Sentiment of two women
Sentiment analysis and social media

Lillian Lee
Cornell University

Bo Pang

“Romance should never begin with sentiment. It should begin with science and end with a settlement.”

--- Oscar Wilde, An Ideal Husband
the rise of sentiment analysis

many applications and interesting questions

+ improvements in analytical techniques

+ availability of new data sets for tuning these techniques

= huge interest and progress in recognizing and analyzing subjective language
our three main points

❖ sentiment analysis is **harder** than many people think

❖ sentiment analysis is **more important** than many people think

❖ sentiment analysis involves **many problems and many approaches** --- let’s talk about these in the Q&A
an example
(romance should begin with a search)
(romance should begin with a search)

what is sxsw?
(romance should begin with a search)

what is sxsw?

easy: first result on any search engine
(romance should begin with a search)

what is sxsw?

easy: first result on any search engine

it’s a lot of money; is it worth it?
(romance should begin with a search)

what is sxsw?

easy: first result on any search engine

it’s a lot of money; is it worth it?

?
is sxsw | worth it

is sxsw worth it
is sxsw free
is sxsw interactive worth it

(automatic query suggestion can tell you a lot)
is sxsw worth it
is sxsw worth it
is sxsw free
is sxsw interactive worth it

(automatic query suggestion can tell you a lot)

Stephanie Keller: Is SXSW Worth It?
Mar 24, 2009 ... One question arises in the Lone Star soaked afterglow of SXSW: Do we Austinites love it or hate it?
www.huffingtonpost.com/.../is-sxsw-worth-it_b_178772.html - Cached - Similar

Should a band go to SXSW at all costs? - sxsw southbysouthwest ...
Jan 30, 2010 ... It might be worth it many times over to come to SXSW, but not if no one sees or hears you! There are plenty of MeFites in Austin who'd ...
ask.metafilter.com/.../Should-a-band-go-to-SXSW-at-all-costs - Cached - Similar

I ideas for promoting a band at SXSW? - SXSW music promotion
hotel suggestions for sxsw (music) - sxsw southbysouthwest austin ...
SXSW On the Cheap - SXSW cheap travel
SXSW Interactive or O'Reilly Emerging Technology? - webdesigners ...
More results from ask.metafilter.com »

Restaurants Near SXSW Worth Ditching the Conference - SXSW Maps ...
Mar 9, 2011 ... Photos: dennis / Flickr, Uchi / Facebook, Bouldin Creek / Facebook, fleur-design / Flickr] We know the joys of the Warehouse District and ...
austin.eater.com/.../restaurants-near-sxsw-worth-ditching-the-conference.php - Cached
is sxsw worth it

is sxsw worth it

is sxsw free

is sxsw interactive worth it

(automatic query suggestion can tell you a lot)
but i'm not asking whether Austinites like it!

(automatic query suggestion can tell you a lot)
let me try my favorite review site for this...
let me try my favorite review site for this...

South By Southwest

Categories: Local Flavor, Festivals [Edit]

Austin, TX 78751
(512) 467-7979
sxsw.com

Good for Kids: No
let me try my favorite review site for this...

too many to read through (for busy and important people) ...
let me try my favorite review site for this...

too many to read through (for busy and important people) ...

... yet too small a sample to draw conclusions from the average rating
sentiment analysis is harder than many people think
a piece of cake (...?)

an extremely simple setting: is a given review on a known topic positive or negative?

“It may be a bit early to make such judgments, but Battlefield Earth may well turn out to be the worst movie of this century.” (Elvis Mitchell, May 12, 2000)
a piece of cake (...?)

an extremely simple setting: is a given review on a known topic positive or negative?

“It may be a bit early to make such judgments, but Battlefield Earth may well turn out to be the worst movie of this century.” (Elvis Mitchell, May 12, 2000)
a piece of cake (...?)

an extremely simple setting: is a given review on a known topic positive or negative?

“It may be a bit early to make such judgments, but Battlefield Earth may well turn out to be the worst movie of this century.” (Elvis Mitchell, May 12, 2000)

don’t we just need to look for “worst”, “best”, “love”, hate”, etc.?
a piece of cake (...?)

but people aren’t that good at picking indicative cues.

