Large-scale analysis and summarization of user opinions

Regina Barzilay
MIT

Joint work with
Branavan, Harr Chen, Jacob Eisenstein, Benjamin Snyder
The Buyer’s Dilemma

Maclaren Techno XT
$300

Evenflo Aura Stroller
$90
### Love this stroller, January 8, 2007

By **drew m** (maryland United States) - See all my reviews

I've had this stroller for my now 11 month old and love it. I put her in when she was 3 months, cell phone in one hand). It can do a bit of off-roading if needed, the wheels will take a licking as feature. Also great is the adjustable handle bars, my 6'2 husband loves to stroll her in this. My reclines almost all the way). While a tray would be nice, I've found placing a snack cup of dry for sippy.

The one bad thing about the stroller is that it is impossible to fold with one hand.

### Needs better wheels

by **Semper Gumby** , Oct 17 '01

When I first saw this stroller I loved it. It's very stylish and nice for a tall guy. (I'm 6'5") Then after only a couple weeks the carts). I don't mind spending a lot of money if I...

[Read the full review]

### Favorite Baby Purchase, May 6, 2007

By **Lael** - See all my reviews

This stroller is great! It is light-weight and folds compactly. Our two-month old reclines in it comfortably. But what I also comes with a clear plastic rain cover. I bought

### Great stroller!

by **rubysmom** , Jul 06 '03

We bought our Maclaren stroller because we wanted something I...

[Read the full review]
Okay Stroller, HORRIBLE Car Seat
by rbbm1103, Aug 22 '05

I was so excited when my parents told me they were going to purchase a travel system so that I would not have to bri... Read the full review

Terrible Product - Good Company Resolution
by champsmommy, Aug 10 '06

My experience with this stroller AND carseat were both pretty bad, however I feel that Evenflo did their best to resolve... Read the full review

Do not buy Evenflow Aura Travel System, the stroller is very unsafe and poorly made.
by marni mae, Dec 21 '05

Moms & Dads do not buy this, it is dangerous stroller and system overall. The stroller's wheels came off with my 4 1/2 month old son in it in the mall and it was only lightly used maybe 4 times. My son was a summer baby hot weather)thank goodness my husband had a firm grip on the handle b/c it would have tipped over with my son in it!! We use the car seat and the shade const down and gets caught on things, especially when you are trying to get the baby in and out of the car. I received this as a baby shower gift, so I didn't pick this out and it was not cheap at all. I learned now, and it does not have good reviews and apparently on Evenflo stroller's the wheels fall off often enough, too often, by the way the screws on the wheels are not strong b/c they are PLASTIC, NOT METAL, CHEAP PLASTIC!!
How Much to Read?

I have two children with a third on the way. The Maclaren Techno is my fourth stroller and the very BEST I've ever owned! I went through a Gerry Convenience Stroller (which squeaked horribly, was heavy, and didn't steer well), a Graco (can't remember the model, but it looked like the Convenience Stroller, was heavy, and the handles/cup holder assembly fell apart after only 1 1/2 years), and a cheap Kolcraft umbrella stroller (shower gift that was flimsy and steered horribly and the handlebars were so low that it was most uncomfortable to push around). After the Graco started to fall apart and knowing that I was expecting a second child, I tossed price to the wind and went in pursuit of the perfect stroller. I wanted something durable, lightweight, easy to push and steer, and sturdy enough to handle a stroller attachment for my older child. In all my research, the Maclaren brand name stood out above the rest, so I decided to go for it.

I just bought the 2004 Techno XT, and after only using it for 1 day (a few short hours actually), the front wheel snapped right off! I couldn't believe it...I was in total shock. I was loving this stroller too! We had a Volvo, which was causing major neck pain (my husband and I are very tall people...that is the main reason why we wanted the Techno XT...long handle heights). I also liked how smooth the ride was, but what good is a stroller that has no front wheel? I read that the reason why it may have broke, was from the cold weather (we are in NY, and last night it was about 20 degrees) and that the plastic got stressed.
Maclaren Techno XT - Cornflower Umbrella Stroller

Common attributes of Strollers

Summary

This summary is based on 28 reviews.

