Introduction to sentiment analysis

Christopher Potts

CS 244U: Natural language understanding
Feb 26 & 28

Overview

1. Applications, data, and resources
2. Sentiment lexicons (off-the-shelf and custom)
3. Basic feature extraction (tokenization, stemming, POS-tagging)
4. Sentiment and syntax (dependencies and sentiment rich phrases)
5. Probabilistic classifier models (with and without classification)
6. Sentiment and context (discourse, identity, perspective)
7. Sentiment as social

Goals

- Sharper conceptualization of the problem
- Practical advice for sentiment analysis
- General lessons about feature extraction and modeling
- General computational methods for extracting meaning
- A glimpse at the next frontier for sentiment: sentiment expression as a social phenomenon

Applications

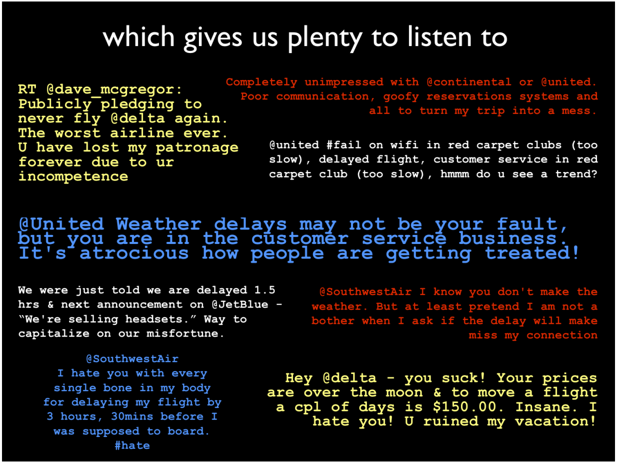


Figure: Understanding customer feedback. From Jeffrey Breen's 'R by example: mining Twitter for attitudes towards airlines': http://jeffreybreen.wordpress.com/2011/07/04/twitter-text-mining-r-slides/

Applications


10 of 120 people found the following review helpful:

I'll buy this book ..., March 15, 2010

By T Boyer "seattleparent" (Seattle) - See all my reviews

This review is from: The Big Short: Inside the Doomsday Machine (Hardcover)

the moment there is a 9.99 Kindle edition. I'll give it a four star rating just so I'm not drawn and quartered by the mob. (Though if you're buying a book based on average stars, without reading the reviews, well how much of a reader are you really?) I'm a big Michael Lewis fan, and I'm sorry his publisher is more interested in winning a pricing war with Amazon than with making the book available to E-book readers.

Help other customers find the most helpful reviews | Report abuse | Permalink

Was this review helpful to you? Yes No | Comments (14)

19 of 394 people found the following review helpful:

Kindle Users get The Big Short !!, March 15, 2010

By JayRye - See all my reviews

This review is from: The Big Short: Inside the Doomsday Machine (Hardcover)

Yes, we kindle users certainly got "The Big Short" on this title. It's really unfortunate. Kindle users take note, the Publisher is W.W. Norton and this decision to not publish a kindle version highlights that greed is not limited to the banking industry.

Help other customers find the most helpful reviews | Report abuse | Permalink

Was this review helpful to you? Yes No | Comments (14)

Figure: Reviews of Michael Lewis's *The Big Short*. These reviews are not critical of the book, but rather of a decision by the publisher about when to release an electronic edition.

Applications

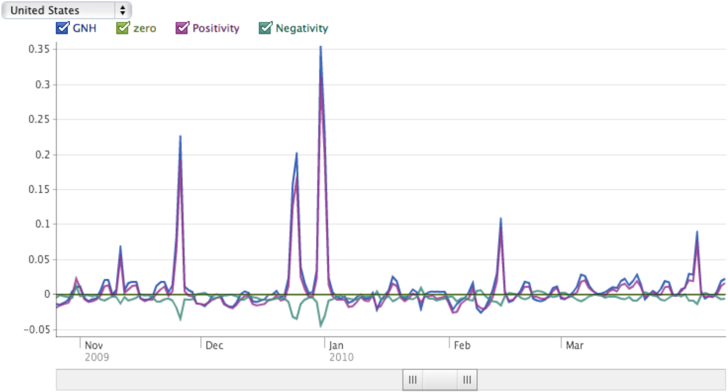


Figure: Facebook's Gross National Happiness interface (defunct?). Holidays register large happiness spikes. The happiness dips in January correspond roughly with the earthquake in Haiti (Jan 12) and its most serious aftershock (Jan 20).

Applications

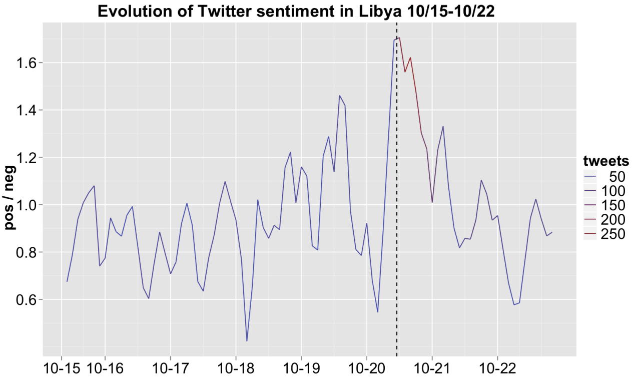


Figure: Twitter sentiment in tweets about Libya, from the project ‘Modeling Discourse and Social Dynamics in Authoritarian Regimes’. The vertical line marks the timing of the announcement that Gaddafi had been killed.

Applications

The media, the President, and the horse race:

> BROOKE GLADSTONE: How do you measure positive and negative press, ’cause you’re talkin’ about news coverage as much as editorial and opinion.
>
> MARK JURKOWITZ: Yes we are, and this is kind of a new research tool for us. It was a computer algorithm developed by a company called Crimson Hexagon. And we actually used our own human researchers and coders to **train the computer basically to look for positive, negative and neutral assertions**. Our sample was over 11,000 different media outlets.

http://www.onthemedia.org/2011/oct/21/media-president-and-horse-race/transcript/

Data

- Data from Lillian Lee's group: `http://www.cs.cornell.edu/home/llee/data/`
- Data from Bing Liu: `http://www.cs.uic.edu/~liub/`
- Large movie review dataset: `http://ai.stanford.edu/~amaas/data/sentiment/`
- Pranav Anand & co. (`http://people.ucsc.edu/~panand/data.php`)
  - Internet Argument Corpus
  - Annotated political TV ads
  - Focus of negation corpus
  - Persuasion corpus (blogs)
- Data on AFS:
  - `/afs/ir/data/linguistic-data/mnt/mnt4/PottsCorpora`
    `README.txt, Twitter.tgz, imdb-english-combined.tgz, opentable-english-processed.zip`
  - `/afs/ir/data/linguistic-data/mnt/mnt9/PottsCorpora`
    `opposingviews, product-reviews, weblogs`
- Twitter data collected and organized by Moritz!
  `/afs/ir.stanford.edu/data/linguistic-data/mnt/mnt3/TwitterTopics/`

Resources

- Basic sentiment tokenizer and some tools: http://sentiment.christopherpotts.net/
- Twitter NLP and Part-of-Speech Tagging: http://www.ark.cs.cmu.edu/TweetNLP/
- Bing Liu's tutorial: http://www.cs.uic.edu/~liub/FBS/Sentiment-Analysis-tutorial-AAAI-2011.pdf
- My tutorial: http://sentiment.christopherpotts.net/
- My course with Dan Jurafsky: http://www.stanford.edu/class/linguist287/
- PDF and BibTEX database for Pang and Lee 2008: http://www.cs.cornell.edu/home/llee/opinion-mining-sentiment-analysis-survey.html

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.
7. Kim bought that damn bike.

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.
7. Kim bought that damn bike.
8. Oh, you're terrible!

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.
7. Kim bought that damn bike.
8. Oh, you’re terrible!
9. Here’s to ya, ya bastard!

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in Arizona.
2. The team failed to complete the physical challenge. (We win/lose!)
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.
7. Kim bought that damn bike.
8. Oh, you're terrible!
9. Here's to ya, ya bastard!
10. Of 2001, "Many consider the masterpiece bewildering, boring, slow-moving or annoying, . . . "

Affect and emotion

Type of affective state: brief definition (*examples*)	Intensity	Duration	Syn-chroni-zation	Event focus	Appraisal elicita-tion	Rapid-ity of change	Behav-ioral impact
Emotion: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an external or internal event as being of major significance (*angry, sad, joyful, fearful, ashamed, proud, elated, desperate*)	+ + –+ + +	+	+ + +	+ + +	+ + +	+ + +	+ + +
Mood: diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause (*cheerful, gloomy, irritable, listless, depressed, buoyant*)	+–+ +	++	+	+	+	++	+
Interpersonal stances: affective stance taken toward another person in a specific interaction, colouring the interpersonal exchange in that situation (*distant, cold, warm, supportive, contemptuous*)	+–+ +	+–+ +	+	++	+	+ + +	++
Attitudes: relatively enduring, affectively coloured beliefs, preferences, and predispositions towards objects or persons (*liking, loving, hating, valueing, desiring*)	0–+ +	+ +–+ + +	0	0	+	0–+	+
Personality traits: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person (*nervous, anxious, reckless, morose, hostile, envious, jealous*)	0–+	+ + +	0	0	0	0	+

0: low, +: medium, ++: high, + + +: very high, –: indicates a range.

Figure: Scherer's (1984) typology of affective states provides a broad framework for understanding sentiment. In particular, it helps to reveal that emotions are likely to be just one kind of information that we want our computational systems to identify and characterize.

Sentiment is hard

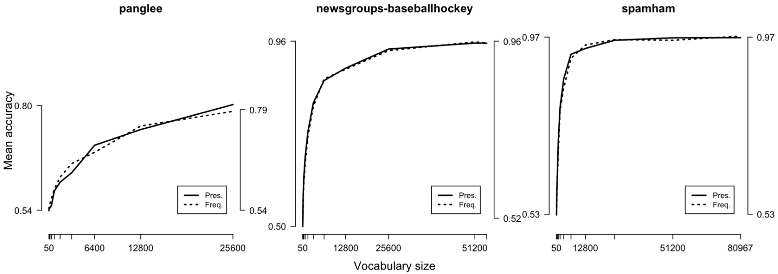


Figure: A single classifier model (MaxEnt) applied to three different domains at various vocabulary sizes. panglee is the widely used movie review corpus distributed by Lillian Lee's group. The 20 newsgroups corpus is a collection of newsgroup discussions on topics like sports, religion, and motorcycles, each with subtopics. spamham is a corpus of spam and ham email messages.

Sentiment lexicons

Understanding and deploying existing sentiment lexicons, or building your own from scratch using unsupervised methods.

Bing Liu's Opinion Lexicon

- http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
- Positive words: 2006
- Negative words: 4783
- Useful properties: includes mis-spellings, morphological variants, slang, and social-media mark-up

MPQA subjectivity lexicon

http://www.cs.pitt.edu/mpqa/

1.	type=weaksubj	len=1	word1=abandoned	pos1=adj	stemmed1=n	priorpolarity=negative
2.	type=weaksubj	len=1	word1=abandonment	pos1=noun	stemmed1=n	priorpolarity=negative
3.	type=weaksubj	len=1	word1=abandon	pos1=verb	stemmed1=y	priorpolarity=negative
4.	type=strongsubj	len=1	word1=abase	pos1=verb	stemmed1=y	priorpolarity=negative
5.	type=strongsubj	len=1	word1=abasement	pos1=anypos	stemmed1=y	priorpolarity=negative
6.	type=strongsubj	len=1	word1=abash	pos1=verb	stemmed1=y	priorpolarity=negative
7.	type=weaksubj	len=1	word1=abate	pos1=verb	stemmed1=y	priorpolarity=negative
8.	type=weaksubj	len=1	word1=abdicate	pos1=verb	stemmed1=y	priorpolarity=negative
9.	type=strongsubj	len=1	word1=aberration	pos1=adj	stemmed1=n	priorpolarity=negative
10.	type=strongsubj	len=1	word1=aberration	pos1=noun	stemmed1=n	priorpolarity=negative
11.	type=strongsubj	len=1	word1=abhor	pos1=anypos	stemmed1=y	priorpolarity=negative
12.	type=strongsubj	len=1	word1=abhor	pos1=verb	stemmed1=y	priorpolarity=negative
13.	type=strongsubj	len=1	word1=abhorred	pos1=adj	stemmed1=n	priorpolarity=negative
14.	type=strongsubj	len=1	word1=abhorrence	pos1=noun	stemmed1=n	priorpolarity=negative
15.	type=strongsubj	len=1	word1=abhorrent	pos1=adj	stemmed1=n	priorpolarity=negative
16.	type=strongsubj	len=1	word1=abhorrently	pos1=anypos	stemmed1=n	priorpolarity=negative
17.	type=strongsubj	len=1	word1=abhors	pos1=adj	stemmed1=n	priorpolarity=negative
18.	type=strongsubj	len=1	word1=abhors	pos1=noun	stemmed1=n	priorpolarity=negative
19.	type=strongsubj	len=1	word1=abidance	pos1=adj	stemmed1=n	priorpolarity=positive
20.	type=strongsubj	len=1	word1=abidance	pos1=noun	stemmed1=n	priorpolarity=positive
⋮						
8221.	type=strongsubj	len=1	word1=zest	pos1=noun	stemmed1=n	priorpolarity=positive