<table>
<thead>
<tr>
<th>Positive Words</th>
<th>Lesser Positive Words</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>dazzling brilliant phenomenal excellent fantastic</td>
<td>suck terrible awful unwatchable hideous</td>
<td>58%</td>
</tr>
<tr>
<td>gripping mesmerizing riveting spectacular cool awesome thrilling badass</td>
<td>bad cliched sucks boring stupid slow</td>
<td>64%</td>
</tr>
<tr>
<td>excellent moving exciting</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**a piece of cake (...?)**

**but** people aren’t that good at picking indicative cues.

<table>
<thead>
<tr>
<th>Positive Words</th>
<th>Negative Words</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>dazzling, brilliant, phenomenal, excellent, fantastic</td>
<td>suck, terrible, awful, unwatchable, hideous</td>
<td>58%</td>
</tr>
<tr>
<td>gripping, mesmerizing, riveting, spectacular, cool, awesome, thrilling, badass, excellent, moving, exciting</td>
<td>bad, cliched, sucks, boring, stupid, slow</td>
<td>64%</td>
</tr>
<tr>
<td>love, wonderful, best, great, superb, beautiful, still</td>
<td>bad, worst, stupid, waste, boring</td>
<td>69%</td>
</tr>
</tbody>
</table>

Tuesday, March 22, 2011
a piece of cake (...?)
but people aren’t that good at picking indicative cues.

<table>
<thead>
<tr>
<th>Positive Words</th>
<th>Negative Words</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>dazzling, brilliant, phenomenal, excellent, fantastic</td>
<td>suck, terrible, awful, unwatchable, hideous</td>
<td>58%</td>
</tr>
<tr>
<td>gripping, mesmerizing, riveting, spectacular, cool, awesome, thrilling, badass, excellent, moving, exciting</td>
<td>bad, cliched, sucks, boring, stupid, slow</td>
<td>64%</td>
</tr>
<tr>
<td>love, wonderful, best, great, superb, beautiful</td>
<td>bad, worst, stupid, waste, boring, ?!, still</td>
<td>69%</td>
</tr>
</tbody>
</table>

Tuesday, March 22, 2011
we need to move beyond indicative terms

1. This laptop is a great deal.

2. The release of this laptop caused a great deal of hoopla.

3. Yeah, this laptop is a great deal ... and I’ve got a nice bridge you might be interested in.
we need to move beyond indicative terms

1. This laptop is a great deal.

2. The release of this laptop caused a great deal of hoopla.

3. Yeah, this laptop is a great deal ... and I’ve got a nice bridge you might be interested in.
This film should be brilliant. It sounds like a great plot, the actors are first [rate], and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up.

we need to move beyond indicative terms
This film should be brilliant. It sounds like a great plot, the actors are first [rate], and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up.

we need to move beyond indicative terms
This film should be brilliant. It sounds like a great plot, the actors are first [rate], and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up.
This film should be brilliant. It sounds like a great plot, the actors are first rate, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it can't hold up.
what phrases indicate what, anyway?

the phrases
   “amazing camera”
and
   “decent quality”
in reviews have *negative* effects on demand for the corresponding products.  (Archak, Ghose, and Ipeirotis 2007)
sometimes there aren’t any indicative terms

She ran the gamut of emotions from A to B.
sometimes there aren’t any indicative terms

She ran the gamut of emotions from A to B.

Read the book.
what the machine gets to see
what the machine gets to see

“bag-of-words representation”

and Austen beat books can’t conceal dig Everytime frenzy from her I Jane madden me my over own Pride read reader shin-bone skull so that the to up
what the machine gets to see

“bag-of-words representation”

and Austen beat books can’t
conceal dig Everytime frenzy from
her I Jane madden me
my over own Prejudice Pride
read reader shin-bone skull so
that the to up
what the machine gets to see

“bag-of-words representation”

and Austen beat books can't
conceal dig Everytime frenzy from
her I Jane madden me
my over own Prejudice Pride
read reader shin-bone skull so
that the to up

what now???
Jane Austen’s books madden me so that I can’t conceal my frenzy from the reader. *Everytime I read ‘Pride and Prejudice’ I want to dig her up and beat her over the skull with her own shin-bone.*
Jane Austen’s books madden me so that I can’t conceal my frenzy from the reader. *Everytime I read ‘Pride and Prejudice’ I want to dig her up and beat her over the skull with her own shin-bone.*

-- Mark Twain
Jane Austen’s books madden me so that I can’t conceal my frenzy from the reader. Everytime I read ‘Pride and Prejudice’ I want to dig her up and beat her over the skull with her own shin-bone.