Average Ratings from all reviews.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Rating</td>
<td>★★★★☆☆</td>
</tr>
<tr>
<td>Durability</td>
<td>★★★★☆☆</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>★★★★☆☆</td>
</tr>
</tbody>
</table>

Pros

- Comfortable ride
- Swiveling front wheel
- Easy to glide
- Full size features in lightweight package
- Comfort for child
- Excellent handling

Cons

- Fragile canopies
- No full reclining seat
- Large when folded
Evenflo Aura - Laxford Walnut Standard Stroller

Common attributes of Strollers

This summary is based on 5 reviews.

Average Ratings from all reviews.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Rating</td>
<td>3 stars</td>
</tr>
<tr>
<td>Durability</td>
<td>3 stars</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>3 stars</td>
</tr>
</tbody>
</table>

Pros
- Great for travel
- Comfort for child
- Nice ride

Cons
- Not for rugged terrain
- Kind of heavy
- Scratches easily
- Difficult collapse
- Minimal storage
- Difficult to steer
- Large for car trunk
Automatically summarize reviews using pros/cons

- Don't have to read lots of reviews...
- Comparison of products
- Smarter search & access to content
- Mobile access
Generating Pros & Cons

- Find salient product features
  - Good Reception
  - Clarity
  - Portability
  - Battery Life

- Analyze review in terms of the features

  i liked this phone lots of great features and extras email web and it is very light weight and is small and cute it fits right in my pocket yeah right in my pocket my overall rating for this phone is nine thumbs up

  ✗ Good Reception
  ✓ Clarity
  ✓ Portability
  ✗ Battery Life

- Merge features from multiple reviews

  ✗ Good Reception
  ✓ Clarity
  ✓ Portability
  ✗ Battery Life  →  ✓ Good Reception
  ✓ Clarity
  ✓ Portability
  ✗ Battery Life
Key Technical Challenge

- Semantic inference of properties
  - Input: Property, Document
  - Output: Property holds / doesn't hold

Example:

Property: **Portable**

I liked this phone lots of great features and extras - email, web, and it is very light weight and is small and cute. It fits right in my pocket, yeah right in my pocket. My overall rating for this phone is nine thumbs up!
Related Work

- **Sentiment assessment**
  
  (Hatzivassiloglou & McKeown, 1997; Wiebe, 2000; Pang & Lee, 2002; ...)
  
  - Orthogonal to semantic inference.

- **Product feature extraction**
  
  (Hu & Liu, 2004; Popescu & Etzioni, 2005; ...)
  
  - Does not support semantic comparison.
  
  - These methods can be viewed as a preprocessing step for our system.

- **Semantic entailment**
  
  (Degan et al. 2002, ...)
  
  - Stricter definition of semantic equivalence.
  
  - Existing methods assume clean training data.
What data do we have to learn from?

Review Summary

Well swash my buckle! It's *Pirates of the Caribbean: At World's End*

**Pros**
- Visuals
- Appropriately questionable pirate morality

**Cons**
- Sequelitis, plot convolution, the whole thing is getting tired

Full Review

Okay, I didn't expect *At World's End* to revive the brilliance of *The Curse of the Black Pearl*. That's the kind of supreme karma that strikes once — and trying to replicate it is like trying to catch lightning in a bottle. One of the main reasons the first film is so delightful is because Jack Sparrow is so undeniably, delightfully weird — experiencing that for the first time is just fabulous in every way. That Johnny Depp pulled this character out of his psyche is both awe inspiring and a little scary. But like all good things, enough is enough at some point. We've gotten over the novelty of the character and since he is such a draw, no one saw fit to change him in any way for *At World's End*. Eventually it all gets old. Even so, *At World's End* is a reasonable piece of summer fluff. Nothing of substance, but it is a fun way to while away a couple of hours on a lazy afternoon with some of your favorite pirates. You won't be disappointed by the stunningly gorgeous visuals — special effects, beautiful scenery and lovely, lovely lead actors — and it is always fun to gaze upon such glorious excess on the big screen. But you won't come out of the theater bursting with enthusiasm for more, more, more! I've had enough *Pirates of the Caribbean*. It's been a fun ride, but it's time to let it be before every last bit of soul is sucked from what was once a fabulous fluke of Piratical Perfection.
What data do we have to learn from?

