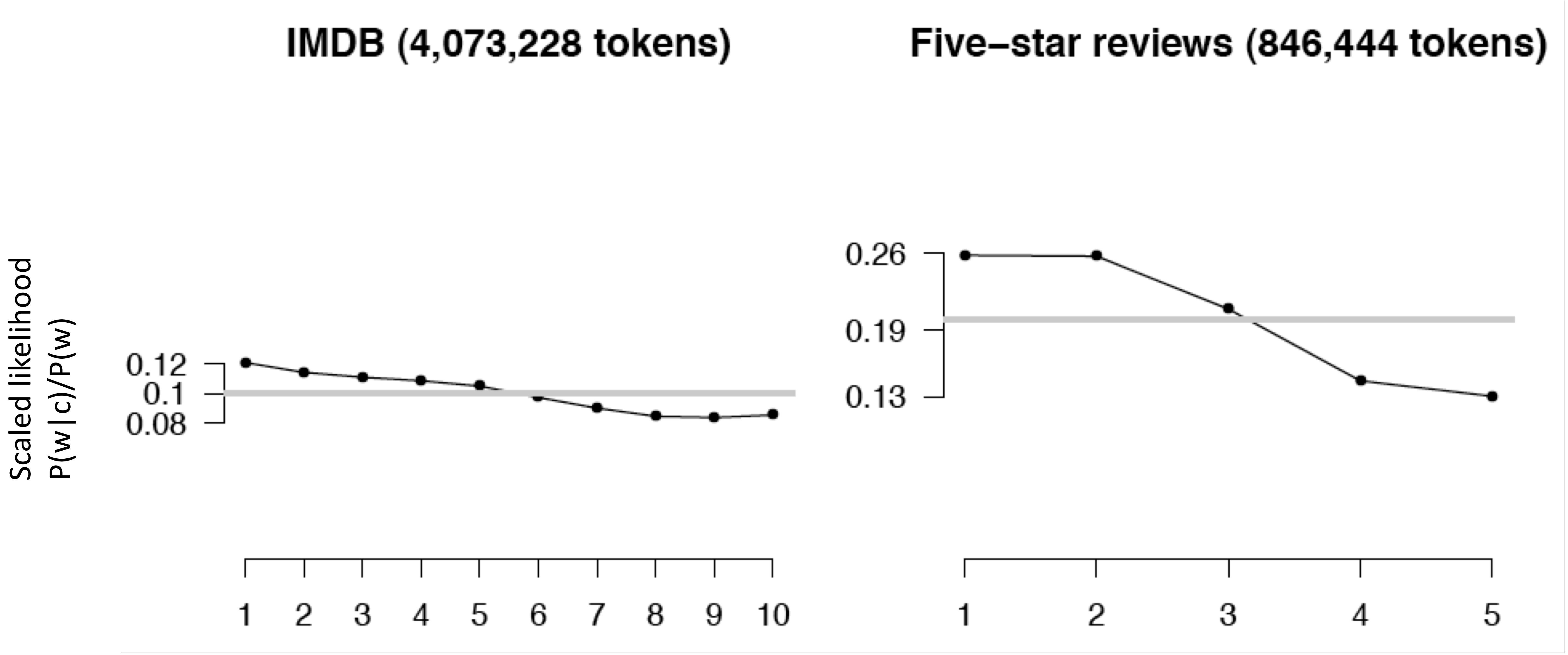


Other sentiment feature: Logical negation

Potts, Christopher. 2011. On the negativity of negation. SALT 20, 636-659.

- 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



Sentiment Analysis

Sentiment Lexicons

Sentiment Analysis

Other Sentiment Tasks

Finding sentiment of a sentence

- Important for finding aspects or attributes
  - Target of sentiment

- `The food was great but the service was awful`

Finding aspect/attribute/target of sentiment

M. Hu and B. Liu. 2004. Mining and summarizing customer reviews. In Proceedings of KDD.
S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop.