-- Mark Twain

the author might not be Mark Twain, but the task could get harder...
Jane Austen’s books madden me so that I can’t conceal my frenzy from the reader. *Everytime I read ‘Pride and Prejudice’ I want to dig her up and beat her over the skull with her own shin-bone.*

-- Mark Twain

the author might not be Mark Twain, but the task could get harder...

**opinion, about what**
opinions about sxsw
opinions about sxsw
opinions about sxsw

@stse rocks the pool table at film party
#sxsw http://twitpic.com/48p2ae
32 minutes ago via Seesmic for Android

#iPad2 rocks #SXSW (@ Apple POP UP Store) http://picplz.com/X0hm
about 1 hour ago via picplz

Missing you at #sxsw! How's Germany? @thepaulmeyers @audi rocks!:-)
about 2 hours ago via Seesmic for Android

@HRMargo Glad to have you along mam. #hirefriday rocks!! #SxSW rocks as well. ((hugs))
about 2 hours ago via web
ukuleila Installed #sxsw mobile app SUCKS! Broke my phone :
about 1 hour ago via Twitter for Android

Cocoabebe Settling in ans excited to be apart of #sxsw but sucks the hotel stops roomservice at 11pm and I called at 11:20 :( http://4sq.com/hzbDov
about 2 hours ago via foursquare

quikness Fighting a cold at #SXSW sucks. iPad2 helping to make it suck less ;-) 
about 2 hours ago via HootSuite

Sunfell @acarvin I wish I was there– SXSW unfortunately falls during a time of year when I cannot get leave from work. That sucks.
about 3 hours ago via web
ukuleila Installed #sxsw mobile app SUCKS! Broke my phone :(  
About 1 hour ago via Twitter for Android

Cocoabebe Settling in ans excited to be apart of #sxsw but sucks the hotel stops roomservice at 11pm and I called at 11:20 :(
http://4sq.com/hzbDov  
About 2 hours ago via foursquare

quikness Fighting a cold at #SXSW sucks. iPad2 helping to make it suck less ;-(
About 2 hours ago via HootSuite

Sunfell @acarvin I wish I was there– SXSW unfortunately falls during a time of year when I cannot get leave from work. That sucks.
About 3 hours ago via web
**Search tip**

Use :( to find tweets with a negative or sad attitude. Example: flight :( will find tweets containing "flight" and with a negative attitude.

---

**leila** Installed #sxsw mobile app SUCKS! Broke phone :( about 1 hour ago via Twitter for Android

---

**Cocoabebe** Settling in ans excited to be apart of #sxsw but sucks the hotel stops roomservice at 11pm and I called at 11:20 :( [http://4sq.com/hzbDov](http://4sq.com/hzbDov) about 2 hours ago via foursquare

---

**quikness** Fighting a cold at #SXSW sucks. iPad2 helping to make it suck less ;-) about 2 hours ago via HootSuite

---

**Sunfell** @acarvin I wish I was there– SXSW unfortunately falls during a time of year when I cannot get leave from work. That sucks. about 3 hours ago via web
Search tip
Use :( to find tweets with a negative or sad attitude.
Example: flight :( will find tweets containing "flight" and with a negative attitude.

SXWhit home from #sxsw night 1– had a great time tonight! just missed the @rviplounge at the ginger man though– totally bummed about that :( 4 minutes ago via web

HannahDeMulta Oh no :( RT @redshoes: Lost passport near Hilton Austin. If you find it/hear of it please ping me. Australian passport name: Kim McKay #sxsw 9 minutes ago via web

barrettlane @lgilchrist @jdmoran @zmsmith so jealous you are all at #sxsw right now. Wish I was there with team @yipit :( 16 minutes ago via Twitter for iPhone
Realtime results for "sxsw sucks"