positive examples

visuals, pirate morality

Okay, I didn't expect At World's End to revive the brilliance of The Curse of the Black Pearl. That's the kind of supreme karmic that strikes once—and trying to replicate it is like trying to catch lightning in a bottle. One of the main reasons the first film is so delightful is because Jack Sparrow is so undeniably, delightfully weird—experiencing that for the first time is just fabulous in every way. That Johnny Depp pulled this character out of his psyche is both awe-inspiring and a little scary. But like all good things, enough is enough at some point. We've gotten over the novelty of the character and since he is such a draw, no one saw fit to change him in any way for At World's End. Eventually it all gets old. Even so, At World's End is a reasonable piece of summer fluff. Nothing of substance, but it is a fun way to white away a couple of hours on a lazy afternoon with some of your favorite pirates. You won't be disappointed by the stunningly gorgeous visuals—special effects, beautiful scenery, and lovely, lovely lead actors—and it's always fun to gaze upon such glorious excess on the big screen. But you won't come out of the theater bursting with enthusiasm for more, more, more! I've had enough Pirates of the Caribbean. It's been a fun ride, but it's time to let it be before every last bit of soul is sucked from what was once a fabulous fluke of Piratical Perfection.

great music, visuals

The stuns were pulled off excellent and looked amazing. The film did use CGI, which also made it's appearance in the first. The monster pirates and Davy Jones are all computer generated imagery. They looked really good I'll admit, but you can tell of it's CGI while the first, you couldn't tell at all. The make-up effects looked extremely well and the music is extraordinary. I love this series' music. It's all up-tempo stuff and it fits this movie really well. I'd buy the soundtrack! I believe the music makes this film work on so many levels. The action is more intense and the scenery is more beautiful just by the music! The scenery in this film is just beautiful. Lush forests, foggy swamps, white sandy beaches, the old-fashioned ships, etc. captures this film with it's looks and makes everything work, whether it's the action or the adventure. You will love where this film takes you I promise you.

great special effects, pirate morality

The acting this time is more than fascinating. By now, we are all accustomed to the fact that Johnny Depp makes a funny entrance at the beginning of each movie of the franchise. Well, this time his entrance is unlike anything you've seen before. I don't want to spoil it for you, you just have to go and see it for yourself! Johnny Depp, as usual, delivers a performance that really shows how special an actor he is. But this time, I was surprised by how great the acting of some other characters was: Geoffrey Rush (Hektor Barbossa) makes every scene that he appears in very special. Also Tom Hollander (Cutler Beckett) gets a bigger screen time this movie and delivers a great performance as one of the main antagonists.

Most noticably, Hans Zimmer takes a giant leap in this movie to compose a score that doesn't fail the epic proportions of the movie. Although I thought that the score for Dead Man's Chest didn't meet my expectations after what I've heard in The Curse of the Black Pearl, the music in this movie is the best so far, with a lot of new themes that really bring the movie to life.

too long

What is wrong with everyone reviewing these? The movie is unbearably long, the plot (or should I say, the ten thousand sub-plots) are incoherent and convoluted, and the late action and special effects don't make up for it.

I don't understand why there was a need to try and continue the series past the first movie. The first movie was great, and it was excellent by itself. The second and third destroyed all the greatness and character development and charm of the first.
A Simple Solution?

- Use a discriminative classifier!
  - positive examples : Reviews with phrase in pros/cons
  - negative examples : Reviews without phrase in pros/cons

- Performance is worse than random!
  - Training data is noisy…
Noisy Training Data

- Different phrases express the same topic

---

**Review Summary**

**Well swash my buckle! It's *Pirates of the Caribbean: At World's End***

*May 30 '07*

🌟🌟🌟

**Pros**

Visuals, visuals, visuals, appropriately questionable pirate morality

**Cons**

Sequels, plot convolution, the whole thing is getting tired

---

**Review Summary**

**The POTC Franchise Ends with a Fizzle, Not a Sizzle.**

*May 24 '07 (Updated May 24 '07)*

🌟🌟

**Pros**

Some funny moments, good acting, great battle scenes and special effects.

**Cons**

Generally dismal plot and editing, too many recycled gags.

**The Bottom Line**

It's probably about what I expected, which is not necessarily a compliment.