- 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

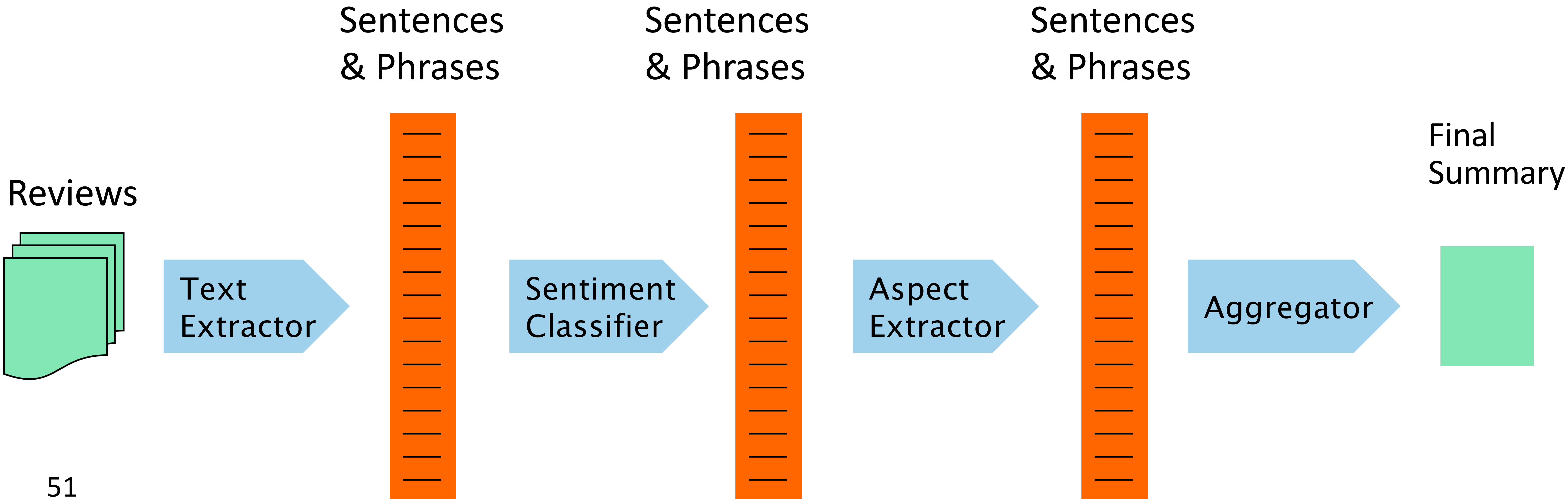
Casino	casino, buffet, pool, resort, beds
Children's Barber	haircut, job, experience, kids
Greek Restaurant	food, wine, service, appetizer, lamb
Department Store	selection, department, sales, shop, clothing

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 asentence
    - “Given this sentence, is the aspect *food, décor, service, value,* or *NONE*”

Putting it all together: Finding sentiment for aspects

S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop



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

Baseline methods assume classes have equal frequencies!

- If not balanced (common in the real world)
  - can't use accuracies as an evaluation
  - need to use F-scores
- Severe imbalancing also can degrade classifier performance
- Two common solutions:
  1. Resampling in training
     - Random undersampling
  2. Cost-sensitive learning
     - Penalize SVM more for misclassification of the rare thing

How to deal with 7 stars?

Bo Pang and Lillian Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. ACL, 115–124

1. Map to binary
2. Use linear or ordinal regression
    - Or specialized models like metric labeling

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

Scherer Typology of Affective States

- **Emotion**: brief organically synchronized … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood**: diffuse non-caused 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
  - *nervous, anxious, reckless, morose, hostile, jealous*

Computational work on other affective states

- **Emotion**:
  - Detecting annoyed callers to dialogue system
  - Detecting confused/frustrated versus confident students
- **Mood**:
  - Finding traumatized or depressed writers
- **Interpersonal stances**:
  - Detection of flirtation or friendliness in conversations
- **Personality traits**:
  - Detection of extroverts

Detection of Friendliness

Ranganath, Jurafsky, McFarland

- Friendly speakers use collaborative conversational style
  - Laughter
  - Less use of negative emotional words
  - More sympathy
    - `That's too bad     I'm sorry to hear that`
  - More agreement
    - `I think so too`
  - Less hedges
    - `kind of    sort of    a little …`

Sentiment Analysis

Other Sentiment Tasks