SentiWordNet

POS	ID	PosScore	NegScore	SynsetTerms	Gloss
a	00001740	0.125	0	able#1	(usually followed by 'to') having the necessary means or [...]
a	00002098	0	0.75	unable#1	(usually followed by 'to') not having the necessary means or [...]
a	00002312	0	0	dorsal#2 abaxial#1	facing away from the axis of an organ or organism; [...]
a	00002527	0	0	ventral#2 adaxial#1	nearest to or facing toward the axis of an organ or organism; [...]
a	00002730	0	0	acroscopic#1	facing or on the side toward the apex
a	00002843	0	0	basiscopic#1	facing or on the side toward the base

- Project homepage: `http://sentiwordnet.isti.cnr.it`
- Python/NLTK interface: `http://compprag.christopherpotts.net/wordnet.html`

Harvard General Inquirer

	Entry	Positiv	Negativ	Hostile	... (184 classes)	Othtags	Defined
1	A					DET ART	...
2	ABANDON		Negativ			SUPV	
3	ABANDONMENT		Negativ			Noun	
4	ABATE		Negativ			SUPV	
5	ABATEMENT					Noun	
⋮							
35	ABSENT#1		Negativ			Modif	
36	ABSENT#2					SUPV	
⋮							
11788	ZONE					Noun	

Table: '#*n*' differentiates senses. Binary category values: 'Yes' = category name; 'No' = blank. Heuristic mapping from Othtags into {a,n,r,v}.

- Download: `http://www.wjh.harvard.edu/~inquirer/spreadsheet_guide.htm`
- Documentation: `http://www.wjh.harvard.edu/~inquirer/homecat.htm`

Linguistic Inquiry and Word Counts (LIWC)

Linguistic Inquiry and Word Counts (LIWC) is a propriety database ($90) consisting of a lot of categorized regular expressions.

Category	Examples
Negate	aint, ain't, arent, aren't, cannot, cant, can't, couldnt, . . .
Swear	arse, arsehole*, arses, ass, asses, asshole*, bastard*, . . .
Social	acquainta*, admit, admits, admitted, admitting, adult, adults, advice, advis*
Affect	abandon*, abuse*, abusi*, accept, accepta*, accepted, accepting, accepts, ache*
Posemo	accept, accepta*, accepted, accepting, accepts, active*, admir*, ador*, advantag*
Negemo	abandon*, abuse*, abusi*, ache*, aching, advers*, afraid, aggravat*, aggress*,
Anx	afraid, alarm*, anguish*, anxi*, apprehens*, asham*, aversi*, avoid*, awkward*
Anger	jealous*, jerk, jerked, jerks, kill*, liar*, lied, lies, lous*, ludicrous*, lying, mad

Table: A fragment of LIWC.

Relationships

	MPQA	Opinion Lexicon	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%)
Inquirer			—	520/2306 (23%)	1/204 (0.5%)
SentiWordNet				—	174/694 (25%)
LIWC					—

Table: Disagreement levels for the sentiment lexicons.

- Where a lexicon had POS tags, I removed them and selected the most sentiment-rich sense available for the resulting string.
- For SentiWordNet, I counted a word as positive if its positive score was larger than its negative score; negative if its negative score was larger than its positive score; else neutral, which means that words with equal non-0 positive and negative scores are neutral.
- How to handle the disagreements?

Additional sentiment lexicon resources

- Happy/Sad lexicon (Data_Set_S1.txt) from Dodds et al. 2011
- My NASSLLI 2012 summer course:
  http://nasslli2012.christopherpotts.net
- UMass Amherst Multilingual Sentiment Corpora:
  http://semanticsarchive.net/Archive/jQ0ZGZiM/readme.html
- Developing adjective scales from user-supplied textual metadata
  http://www.stanford.edu/~cgpotts/data/wordnetscales/

The semantic orientation method

1. Get your VSM into shape by weighting and/or dimensionality reduction.
2. Define two seed-sets S_1 and S_2 of words (they should be opposing in some way that is appropriate for your matrix).
3. For a given distance metric *dist* and word *w*:

$$\left(\sum_{w' \in S_1} dist(w, w')\right) - \left(\sum_{w' \in S_2} dist(w, w')\right)$$

The semantic orientation method

1. Get your VSM into shape by weighting and/or dimensionality reduction.
2. Define two seed-sets S_1 and S_2 of words (they should be opposing in some way that is appropriate for your matrix).
3. For a given distance metric *dist* and word w:

$$\left(\sum_{w' \in S_1} \mathit{dist}(w, w')\right) - \left(\sum_{w' \in S_2} \mathit{dist}(w, w')\right)$$

Turney and Littman's (2003:343) hypothesis

The ideas in SO-A can likely be extended to many other semantic aspects of words. The General Inquirer lexicon has 182 categories of word tags [Stone et al. 1966] and this paper has only used two of them, so there is no shortage of future work.

Using our earlier VSM code

```
# Load the source code:
1 source('vsm.R')
# Load the word x word count matrix:
2 imdb = Csv2Matrix('imdb-wordword.csv')
# Build VSM; p. 31/36 of VSM slides for more weighting, LSA, etc:
3 imdb = PMI(imdb, positive=TRUE, discounting=TRUE)
# Define seed-sets; these are from Turney and Littman:
4 neg = c('bad', 'nasty', 'poor', 'negative', 'unfortunate', 'wrong', 'inferior')
5 pos = c('good', 'nice', 'excellent', 'positive', 'fortunate', 'correct', 'superior')
# What is the orientation of 'great'?
6 SemanticOrientation(imdb, word='great', seeds1=neg, seeds2=pos,
  distfunc=CosineDistance)
# For the whole vocabulary:
7 scores = SemanticOrientation(imdb, seeds1=neg, seeds2=pos,
  distfunc=CosineDistance)
# Most like neg:
8 head(scores)
# Most like pos:
9 tail(scores)
```

Pos/neg semantic orientation results (top and bottom 15)

Neighbor	Score
bad	−1.22
worst	−1.13
awful	−1.10
waste	−1.02
terrible	−1.02
worse	−1.00
horrible	−0.95
crap	−0.95
wrong	−0.95
stupid	−0.93
avoid	−0.90
pointless	−0.89
even	−0.89
garbage	−0.88
pathetic	−0.88

Neighbor	Score
excellent	1.17
nice	0.93
great	0.89
superior	0.83
well	0.76
very	0.74
perfect	0.71
role	0.67
performance	0.67
always	0.66
correct	0.66
good	0.65
fantastic	0.65
job	0.65
superb	0.64

Basic feature extraction

- Tokenizing (why this is important)
- Stemming (why you shouldn't)
- POS-tagging (in the service of other goals)

Tokenizing

Raw text

@NLUers: can't wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Tokenizing

Isolate mark-up, and replace HTML entities.

@NLUers: can't wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Tokenizing

Isolate mark-up, and replace HTML entities.

@NLUers: can’t wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Whitespace tokenizer

@NLUers:
can’t
wait
for
the
Mar
12-14
#project
talks!
YAAAAAAY!!!
>:-D
http://stanford.edu/class/cs224u/.

Tokenizing

Isolate mark-up, and replace HTML entities.

@NLUers: can’t wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Treebank tokenizer

@	!
NLUers	YAAAAAAY
:	!
ca	!
n’t	!
wait	>
for	:
the	-D
Mar	http
12-14	:
#	//stanford.edu/class/cs224u/
project	.
talks	

Tokenizing

Isolate mark-up, and replace HTML entities.

@NLUers: can't wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Elements of a sentiment-aware tokenizer

- Isolates emoticons
- Respects Twitter and other domain-specific markup
- Makes use of the underlying mark-up (e.g., <strong> tags)
- Captures those #$%ing masked curses!
- Preserves capitalization where it seems meaningful
- Regularizes lengthening (e.g., *YAAAAAAY*⇒*YAAAY*)
- Captures significant multiword expressions (e.g., *out of this world*)

For regexs and details:
`http://sentiment.christopherpotts.net/tokenizing.html`

Tokenizing

Isolate mark-up, and replace HTML entities.

@NLUers: can’t wait for the Mar 12-14 #project talks! YAAAAAAY!!! >:-D http://stanford.edu/class/cs224u/.

Sentiment-aware tokenizer

@nluers	!
:	YAAAY
can’t	!
wait	!
for	!
the	>:-D
Mar_12-14	http://stanford.edu/class/cs224u/
#project	.
talks	

How much does sentiment-aware tokenizing help?

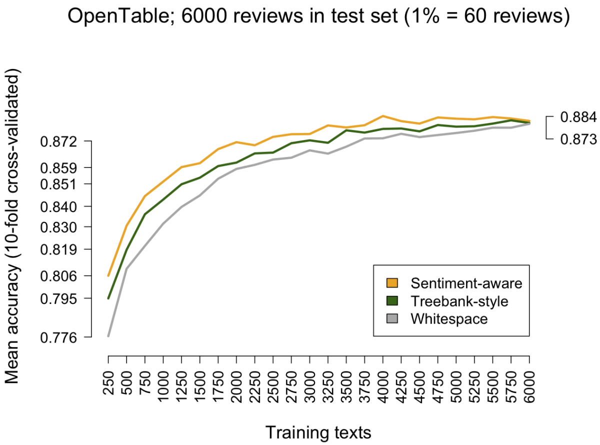


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

How much does sentiment-aware tokenizing help?

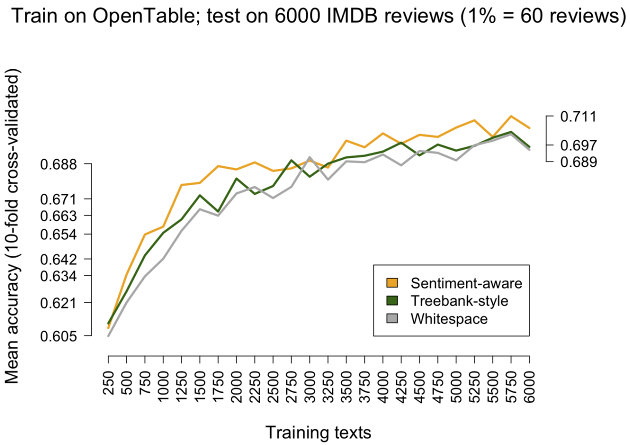


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

Stemming

- Stemming collapses distinct word forms.
- Three common stemming algorithms in the context of sentiment:
  - the Porter stemmer
  - the Lancaster stemmer
  - the WordNet stemmer
- Porter and Lancaster destroy too many sentiment distinctions.
- The WordNet stemmer does not have this problem nearly so severely, but it generally doesn't do enough collapsing to be worth the resources necessary to run it.

Stemming

The Porter stemmer heuristically identifies word suffixes (endings) and strips them off, with some regularization of the endings.

Positiv	Negativ	Porter stemmed
defense	defensive	defens
extravagance	extravagant	extravag
affection	affectation	affect
competence	compete	compet
impetus	impetuous	impetu
objective	objection	object
temperance	temper	temper
tolerant	tolerable	toler

Table: Sample of instances in which the Porter stemmer destroys a Harvard Inquirer Positiv/Negativ distinction.

Stemming

The Lancaster stemmer uses the same strategy as the Porter stemmer.

Positiv	Negativ	Lancaster stemmed
call	callous	cal
compliment	complicate	comply
dependability	dependent	depend
famous	famished	fam
fill	filth	fil
flourish	floor	flo
notoriety	notorious	not
passionate	passe	pass
savings	savage	sav
truth	truant	tru

Table: Sample of instances in which the Lancaster stemmer destroys a Harvard Inquirer Positiv/Negativ distinction.

Stemming

The WordNet stemmer (NLTK) is high-precision. It requires word–POS pairs. Its only general problem for sentiment is that it removes comparative morphology.

Positiv	WordNet stemmed
(exclaims, v)	exclaim
(exclaimed, v)	exclaim
(exclaiming, v)	exclaim
(exclamation, n)	exclamation
(proved, v)	prove
(proven, v)	prove
(proven, a)	proven
(happy, a)	happy
(happier, a)	happy
(happiest, a)	happy

Table: Representative examples of what WordNet stemming does and doesn't do.