---

**Full Review**

Even after having a "meh" reaction to the lackluster second film in the wildly popular *Pirates of the Caribbean* trilogy, I found myself counting down the days until the release of the third film, *Pirates of the Caribbean: At World's End*. I love major movie events, even when I have my doubts about how good the film will be.
Noisy Training Data

- Missing keyphrase labels

---

Review Summary

**Well swash my buckle! It's *Pirates of the Caribbean: At World's End***

**Pros**
- Visuals, visuals, visuals
- **appropriately questionable pirate morality**

**Cons**
- Sequelitis, plot convolution, the whole thing is getting tired

---

**Full Review**

Okay, I didn't expect *At World's End* to revive the brilliance of *The Curse of the Black Pearl*. That's the kind of supreme karma that strikes once—and trying to replicate it is like trying to catch lightning in a bottle. One of the main reasons the first film is so delightful is because Jack Sparrow is so undeniably, delightfully weird—experiencing that for the first time is just fabulous in every way. That Johnny Depp pulled this character out of his psyche is both awe inspiring and a little scary. But like all good things, enough is enough at some point. We've gotten over the novelty of the character and since he is such a draw, no one saw fit to change him in any way for *At World's End*. Eventually it all gets old. Even so, *At World's End* is a reasonable piece of summer fluff. Nothing of substance, but it is a fun way to while away a couple of hours on a lazy afternoon with some of your favorite pirates. You won't be disappointed by the stunningly gorgeous visuals—special effects, beautiful scenery and **lovely, lovely lead actors**—and it is always fun to gaze upon such glorious excess on the big screen. But you won't come out of the theater bursting with enthusiasm for more, more, more! I've had enough *Pirates of the Caribbean*. It's been a fun ride, but it's time to let it be before every last bit of soul is sucked from what was once a fabulous fluke of Piratical Perfection.
Our Approach

- Training using free-text annotations available on the web.
  - Handles paraphrasing
  - Handles missing labels
Given a collection of reviews, partially annotated with keyphrases:

- Find a set of domain properties (the semantic classes of keyphrases)
  
  - **size**
    - small size
    - compact size
  
  - **battery life**
    - poor battery life
    - short battery life
  
  - **style**
    - cute
    - looks

- Learn to predict whether a document supports a property

The Task

Okay, I didn’t expect *At World’s End* to revive the brilliance of *The Curse of the Black Pearl*. That’s the kind of surprise I’ve been waiting for — and trying to reproduce. I’m like trying to catch lightning in a bottle. One of the main reasons the first film is so delightful is because Jack Sparrow and Elizabeth are such a joy. Delightfully written — experiencing that for the first time is just fabulous. And yet, Jack’s character was a bit of a shadow. But as for what goes on, enough is enough at some point. We’ve been over the roof of the character and in fact, it’s such a joy to watch and enjoy. No one can tell whether it’s in any way for *At World’s End*. Eventually it all gets old. Even so, *At World’s End* is a reasonable piece of summer fluff. Nothing of substance, but it is a fun way to while away a couple of hours on a lazy afternoon with some of your favorite pirates. You won’t be disappointed by the stunningly gorgeous visuals – special effects, beautiful scenery and lovely, lovely ship designs. And it is always fun to see such glorious excess on the big screen. But you won’t come out of the theater feeling with enthusiasm for more. More, more, more if we had enough. *Pirates of the Caribbean*.

visuals

Okay, I didn’t expect *At World’s End* to revive the brilliance of *The Curse of the Black Pearl*. That’s the kind of surprise I’ve been waiting for — and trying to reproduce. I’m like trying to catch lightning in a bottle. One of the main reasons the first film is so delightful is because Jack Sparrow and Elizabeth are such a joy. Delightfully written — experiencing that for the first time is just fabulous. And yet, Jack’s character was a bit of a shadow. But as for what goes on, enough is enough at some point. We’ve been over the roof of the character and in fact, it’s such a joy to watch and enjoy. No one can tell whether it’s in any way for *At World’s End*. Eventually it all gets old. Even so, *At World’s End* is a reasonable piece of summer fluff. Nothing of substance, but it is a fun way to while away a couple of hours on a lazy afternoon with some of your favorite pirates. You won’t be disappointed by the stunningly gorgeous visuals – special effects, beautiful scenery and lovely, lovely ship designs. And it is always fun to see such glorious excess on the big screen. But you won’t come out of the theater feeling with enthusiasm for more. More, more, more if we had enough. *Pirates of the Caribbean*. It’s been a fun ride, but it’s time to let it be before every last bit of soul is sucked from what was once a fabulous role of Practical Perfection.
Comparing Keyphrases