**quikness** Fighting a cold at #SXSW sucks. iPad2 helping to make it suck less ;‐)
about 1 hour ago via HootSuite

**samdotcom** Being in 'the industry' and not being at #SxSW sucks...excruciating on social media.
#employedpeopleproblems
about 4 hours ago via Twitter for Mac

**paultreichler** verizon wireless broadband connectivity at #sxsw sucks
about 8 hours ago via web

**borisraskin** sqlite is f'ing awesome. i just wanted to say that. also, i hope that #sxsw sucks this year.
#notjealous
about 9 hours ago via web
fine-grained analysis is harder
fine-grained analysis is harder

• tasks with an “extraction” flavor will be harder
fine-grained analysis is harder

- tasks with an “extraction” flavor will be harder
- need more sophisticated analysis (that I will not go into) to model the complexity of language:
  - POS, parse trees, negation (not great != terrible)
  - sarcasm, ...
whose opinion?

Amazon Glitch Unmasks War Of Reviewers

By AMY HARMON
Published: February 14, 2004

Close observers of Amazon.com noticed something peculiar this week: the company's Canadian site had suddenly revealed the identities of thousands of people who had anonymously posted book reviews on the United States site under signatures like "a reader from New York."

The weeklong glitch, which Amazon fixed afterouted reviewers complained, provided a rare glimpse at how writers and readers are wielding the online reviews as a tool to promote or pan a book -- when they think no one is watching.
whose opinion?
whose opinion?

Amazon Glitch Unmasked

By AMY HARMON
Published: February 14, 2004

Close observers of Amazon.com noticed the company's Canadian site had thousands of people who had another United States site under signature.

The weeklong glitch, which Amazon.com self-complained, provided a rare glimpse wielding the online reviews as a tool they think no one is watching.

Cyber Blackmail ala Trip Advisor

Nov 30, 2010, 6:00 PM

Destination Expert
for Tanzania, Serengeti National Park

I didn't want to hijack another thread so I thought I would start my own fun topic:

Ah yes, a new phenomenon in this tech age: Cyber Blackmail ala TripAdvisor. It has happened to me. You didn't read about it here because I paid the blackmailer. Read on:
sentiment analysis is **more important** than many people think
consumers really (think they) care about opinions

consumers report being willing to pay 20%-99% more for a 5-star-rated item than a 4-star-rated item
consumers really (think they) care about opinions

customers report being willing to pay 20\%-99\% more for a 5-star-rated item than a 4-star-rated item

academics are divided on actual impact:

Houser and Wooders 2006: of eBay's $1.6 billion of transactions in the fourth quarter of 2000, sellers' positive reputations added an estimated $55 million to sales, while non-positives reduced sales by about $15 million.

Duan, Gu, Whinston 2008: positive vs negative ratings have no effect on box office, only volume does
political opinions really (ought to) matter, too

in 2008, of the 44% of US residents (60% of Internet users) who got political information from the Internet:

- 33% said most sites they use share their point of view
- 21% said that most challenge their point of view
political opinions really (ought to) matter, too

in 2008, of the 44% of US residents (60% of Internet users) who got political information from the Internet:

- 33% said most sites they use share their point of view
- 21% said that most challenge their point of view
  (so maybe it’s not all about polarization)
political opinions really (ought to) matter, too

in 2008, of the 44% of US residents (60% of Internet users) who got political information from the Internet:

- 33% said most sites they use share their point of view
- 21% said that most challenge their point of view

(so maybe it’s not all about polarization)

the US gov’t is trying to sort through these opinions. But, people sent 400,000 comments about just one rule (labeling organic food).
opinions are a lens on human behavior

**economics:** “Does a positive reputation increase revenue, or is word of mouth enough?”

**political science:** “How has public opinion about the death penalty changed, and what are the causes?”

**sociology and social psychology:** “How are people's opinions influenced by the opinions of their friends?”
some key problems

- transferring between languages and/or domains is “unpredictable” positive or negative? “impressionante” is both, but “impressive” is positive

- summarization and calibration is your “3 stars” the same as mine?

- how do opinions spread and change?

- which opinions matter? example issues: unhelpful reviews, review spam
more on applications, approaches, research directions, connections to other fields, ...

**opinion mining and sentiment analysis**

Bo Pang and Lillian Lee


135 pp, 330+ references, full pdf posted includes bibliographies, pointers to datasets, etc.