How much does stemming help/hurt?

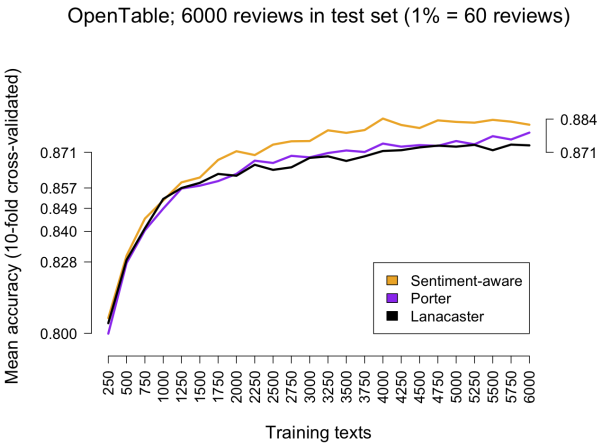


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

Part-of-speech tagging

Word	Tag1	Val1	Tag2	Val2
arrest	jj	Positiv	vb	Negativ
even	jj	Positiv	vb	Negativ
even	rb	Positiv	vb	Negativ
fine	jj	Positiv	nn	Negativ
fine	jj	Positiv	vb	Negativ
fine	nn	Negativ	rb	Positiv
fine	rb	Positiv	vb	Negativ
help	jj	Positiv	vbn	Negativ
help	nn	Positiv	vbn	Negativ
help	vb	Positiv	vbn	Negativ
hit	jj	Negativ	vb	Positiv
mind	nn	Positiv	vb	Negativ
order	jj	Positiv	vb	Negativ
order	nn	Positiv	vb	Negativ
pass	nn	Negativ	vb	Positiv

Table: Harvard Inquirer POS contrasts.

How much does POS tagging help/hurt?

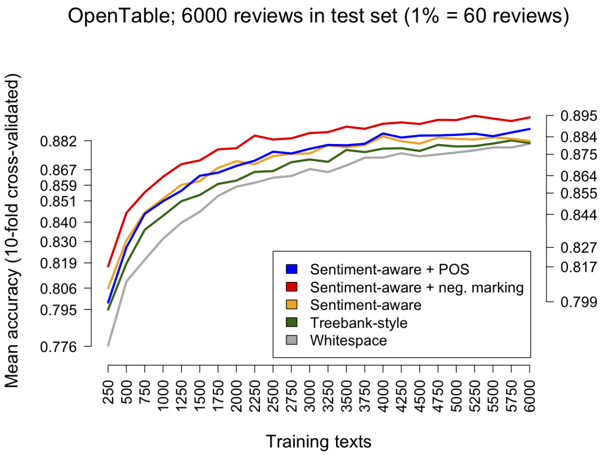
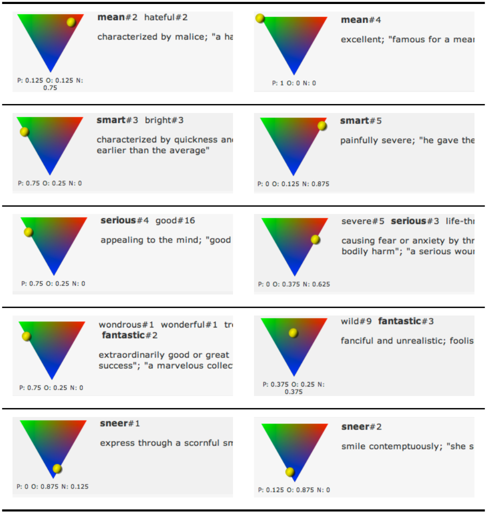


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

SentiWordNet lemma contrasts

1,424 cases where a (word, tag) pair is consistent with pos. and neg. lemma-level sentiment



Word	Tag	ScoreDiff
mean	s	1.75
abject	s	1.625
benign	a	1.625
modest	s	1.625
positive	s	1.625
smart	s	1.625
solid	s	1.625
sweet	s	1.625
artful	a	1.5
clean	s	1.5
evil	n	1.5
firm	s	1.5
gross	s	1.5
iniquity	n	1.5
marvellous	s	1.5
marvelous	s	1.5
plain	s	1.5
rank	s	1.5
serious	s	1.5
sheer	s	1.5
sorry	s	1.5
stunning	s	1.5
wickedness	n	1.5
	[...]	
unexpectedly	r	0.25
velvet	s	0.25
vibration	n	0.25
weather-beaten	s	0.25
well-known	s	0.25
whine	v	0.25
wizard	n	0.25
wonderland	n	0.25
yawn	v	0.25

Sentiment and syntax

1. Marking negation (big gains with a simple technique)
2. Dependency representations for sentiment features
3. Emphasis, attenuation, and negation
4. Sentiment and semantic scope
5. The general problem of sentiment and compositionality

Compositional and non-compositional effects

Sentiment is often, but not always, influenced by the immediate syntactic context.

1. That was fun :)
2. That was miserable :(
3. That was not :)
4. I stubbed my damn toe.
5. What’s with these friggin QR codes?
6. What a view!
7. They said it would be wonderful, but they were wrong: it was awful!
8. This “wonderful” movie turned out to be boring.

Negation

The phenomenon

1. I didn’t enjoy it.
2. I never enjoy it.
3. No one enjoys it.
4. I have yet to enjoy it.
5. I don’t think I will enjoy it.

Negation

The method (Das and Chen 2001; Pang et al. 2002)

- Append a _NEG suffix to every word appearing between a negation and a clause-level punctuation mark.
- For regex details: `http://sentiment.christopherpotts.net/lingstruc.html`

Negation

No one enjoys it.	no one_NEG enjoys_NEG it_NEG .
I don't think I will enjoy it, but I might.	i don't think_NEG i_NEG will_NEG enjoy_NEG it_NEG , but i might .

How much does negation-marking help?

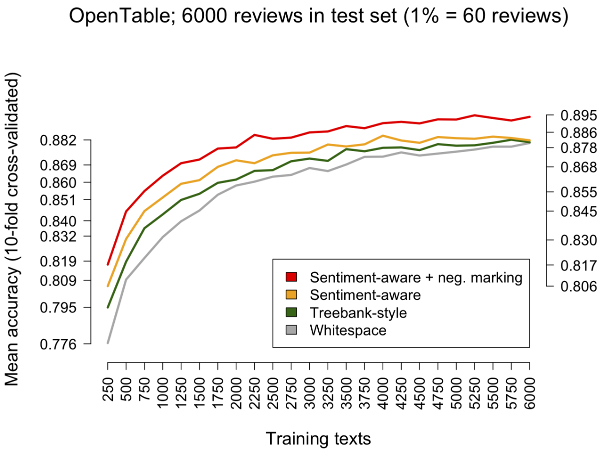


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

How much does negation-marking help?

Train on OpenTable; test on 6000 IMDB reviews (1% = 60 reviews)

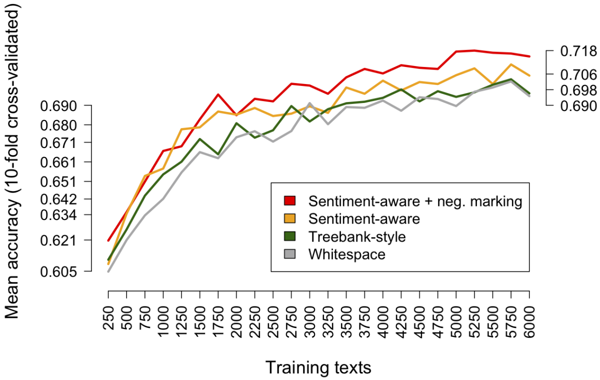


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars). MaxEnt classifier.

Dependency parsing

Treebank-style parsetree

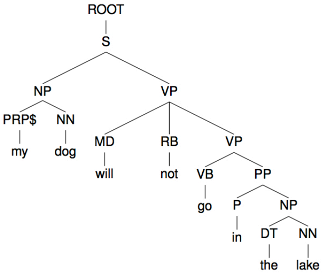
```
(ROOT
  (S
    (NP (PRP$ My) (NN dog))
    (VP (MD will) (RB not)
      (VP (VB go)
        (PP (IN in)
          (NP (DT the) (NN lake)))))
    (. .)))
```

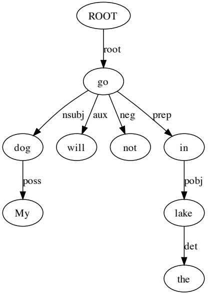
Dependencies

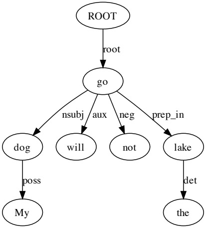
```
poss(dog-2, My-1)
nsubj(go-5, dog-2)
aux(go-5, will-3)
neg(go-5, not-4)
root(ROOT-0, go-5)
prep(go-5, in-6)
det(lake-8, the-7)
pobj(in-6, lake-8)
```

Collapsed dependencies

```
poss(dog-2, My-1)
nsubj(go-5, dog-2)
aux(go-5, will-3)
neg(go-5, not-4)
root(ROOT-0, go-5)
det(lake-8, the-7)
prep_in(go-5, lake-8)
```







A few sentiment-relevant dependencies

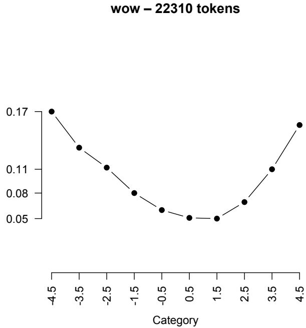
1. amod(student, happy)
2. det(no, student)
3. advmod(amazing , absolutely)
4. aux(VERB, MODAL)
   [MODAL $\in$ {can,could,shall,should,will,would,may,might,must}]
5. nsubj(VERB, NOUN) [subjects generally agents/actors]
6. dobj(VERB, NOUN) [objects generally acted on]
7. ccomp(think, VERB) [clausal complements
8. xcomp(want, VERB) often express attitudes]

Counting and visualizing: IMDB

A Rating	B Count	C Total	D $\Pr(w\|r)$	E $\Pr(r\|w)$
−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

$$\Pr(w|r) \stackrel{def}{=} \text{Count}(w,r)/\text{Total}(r)$$

$$\boxed{\Pr(r|w) \stackrel{def}{=} \frac{\Pr(w|r)}{\sum_{x\in\text{Rating}}\Pr(w|x)}}$$

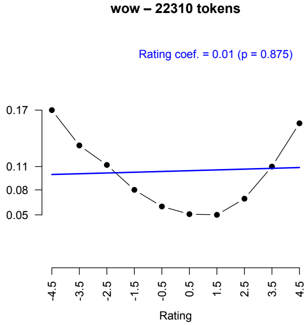


Counting and visualizing: IMDB

A Rating	B Count	C Total	D $\Pr(w\|r)$	E $\Pr(r\|w)$
−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

$$\Pr(w|r) \stackrel{def}{=} \mathrm{Count}(w, r)/\mathrm{Total}(r)$$

$$\boxed{\Pr(r|w) \stackrel{def}{=} \frac{\Pr(w|r)}{\sum_{x \in \mathrm{Rating}} \Pr(w|x)}}$$



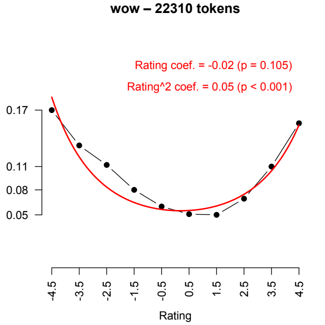
$$\Pr(wow) = \mathrm{logit}^{-1}\left(\begin{array}{l}\texttt{intercept} + \\ \texttt{rating}\end{array}\right)$$

Counting and visualizing: IMDB

A Rating	B Count	C Total	D $\Pr(w\|r)$	E $\Pr(r\|w)$
−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

$$\Pr(w|r) \stackrel{def}{=} \text{Count}(w,r)/\text{Total}(r)$$

$$\Pr(r|w) \stackrel{def}{=} \frac{\Pr(w|r)}{\sum_{x \in \text{Rating}} \Pr(w|x)}$$



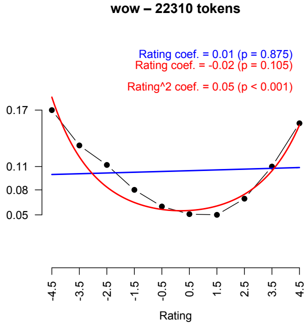
$$\Pr(wow) = \text{logit}^{-1}\begin{pmatrix} \texttt{intercept} & + \\ \texttt{rating} & + \\ \texttt{rating}^2 & \end{pmatrix}$$

Counting and visualizing: IMDB

A Rating	B Count	C Total	D Pr(w|r)	E 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

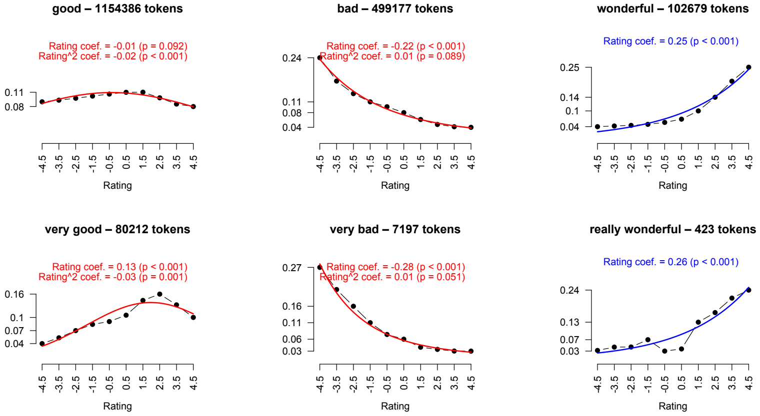
$$\Pr(w|r) \stackrel{def}{=} \text{Count}(w, r)/\text{Total}(r)$$

$$\boxed{\Pr(r|w) \stackrel{def}{=} \frac{\Pr(w|r)}{\sum_{x \in \text{Rating}} \Pr(w|x)}}$$



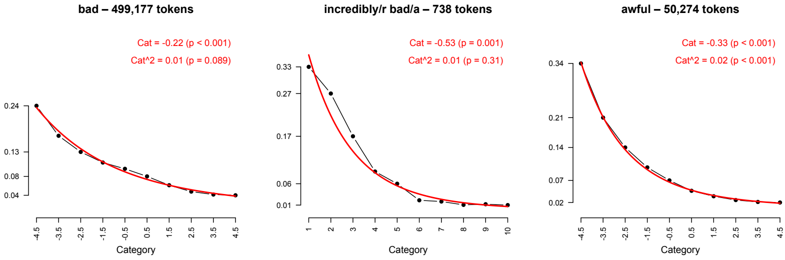
Degree modification (advmod relation)

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



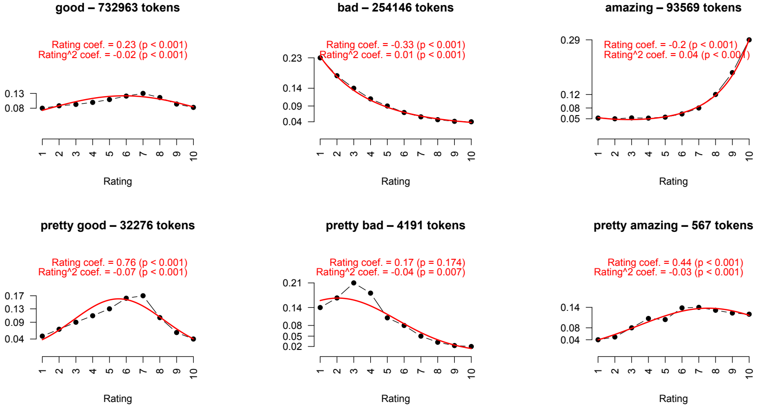
Intensification: the weak overtake the strong

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



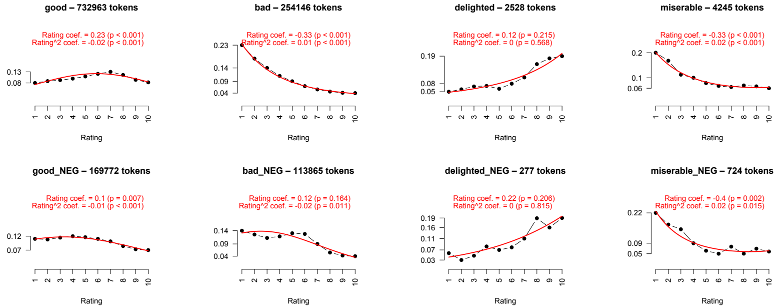
Attenuators (advmod relation)

Adverbials like *pretty* weaken/attenuate sentiment:



Negation (neg relation and the lexicon)

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

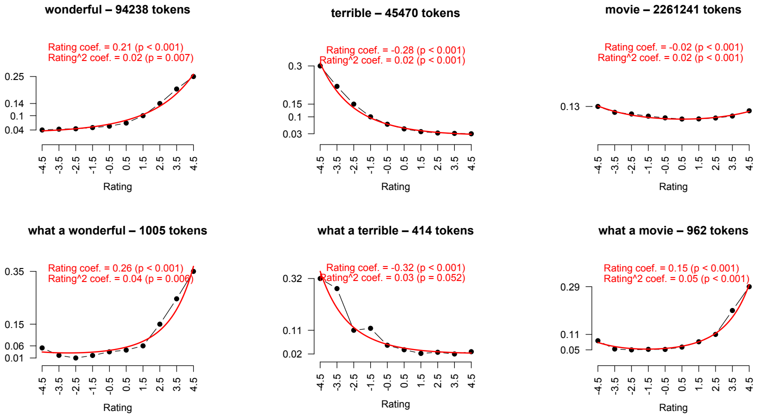


not good $\approx$ *bad*
not bad $\approx$ *good*

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

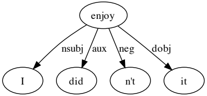
Exclamatives

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

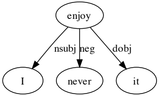


Tracking the influence of negation: semantic scope

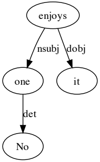
I didn’t enjoy it.



I never enjoy it.

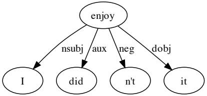


No one enjoys it.

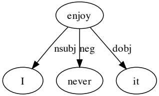


Tracking the influence of negation: semantic scope

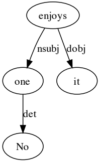
I didn’t enjoy it.



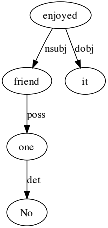
I never enjoy it.



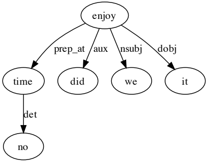
No one enjoys it.



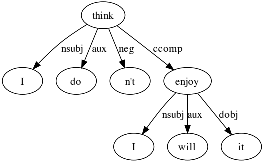
No one’s friend enjoyed it.



At no time did we enjoy it.



I don’t think I will enjoy it.



Seeking a general solution

- On Jan 22, Richard Socher reported to us on his general models of sentiment and compositionality (Socher et al. 2011a,b, 2012)
- While there is a lot of work on semantic compositionality in vector spaces right now (Mitchell and Lapata 2010; Grefenstette et al. 2011), Richard is (as far as I know) the only one directly incorporating labeled data and the only one tackling compositionality in its full generality.
- This is vital because the selective approach described above obviously only begins to address the compositional effects.

Probabilistic classifier models for sentiment

Naive Bayes vs. MaxEnt — who wins? Plus, beyond classification.

Naive Bayes

1. Estimate the probability $P(c)$ of each class $c \in C$ by dividing the number of words in documents in c by the total number of words in the corpus.
2. Estimate the probability distribution $P(w \mid c)$ for all words w and classes c. This can be done by dividing the number of tokens of w in documents in c by the total number of words in c.
3. To score a document d for class c, calculate

$$\textbf{score}(d, c) = P(c) \times \prod_{i=1}^{n} P(w_i \mid c)$$

4. If you simply want to predict the most likely class label, then you can just pick the c with the highest score value.
5. To get a probability distribution, calculate

$$P(c \mid d) = \frac{\textbf{score}(d, c)}{\sum_{c' \in C} \textbf{score}(d, c')}$$

Naive Bayes

- The model predicts a full distribution over classes.
- Where the task is to predict a single label, one chooses the label with the highest probability.
- This means losing a lot of structure. For example, where the max label only narrowly beats the runner-up, we might want to know that.
- The chief drawback to the Naive Bayes model is that it assumes each feature to be independent of all other features.
- For example, if you had a feature *best* and another *world's best*, then their probabilities would be multiplied as though independent, even though the two are overlapping.

MaxEnt

Definition (MaxEnt)

$$P(\mathit{class} \mid \mathit{text}, \lambda) = \frac{\exp\left(\sum_i \lambda_i f_i(\mathit{class}, \mathit{text})\right)}{\sum_{\mathit{class}'} \exp\left(\sum_i \lambda_i f_i(\mathit{class}', \mathit{text})\right)}$$

Minimize:

$$-\sum_{\mathit{class},\mathit{text}} \log P(\mathit{class} \mid \mathit{text}, \lambda) + \log P(\lambda)$$

Gradient:

$$\mathit{empirical\ count}(f_i, c) - \mathit{predicted\ count}(f_i, \lambda)$$

- 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 $\mathit{empirical\ count}(f_i, c)$ is typically also suppressed/ignored.

Ordered categorical regression

Appropriate for data with definitely ordered rating scales (though take care with the scale — it probably isn't conceptually a total ordering for users, but rather more like a pair of scales, positive and negative).

$$\begin{aligned} P(r > 1|\mathbf{x}) &\quad \ldots \\ P(r > 2|\mathbf{x}) &\quad \ldots \\ &\vdots \\ P(r > n-1|\mathbf{x}) &\quad \ldots \end{aligned}$$

Probabilities for the categories:

$$P(r = k|\mathbf{x}) = P(r > k-1) - P(r > k)$$

I don't know whether any classifier packages can build these models, but R users can fit smaller models using polr (from the MASS library). You can also derive them from a series of binary classifiers.

Others

- Support Vector Machines (likely to be competitive with MaxEnt; see Pang et al. 2002).
- Decision Trees (valuable in situations in which you can intuitively define a sequence of interdependent choices, though I've not seen them used for sentiment).
- Generalized Expectation Criteria (a generalization of MaxEnt that facilitates bringing in expert labels; see Druck et al. 2007, 2008).
- Wiebe et al. (2005) use AdaBoost in the context of polarity lexicon construction.

Comparing Naive Bayes and MaxEnt, in domain

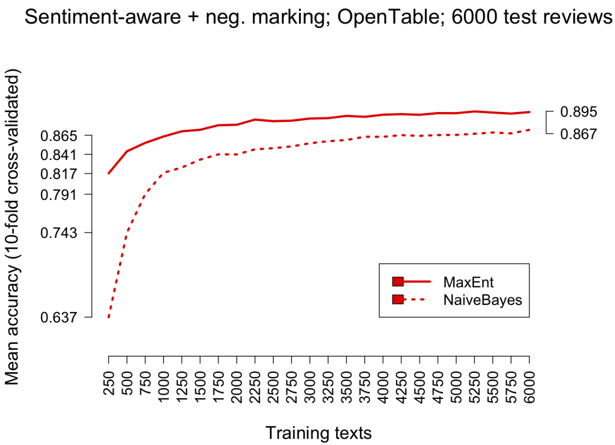


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars).

Comparing Naive Bayes and MaxEnt, in domain

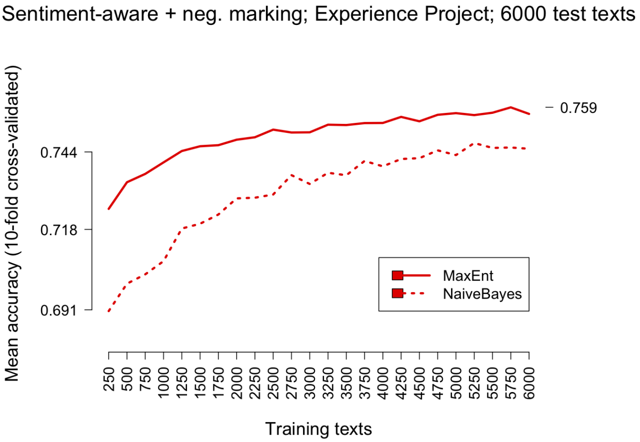


Figure: Training on 15,000 Experience Project texts (5 categories, 3000 in each).

Comparing Naive Bayes and MaxEnt, cross domain

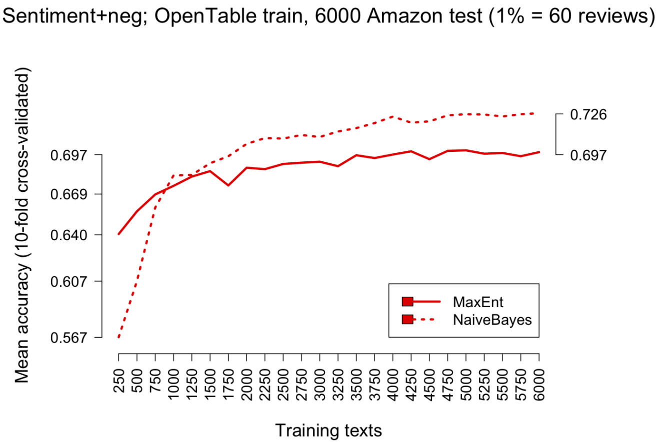


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars).

Comparing Naive Bayes and MaxEnt, cross domain

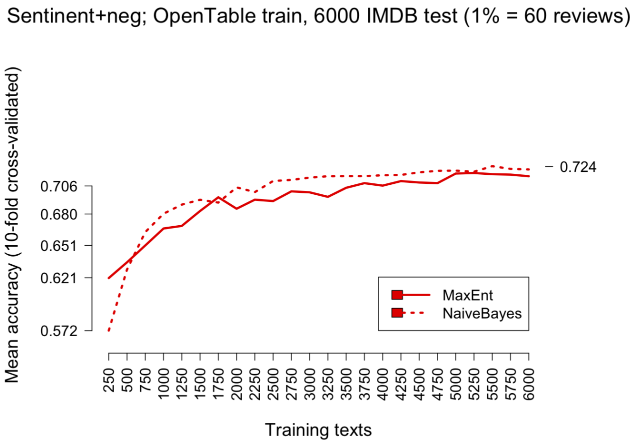


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars).

Overfitting

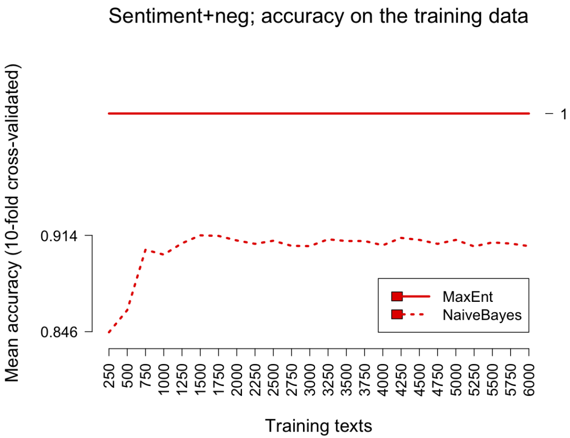


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars).

Feature selection

1. Regularization (strong prior on feature weights).
2. A priori cut-off methods (e.g., top *n* most frequent features; might throw away a lot of valuable information)
3. Select features via mutual information with the class labels (McCallum and Nigam 1998) (liable to make too much of infrequent events!)
4. Sentiment lexicons (potentially unable to detect domain-specific sentiment)

Final comparison

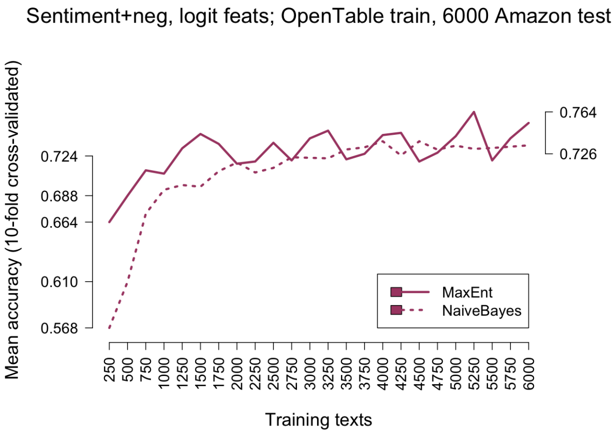


Figure: Training on 12,000 OpenTable reviews (6000 positive/4-5 stars; 6000 negative/1-2 stars).

Beyond classification

This one is for the long-suffering fans, the bittersweet memories, the hilariously embarrassing moments, . . .

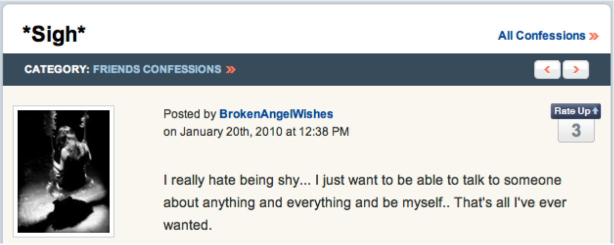
Sentiment as a classification problem

- 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; Wiebe et al. 2005; Pang and Lee 2004; Thomas et al. 2006; Liu et al. 2003; Alm 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.

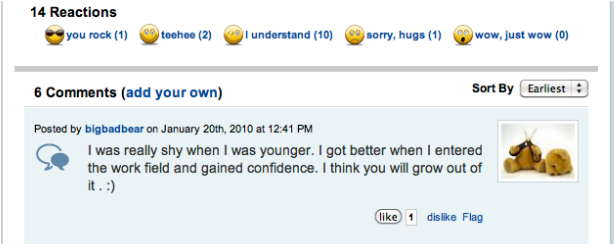
Objections to sentiment as classification

- 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. 2011b), though that will definitely change soon:
    - New datasets emerging
    - Demands from industry
    - New statistical models

Experience Project confessions: blended, continuous sentiment reactions



[. . .]



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

	Texts	Words	Vocab	Mean words/text
Confessions	194,372	21,518,718	143,712	110.71
Comments	405,483	15,109,194	280,768	37.26

Table: The overall size of the corpus.

Reaction distributions

you rock (3) teehee (0) I understand (6) sorry, hugs (1) wow, just wow (0)

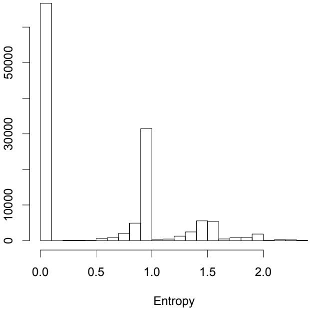
		Category	Reactions
sympathy	←	sorry, hugs	91,222 (22%)
positive exclamative	←	you rock	80,798 (19%)
amused	←	teehee	59,597 (14%)
solidarity	←	I understand	125,026 (30%)
negative exclamative	←	wow, just wow	60,952 (15%)
		Total	417,595

(a) All reactions.

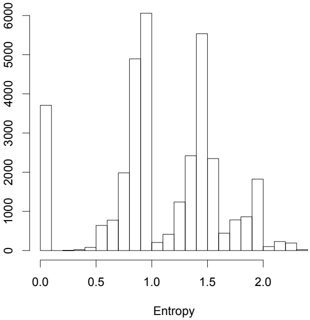
	Texts
$\geqslant 1$	140,467
$\geqslant 2$	92,880
$\geqslant 3$	60,880
$\geqslant 4$	39,342
$\geqslant 5$	25,434

(b) Per text.

Reaction distributions



(a) The full corpus.



(b) $\geqslant 4$ reactions.

Figure: The entropy of the reaction distributions.

A model for sentiment distributions

Definition (MaxEnt with distributional labels)

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

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

$$\sum_{class, text} empiricalProb(class \mid text) \log_2 \left(\frac{empiricalProb(class \mid text)}{P(class|text, \lambda)} \right)$$

Gradient:

$$\sum_{text} empiricalProb(class \mid text) - P(class|text, \lambda)$$

Some results

Features	⩾ 5 reactions KL	⩾ 5 reactions Max Acc.	⩾ 1 reaction KL	⩾ 1 reaction Max Acc.
Uniform Reactions	0.861	20.2	1.275	20.4
Mean Training Reactions	0.763	43.0	1.133	46.7
Bag of Words (All unigrams)	0.637	56.0	1.000	53.4
Bag of Words (Top 5000 unigrams)	0.640	54.9	0.992	54.3
LSA	0.667	51.8	1.032	52.2
Our Method Laplacian Prior	0.621	55.7	0.991	54.7
Our Method Gaussian Prior	0.620	55.2	0.991	54.6

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.

Sentiment and context

A brief look at some of the text-level and contextual features that are important for sentiment:

- Relativization to topics, aspects, and entities
- Isolating the emotional parts of texts
- How perspective and identity influence emotional expression

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)

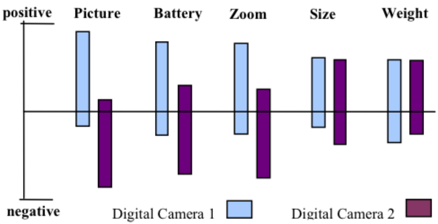
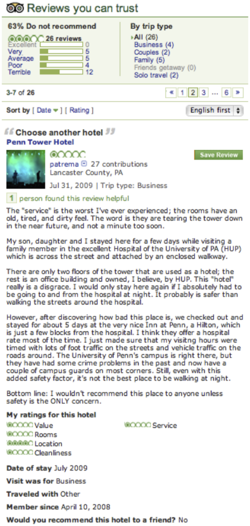


Figure 1: Visual comparison of consumer opinions on two products.

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

Available metadata


Reviews you can trust

63% Do not recommend
26 reviews
Excellent 0
Very 5
Average 5
Poor 4
Terrible 12

By trip type
All (26)
Business (4)
Couples (2)
Family (5)
Friends getaway (0)
Solo travel (2)

3-7 of 26 « 1 2 3 ... 6 »

Sort by [Date] [Rating] English first

"Choose another hotel"
Penn Tower Hotel
Save Review
patrema 27 contributions
Lancaster County, PA
Jul 31, 2009 | Trip type: Business
1 person found this review helpful

The "service" is the worst I've ever experienced; the rooms have an old, tired, and dirty feel. The word is they are tearing the tower down in the near future, and not a minute too soon.

My son, daughter and I stayed here for a few days while visiting a family member in the excellent Hospital of the University of PA (HUP) which is across the street and attached by an enclosed walkway.

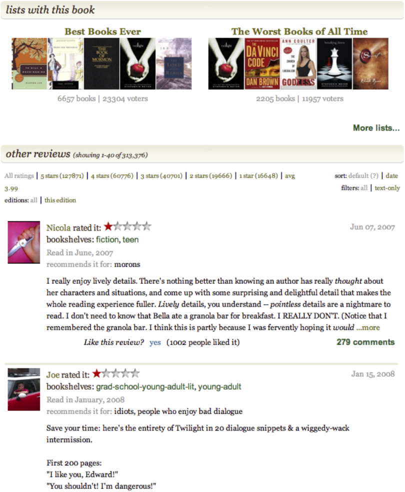
There are only two floors of the tower that are used as a hotel; the rest is an office building and owned, I believe, by HUP. This "hotel" really is a disgrace. I would only stay here again if I absolutely had to be going to and from the hospital at night. It probably is safer than walking the streets around the hospital.

However, after discovering how bad this place is, we checked out and stayed for about 5 days at the very nice Inn at Penn, a Hilton, which is just a few blocks from the hospital. I think they offer a hospital rate most of the time. I just made sure that my visitng hours were timed with lots of foot traffic on the streets and vehicle traffic on the roads around. The University of Penn's campus is right there, but they have had some crime problems in the past and now have a couple of campus guards on most corners. Still, even with this added safety factor, it's not the best place to be walking at night.

Bottom line: I wouldn't recommend this place to anyone unless safety is the ONLY concern.

My ratings for this hotel
Value
Service
Rooms
Location
Cleanliness

Date of stay July 2009
Visit was for Business
Traveled with Other
Member since April 10, 2008
Would you recommend this hotel to a friend? No


lists with this book

Best Books Ever
6657 books | 23304 voters

The Worst Books of All Time
2205 books | 11957 voters

More lists...

other reviews (showing 1-40 of 313,376)

All ratings | 5 stars (127871) | 4 stars (60776) | 3 stars (40701) | 2 stars (19666) | 1 star (16648) | avg 3.99
sort: default (?) | date
filters: all | text-only
editions: all | this edition

Nicola rated it: ★☆☆☆☆ Jun 07, 2007
bookshelves: fiction, teen
Read in June, 2007
recommends it for: morons

I really enjoy lively details. There's nothing better than knowing an author has really *thought* about her characters and situations, and come up with some surprising and delightful detail that makes the whole reading experience fuller. *Lively* details, you understand -- *pointless* details are a nightmare to read. I don't need to know that Bella ate a granola bar for breakfast. I REALLY DON'T. (Notice that I remembered the granola bar. I think this is partly because I was fervently hoping it *would* ...more

Like this review? yes (1002 people liked it) 279 comments

Joe rated it: ★☆☆☆☆ Jan 15, 2008
bookshelves: grad-school-young-adult-lit, young-adult
Read in January, 2008
recommends it for: idiots, people who enjoy bad dialogue

Save your time: here's the entirety of Twilight in 20 dialogue snippets & a wiggedy-wack intermission.

First 200 pages:
"I like you, Edward!"
"You shouldn't! I'm dangerous!"

OpenTable: attribute-level ratings

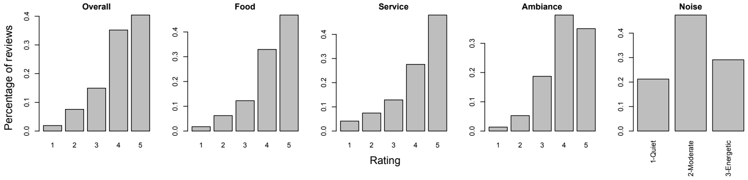
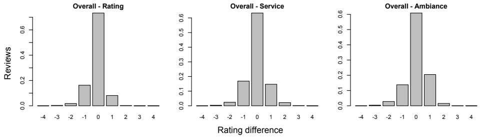


Figure: OpenTable rating distributions. Positive reviews dominate in all categories. ‘Noise’ is different, since it lacks a standard preference ordering.



(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.

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

(b) Correlations.

Figure: OpenTable rating category comparisons.

Narrative structure

38 of 44 people found the following review helpful:

Move over, Robert Jordan., July 19, 1998

By **A Customer**

This review is from: **A Game of Thrones (A Song of Ice and Fire, Book 1) (Mass Market Paperback)**

As a fantasy reader of somewhat high standards, I have always had a proclivity for "epic" fantasy. Nothing else really satisfies my desire for an absorbing story. George R.R. Martin has, with this book, taken the field dominated by such giants as Jordan, Williams, and Kay and blown a great big gust of fresh air into it. Not only does this book have the complicated plot and intricate character development that is common to these three talented authors, but it has a certain brutal realism to it. Granted, we're talking about an invented realm, but never before in all the books that I have read has any author taken his portrayal of all the brutality of human nature to this level. Part of what makes Jordan, Williams, and Kay so brilliant is that they write *human* characters, and good and bad are rarely well delineated. What sets Martin apart is his sheer, brutal, mind-numbing honesty. He doesn't pull any punches, and neither do any of his characters. This ! is life, in all its pain and glory. Honor is not as important as we would like it to be, and things do not all go well as long as we wish for it hard enough. Here, there is no destructive force stronger than the power of men. There is no evil greather than that in the hearts of men. And there is no power, once man has decided to destroy, that can stop him. This novel is a masterpiece; beautifully crafted, shockingly realistic, and a joy to read. However, don't expect to come out of reading this with your ideals intact.

Help other customers find the most helpful reviews

Was this review helpful to you?

Report abuse | Permalink

Comment

(5-star Amazon review)

Narrative structure

41 of 50 people found the following review helpful:

What's left unsaid, February 12, 2004

By **A Customer**

Amazon Verified Purchase (What's this?)

This review is from: **A Game of Thrones (A Song of Ice and Fire, Book 1) (Mass Market Paperback)**

All of the other excellent reviews of this series are correct. The writing is wonderful. The characters are real. The plot is intricate, fascinating, and never predictable. Et cetera. But none of the reviewers complained about the one thing that has led me to stop reading after plugging through the first two books: This is the darkest, bleakest, most depressing book I have ever read! You must never, ever let yourself bond with a hero, a good, kind, strong, resourceful person who in a 'normal' book would win a gratifying victory at the end of the book. This is because chances are your hero will soon die, most likely brutally. Most (eventually all???) of the good guys die in this book! And everyone is always having to look over his shoulder to see which one of his supposed friends is plotting his death. Innocent children are brutally murdered and their heads put up on pikes. Innocent peasants are slowly hanged, kicking, their eyes bulging out. Their rescuers, instead of pulling off a valiant rescue, are themselves captured and tortured. There are innumerable rapes, including several fairly explicit portrayals of vicious gang rapes of peasant women by invading troops. Every time I finished a reading session I felt depressed. I've never seen so much plague, betrayal, death, and destruction in a novel. It's unrelenting. I don't care how wonderful the writing is. I simply couldn't take it anymore. I want to be uplifted by a book, made to smile and feel vicariously triumphant. I don't want to be beaten down and defeated over and over and over. I had to stop reading.

Help other customers find the most helpful reviews | Report abuse | Permalink

Was this review helpful to you? | Comments (2)

(3-star Amazon review)

Narrative structure

Algorithms for text-segmentation

- The TextTiling algorithm (Hearst 1994, 1997)
- Dotplotting (Reynar 1994, 1998)
- Divisive clustering (Choi 2000)
- Supervised approaches (Manning 1998; Beeferman et al. 1999; Sharp and Chibelushi 2008)

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 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 dissapointment , 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: A negative review. Inquirer positive terms in blue, negative in red. There are 20 positive terms and six negative ones, for a Pos:Neg ratio of 3.33.

Thwarted expectations

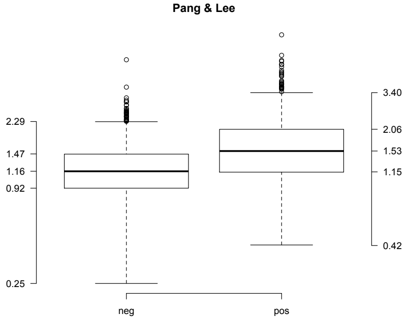


Figure: Inquirer Pos:Neg ratios obtained by counting the terms in the review that are classified as Positiv or Negativ in the Harvard Inquirer (Stone et al. 1966).

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.)

Sentiment, perpective, and identity

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, perpective, and identity

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.

Contextual variables

Age	Texts
teens	5,495
20s	26,564
30s	15,317
40s	7,413
50s	3,600
$\geqslant 60$	1130
unknown	80,948
Total	140,467

(a) Author ages.

Gender	Texts
female	34,921
male	15,333
unknown	90,213
Total	140,467

(b) Author genders.

Group	Texts
crime	312
embarrassing	5,349
family	5,114
friends	13,719
funny	3,692
health	6,467
love	36,242
revenge	1,406
school	1,698
sex	45,538
venting	19,090
work	1,840
Total	140,467

(c) 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

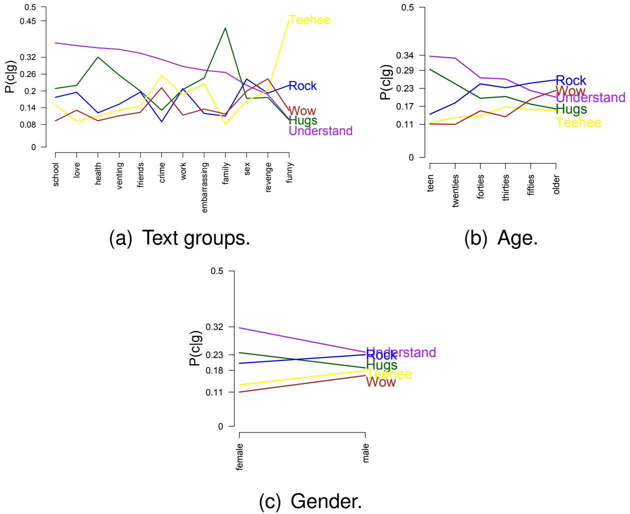
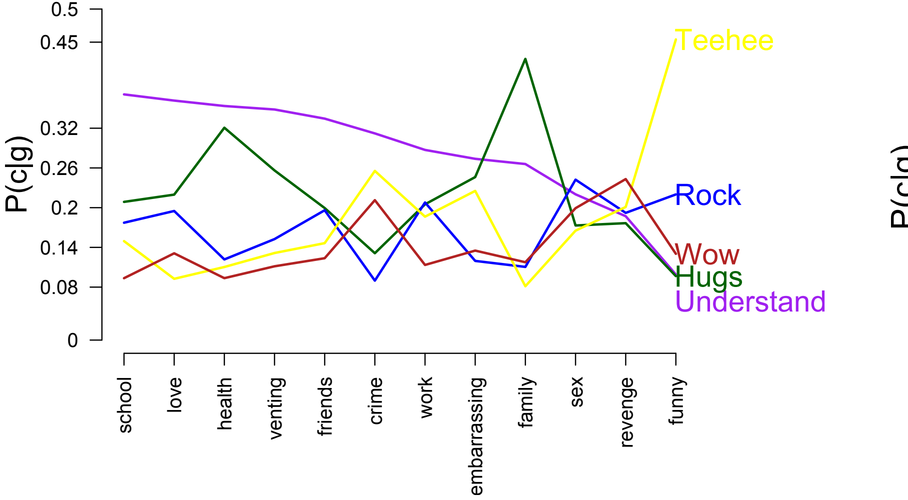


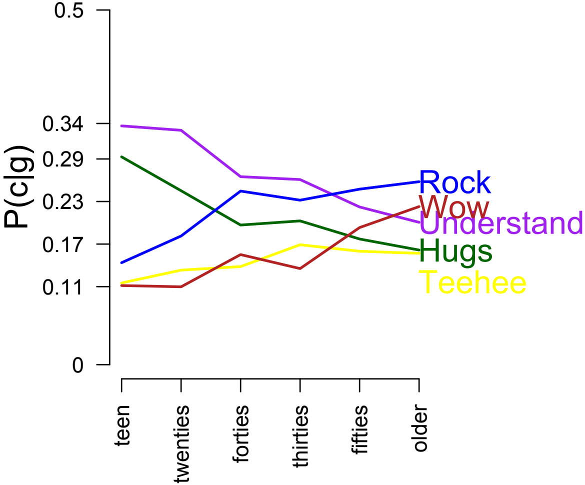
Figure: Text groups show the most variability. Age and gender are more stable by comparison, though the relationships remain interesting.

The influences of context



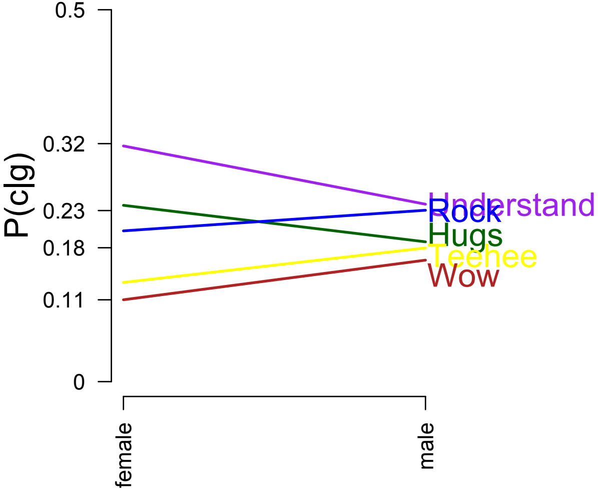
(a) Text groups

The influences of context



(b) Age

The influences of context



(c) Gender

The influences of text groups

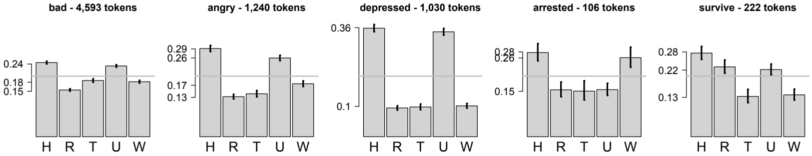


Figure: Words eliciting predominantly ‘You rock’ reactions. The data reveal other dimensions as well, including mixes of light-heartedness, negative exclamativity.

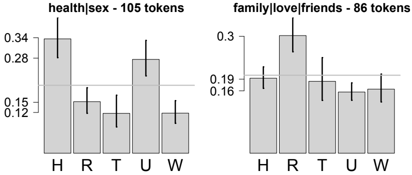


Figure: The bimodal distribution of *survive* seems to derive from an underlying distinction in text group.

The influences of age

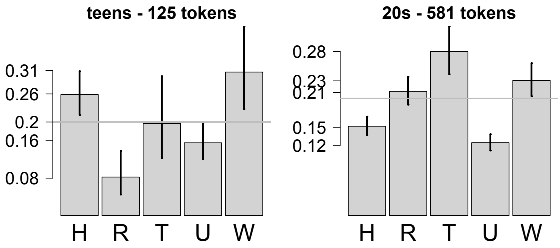


Figure: Age is a source of variation in responses to *drunk*.

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 $P(\textit{class} \mid \textit{text}, \textit{context}, \lambda)$.
- See also Overgoor 2012.

Sentiment as social

- Are your attitudes predicted by the attitudes of your friends?
- How is your emotional expression affected by who you are talking to, what you are talking about, and other facts about the conversational context?

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

Bill	052
Speaker	400011
Party	Democrat
Vote	No
Sample	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 ?

Bill	052
Speaker	400077
Party	Republican
Vote	Yes
Sample	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 .

The Convote corpus

	total	train	test	development
speech segments	3857	2740	860	257
debates	53	38	10	5
average number of speech segments per debate	72.8	72.1	86.0	51.4
average number of speakers per debate	32.1	30.9	41.1	22.6

Table 1: Corpus statistics.

Hierarchy of texts:

Debates (collections of speeches by different speakers)

$\Uparrow$

Speeches (collections of segments by the same speaker)

$\Uparrow$

Speech segments (documents in the corpus)

Basic classification with same-speech links

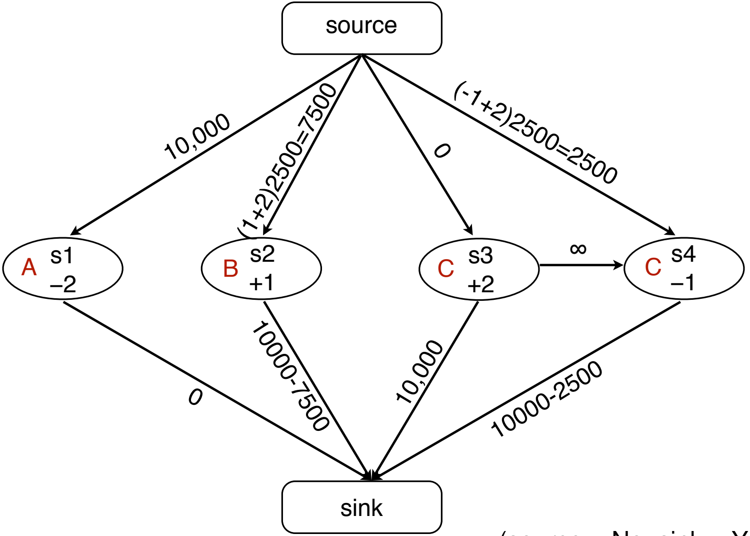
1. SVM classifier with unigram-presence features predicting, for each speech-segment, how the speaker voted (Y or N).
2. For each document s belonging to speech S, the SVM score for s is divided by the standard deviation for all $s' \in S$.
3. Debate-graph construction with minimal cuts:

$$\text{score}(s) \leqslant -2 \Rightarrow \left[\begin{array}{rcl} \text{source} & \overset{0}{\rightarrow} & s \\ s & \overset{10{,}000}{\rightarrow} & \text{sink} \end{array}\right.$$

$$\text{score}(s) \geqslant +2 \Rightarrow \left[\begin{array}{rcl} \text{source} & \overset{10{,}000}{\rightarrow} & s \\ s & \overset{0}{\rightarrow} & \text{sink} \end{array}\right.$$

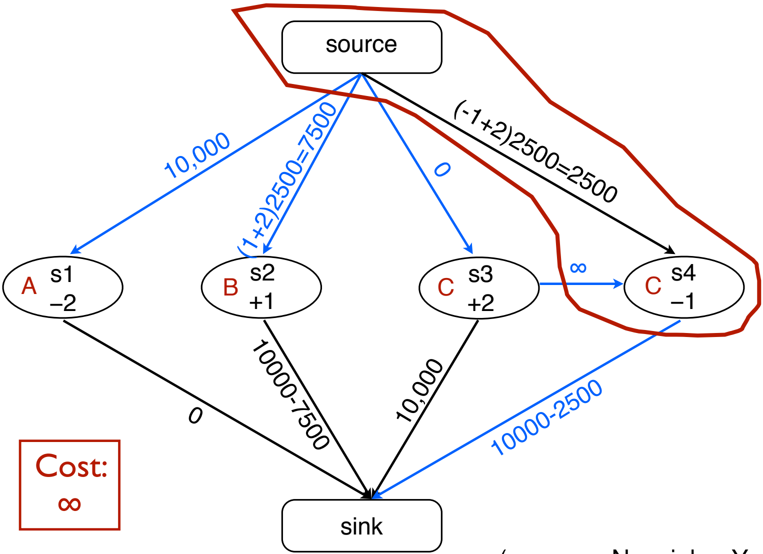
$$\text{else} \Rightarrow \left[\begin{array}{rcl} \text{source} & \overset{x=(\text{score}(s)+2)2500}{\rightarrow} & s \\ s & \overset{10{,}000-x}{\rightarrow} & \text{sink} \end{array}\right.$$

Graph construction and minimal cuts



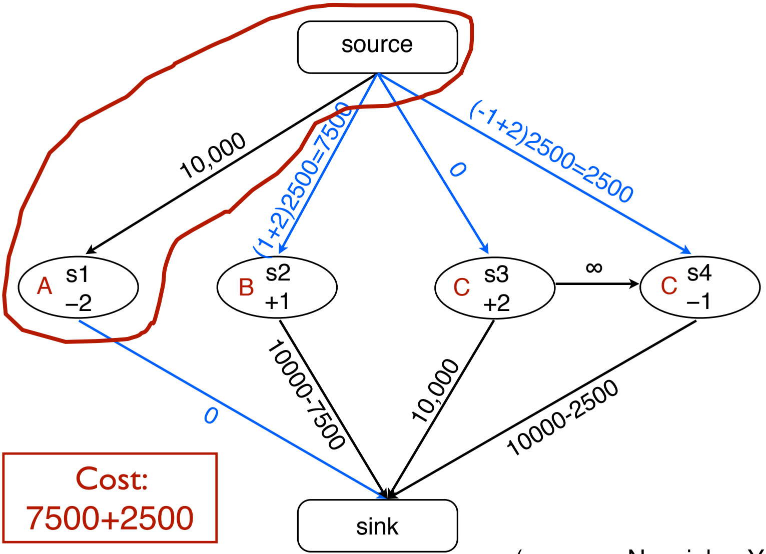
(source = No; sink = Yes)

Graph construction and minimal cuts



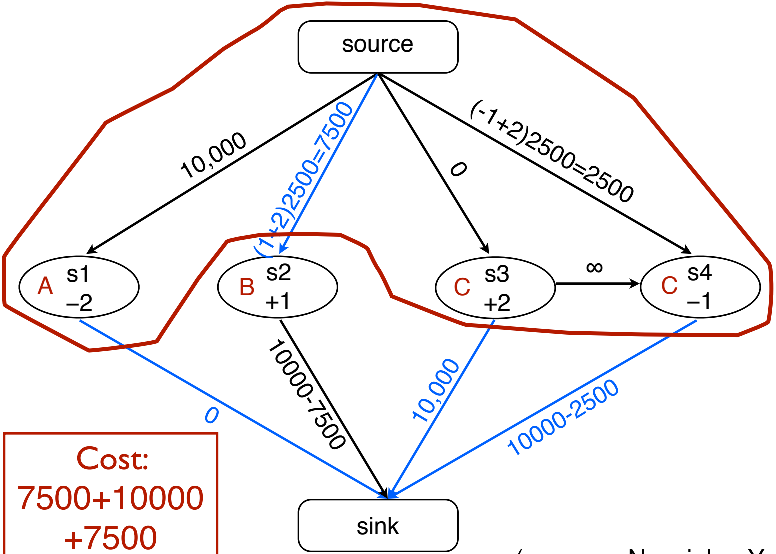
(source = No; sink = Yes)

Graph construction and minimal cuts



(source = No; sink = Yes)

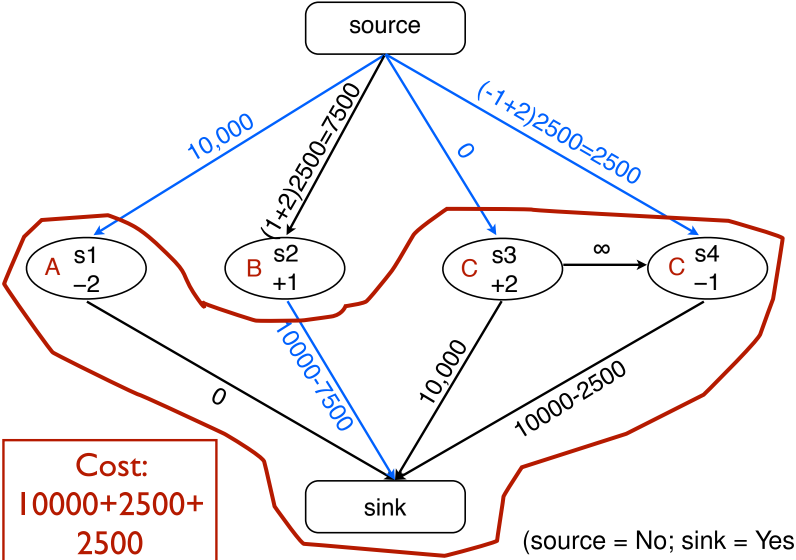
Graph construction and minimal cuts



Cost:
7500+10000
+7500

(source = No; sink = Yes)

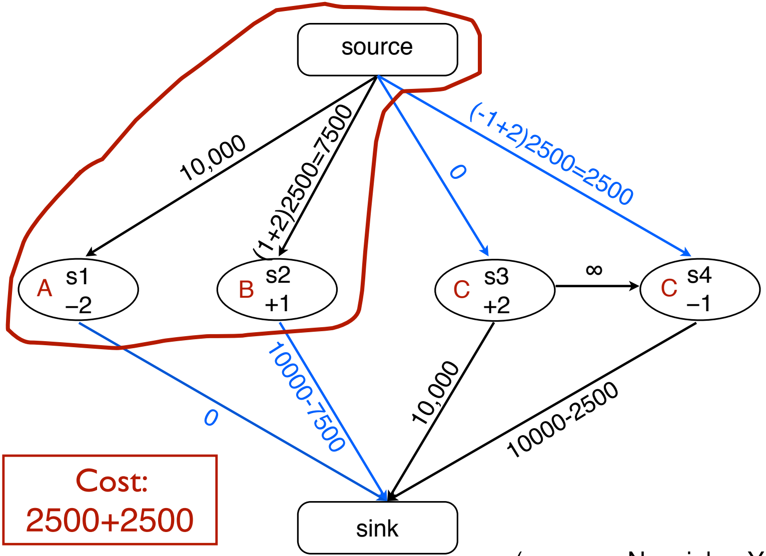
Graph construction and minimal cuts



Cost:
10000+2500+
2500

(source = No; sink = Yes)

Graph construction and minimal cuts



(source = No; sink = Yes)

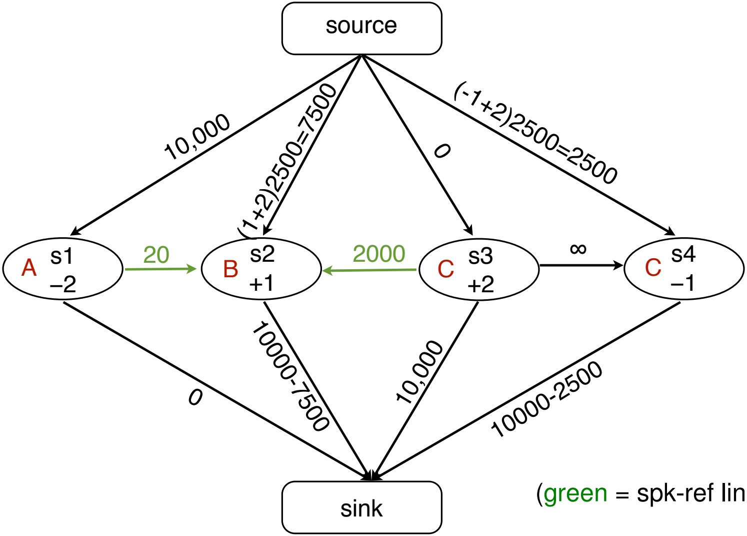
Speaker references

Bill	006
Speaker	400115
Party	Republican
Vote	Yes
Sample	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 .
Bill	006
Speaker	400035
Party	Republican
Vote	Yes
Sample	mr. speaker , i rise in strong support of this balanced rules package . i want to speak particularly to the provisions regarding homeland security . [. . .]

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



(green = spk-ref links)

Results

Support/oppose classifer ("speech segment⇒yea?")	Devel. set	Test set
majority baseline	54.09	58.37
#("support") − #("oppos")	59.14	62.67
SVM [speech segment]	70.04	66.05
SVM + same-speaker links	79.77	67.21
SVM + same-speaker links . . .		
+ agreement links, $\theta_{agr} = 0$	**89.11**	**70.81**
+ agreement links, $\theta_{agr} = \mu$	87.94	71.16

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

θ_{agr} is a free-parameter in the scaling function for speaker agreement scores. The development results suggest that 0 is the better value than μ (a mean of all the debate's scores), but μ 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.

Topic	# users	#t-follow edges		#@ edges		# on-topic tweets
		dir.	mutual	dir.	mutual	
Obama	889	7,838	2,949	2,358	302	128,373
Sarah Palin	310	1,003	264	449	60	21,571
Glenn Beck	313	486	159	148	17	12,842
Lakers	640	2,297	353	1,167	127	35.250
Fox News	231	130	32	37	5	8,479

- 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.

Connected user tend to share topic-relative sentiment (Tan et al. 2011)

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.

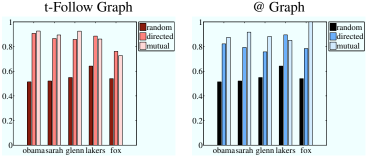


Figure 1: Shared sentiment conditioned on type of connection. Y-axis: probability of two users v_i and v_j having the same sentiment label, conditioned on relationship type. The left plot is for the t-follow graph, while the right one is for the @ graph. "random": pairs formed by randomly choosing users. "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).

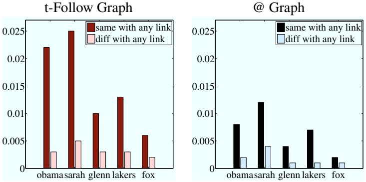
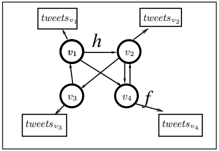


Figure 2: Connectedness conditioned on labels. Y-axis: probability that two users are connected, conditioned on whether or not the users have the same sentiment.

Graphical structure and model (Tan et al. 2011)

Heterogeneous graph



k = user sentiment label $\in \{0,1\}$

l = tweets sentiment label $\in \{0,1\}$

v_i sentiment label $\in \{0,1\}$

unknown label $\in \{0,1\}$ for v_i tweets

$$\log P(\mathbf{Y}) = \Big(\sum_{v_i \in V} \Big[\sum_{t \in tweets_{v_i}, k, \ell} \mu_{k,\ell} f_{k,\ell}(y_{v_i}, \hat{y}_t) + \sum_{v_j \in Neighbors_{v_i}, k, \ell} \lambda_{k,\ell} h_{k,\ell}(y_{v_i}, y_{v_j}) \Big] \Big) - \log Z,$$

vector of user labels for the topic in question

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

= 1 in experiments = 0.125 in experiments

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

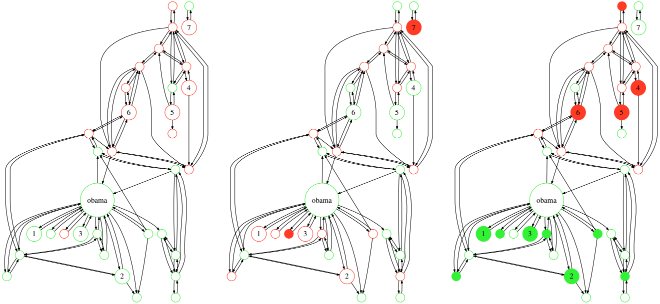
User–tweet factor

= 0.6 in experiments

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

User–user factor

Case study highlighting the value of social information



(a) Ground Truth (b) Text-Only Approach (c) Our algorithm

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

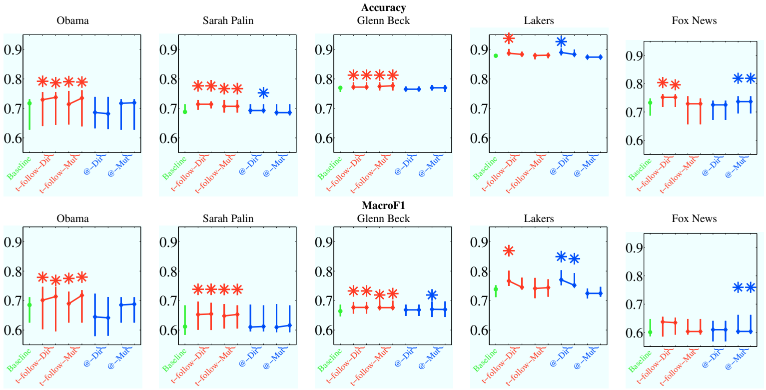


Figure 6: Performance Analysis in Different Topics. **The x-axes are the same as in Figure 5. Bars summarize performance results for our "10-run" experiments: the bottom and top of a bar indicate the 25th and 75th percentiles, respectively. Dots indicate median results; in pairs connected by lines, the 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.**

References I

Alm, Cecilia Ovesdotter; Dan Roth; and Richard Sproat. 2005. Emotions from text: Machine learning for text-based emotion prediction. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*.

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge University Press.

Beeferman, Doug; Adam Berger; and John Lafferty. 1999. Statistical models for text segmentation. *Machine Learning* 34:177–210. doi:\bibinfo{doi}{10.1023/A:1007506220214}. URL http://dl.acm.org/citation.cfm?id=309497.309507.

Bruce, Rebecca F. and Janyce M. Wiebe. 1999. Recognizing subjectivity: A case study in manual tagging. *Natural Language Engineering* 5(2).

Cabral, Luís and Ali Hortaçsu. 2006. The dynamics of seller reputation: Theory and evidence from eBay. Working paper, downloaded version revised in March. URL http://pages.stern.nyu.edu/~lcabral/workingpapers/CabralHortacsu_Mar06.pdf.

Choi, Freddy Y. Y. 2000. Advances in domain independent linear text segmentation. In *1st Meeting of the North American Chapter of the Association for Computational Linguistics*, 26–33. Seattle, WA: Association for Computational Linguistics.

Das, Sanjiv and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In *Proceedings of the 8th Asia Pacific Finance Association Annual Conference*.

References II

Dodds, Peter Sheridan; Kameron Decker Harris; Isabel M. Kloumann; Catherine A. Bliss; and Christopher M. Danforth. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS One* 6(12):1–26.

Druck, Gregory; Gideon Mann; and Andrew McCallum. 2007. Generalized expectation criteria. Technical Report 2007-60, University of Massachusetts Amherst, Amherst, MA.

Druck, Gregory; Gideon Mann; and Andrew McCallum. 2008. Learning from labeled features using generalized expectation criteria. In *Proceedings of ACM Special Interest Group on Information Retreival*.

Ekman, Paul. 1992. An argument for basic emotions. *Cognition and Emotion,* 6(3/4):169–200.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Goldberg, Andrew B. and Jerry Zhu. 2006. Seeing stars when there aren't many stars: Graph-based semi-supervised leaarning for sentiment categorization. In *TextGraphs: HLT/NAACL Workshop on Graph-based Algorithms for Natural Language Processing*.

Grefenstette, Edward; Mehrnoosh Sadrzadeh; Stephen Clark; Bob Coecke; and Stephen Pulman. 2011. Concrete sentence spaces for compositional distributional models of meaning. In *Proceedings of the 9th International Conference on Computational Semantics (IWCS 2011)*, 125–134. Portland, OR: ACL.

Hatzivassiloglou, Vasileios and Janyce Wiebe. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In *Proceedings of the International Conference on Computational Linguistics (COLING)*.

References III

Hearst, Marti A. 1994. Multi-paragraph segmentation of expository text. In *32nd Annual Meeting of the Association for Computational Linguistics*, 9–16. Las Cruces, New Mexico: Association for Computational Linguistics.

Hearst, Marti A. 1997. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics* 23(1):33–64.

Lafferty, John; Andrew McCallum; and Fernando Pereira. 2001. Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proceedings of ICML-01*, 282–289.

Liu, Bing; Minqing Hu; and Junsheng Cheng. 2005. Opinion observer: Analyzing and comparing opinions on the web. In *Proceedings of the 14th International World Wide Web Conference*, 342–351. ACM.

Liu, Hugo; Henry Lieberman; and Ted Selker. 2003. A model of textual affect sensing using real-world knowledge. In *Proceedings of Intelligent User Interfaces (IUI)*, 125–132.

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

Manning, Christopher D. 1998. Rethinking text segmentation models: An information extraction case study. Technical Report SULTRY-98-07-01, University of Sydney.

McCallum, Andrew and Kamal Nigam. 1998. A comparison of event models for naive bayes text classification. In *AAAI/ICML-98 Workshop on Learning for Text Categorization*, 41–48. AAAI Press.

Mitchell, Jeff and Mirella Lapata. 2010. Composition in distributional models of semantics. *Cognitive Science* 34(8):1388–1429.

References IV

Neviarouskaya, Alena; Helmut Prendinger; and Mitsuru Ishizuka. 2010. Recognition of affect, judgment, and appreciation in text. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, 806–814. Beijing, China: COLING 2010 Organizing Committee.

Overgoor, Jan. 2012. *An Investigation of Trust in the CouchSurfing Community*. Master's thesis, Stanford University, Stanford, CA.

Pang, Bo and Lillian Lee. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, 271–278. Barcelona, Spain. doi:\bibinfo{doi}{10.3115/1218955.1218990}. URL http://www.aclweb.org/anthology/P04-1035.

Pang, Bo and Lillian Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, 115–124. Ann Arbor, Michigan: Association for Computational Linguistics.

Pang, Bo and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2(1):1–135.

Pang, Bo; Lillian Lee; and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 79–86. Philadelphia: Association for Computational Linguistics.

References V

Potts, Christopher. 2011. On the negativity of negation. In Nan Li and David Lutz, eds., *Proceedings of Semantics and Linguistic Theory 20*, 636–659. Ithaca, NY: CLC Publications.

Potts, Christopher and Florian Schwarz. 2010. Affective 'this'. *Linguistic Issues in Language Technology* 3(5):1–30.

Reynar, Jeffrey C. 1994. An automatic method for finding topic boundaries. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, 331–333. Las Cruces, New Mexico: Association for Computational Linguistics. doi:\bibinfo{doi}{10.3115/981732.981783}. URL http://www.aclweb.org/anthology/P94-1050.

Reynar, Jeffrey C. 1998. *Topic Segmentation: Algorithms and Applications*. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Riloff, Ellen and Janyce Wiebe. 2003. Learning extraction patterns for subjective expressions. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Russell, James A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39(6):1161–1178.

Scherer, Klaus R. 1984. Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality and Social Psychology* 5(1):37–63.

References VI

Sharp, Bernadette and Caroline Chibelushi. 2008. Text segmentation of spoken meeting transcripts. *International Journal of Speech Technology* 11:157–165. 10.1007/s10772-009-9048-2, URL http://dx.doi.org/10.1007/s10772-009-9048-2.

Snyder, Benjamin and Regina Barzilay. 2007. Multiple aspect ranking using the Good Grief algorithm. In *Proceedings of the Joint Human Language Technology/North American Chapter of the ACL Conference (HLT-NAACL)*, 300–307.

Socher, Richard; Brody Huval; Christopher D. Manning; and Andrew Y. Ng. 2012. Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the 2012 Conference on Empirical Methods in Natural Language Processing*, 1201–1211. Stroudsburg, PA.

Socher, Richard; Cliff Lin; Andrew Y. Ng; and Christopher D. Manning. 2011a. Parsing natural scenes and natural language with recursive neural networks. In *The 28th International Conference on Machine Learning*.

Socher, Richard; Jeffrey Pennington; Eric H. Huang; Andrew Y. Ng; and Christopher D. Manning. 2011b. Semi-supervised recursive autoencoders for predicting sentiment distributions. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 151–161. Edinburgh, Scotland, UK.: Association for Computational Linguistics. URL http://www.aclweb.org/anthology/D11-1014.

Stone, Philip J; Dexter C Dunphry; Marshall S Smith; and Daniel M Ogilvie. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.

References VII

Sutton, Charles and Andrew McCallum. 2010. An introduction to conditional random fields. *Foundations and Trends in Machine Learning* .

Tan, Chenhao; Lillian Lee; Jie Tang; Long Jiang; Ming Zhou; and Ping Li. 2011. User-level sentiment analysis incorporating social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1397–1405. San Diego, CA: ACM Digital Library.

Thomas, Matt; Bo Pang; and Lillian Lee. 2006. Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 327–335. Sydney, Australia: Association for Computational Linguistics. URL http://www.aclweb.org/anthology/W/W06/W06-1639.

Turney, Peter D. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of 40th Annual Meeting of the Association for Computational Linguistics*, 417–424. Philadelphia, PA: Association for Computational Linguistics. doi:\bibinfo{doi}{10.3115/1073083.1073153}. URL http://www.aclweb.org/anthology/P02-1053.

Turney, Peter D. and Michael L. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)* 21:315–346. doi:\bibinfo{doi}{http://doi.acm.org/10.1145/944012.944013}. URL http://doi.acm.org/10.1145/944012.944013.

References VIII

Wiebe, Janyce; Theresa Wilson; and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* 39(2–3):165–210.

Wiebe, Janyce M.; Rebecca F. Bruce; and Thomas P. O'Hara. 1999. Development and use of a gold standard data set for subjectivity classifications. In *Proceedings of the Association for Computational Linguistics (ACL)*, 246–253.

Wilson, Theresa; Janyce Wiebe; and Rebecca Hwa. 2006. Just how mad are you? Finding strong and weak opinion clauses. *Computational Intelligence* 2(22):73–99.