- Represent each keyphrase as a vector of lexical and distributional features
- Compute a matrix of pairwise similarities
Finding Domain Properties

Semantic classes
- good staff
- attentive staff
- great staff
- nice staff
- friendly staff

Property prediction
- excellent atmosphere
- pleasant atmosphere
- relaxing atmosphere
- comfortable atmosphere
- clean atmosphere
- warm atmosphere
- nice atmosphere
- fun atmosphere
The Ugly Truth

- Learning properties based on this matrix alone is noisy

Our solution: learn properties jointly with the inference task
Given a collection of reviews, partially annotated with keyphrases:

- **Find a set of domain properties** (the semantic classes of keyphrases)

- **Learn to predict whether a document supports a property**
If you are looking for picture quality this phone is not for you. but however if you are looking for excellent sound quality, good size, storage of hundreds of text messages and hundreds of good java games to download then this is the phone for you. The loud speaker is also a very clever device on this phone as well, I have found it very useful. The phonebook can store multiple details, such as addresses and email addresses, but it cannot store a picture. It only comes with one built in game (chess) but there are so many out there to download.
Accessing Topics via Word Distributions

If you are looking for picture quality this phone is not for you. But however if you are looking for excellent sound quality, good size, storage of hundreds of text messages and hundreds of good java games to download then this is the phone for you. The loud speaker is also a very clever device on this phone as well, I have found it very useful. The phonebook can store multiple details, such as addresses and email addresses, but it cannot store a picture. It only comes with one built in game (chess) but there are so many out there to download.

Each review discusses multiple topics
If you are looking for **picture quality** this phone is not for you. But however if you are looking for excellent **sound quality**, good size, storage of hundreds of text messages and hundreds of good java **games to download** then this is the phone for you. The **loud speaker** is also a very clever device on this phone as well, I have found it very useful. The **phonebook** can store multiple details, such as **addresses** and **email addresses**, but it cannot store a **picture**. It only comes with one built in **game (chess)** but there are so many out there to **download**.
Accessing Topics via Word Distributions

We are given only the documents

Hidden information:
- What are the topics?
- How are the documents divided according to these topics.

There are standard Bayesian tools which enable the underlying topic structure to be inferred.

If you are looking for **picture quality** this phone is not for you. But however if you are looking for excellent **sound quality**, good size, storage of hundreds of text messages and hundreds of good java **games** to **download** then this is the phone for you. The **loud speaker** is also a very clever device on this phone as well, I have found it very useful. The **phonebook** can store multiple details, such as **addresses** and **email addresses**, but it cannot store a **picture**. It only comes with one built in game (**chess**) but there are so many out there to **download**.
Assessing Topics via Word Distributions

Topic 1

- sound
- clear
- download
- speaker

Topic 2

- phonebook
- address
- email
What about Keyphrases?

If you are looking for picture quality this phone is not for you. But however if you are looking for excellent sound quality, good size, storage of hundreds of text messages and hundreds of good java games to download then this is the phone for you. The loud speaker is also a very clever device on this phone as well, I have found it very useful. The phonebook can store multiple details, such as addresses and email addresses, but it cannot store a picture. It only comes with one built in game (chess) but there are so many out there to download.
What about Keyphrases?

Document

Keyphrases

- clear sound
- great phone book

Text

If you are looking for picture quality this phone is not for you. But however if you are looking for excellent sound quality, good size, storage of hundreds of text messages and hundreds of good java games to download then this is the phone for you. The loud speaker is also a very clever device on this phone as well, I have found it very useful. The phonebook can store multiple details, such as addresses and email addresses, but it cannot store a picture. It only comes with one built in game (chess) but there are so many out there to download.
The Model

- keyphrase clustering
- similarity matrix

- keyphrase
- topic

- topic probability
- word topic

- word

- topic language models

$N_d$ document words

$D$ documents
The Model

\[ \psi \sim \text{Dirichlet}(\psi_0) \]
\[ x_\ell \sim \text{Multinomial}(\psi) \]
\[ s_{\ell, \ell'} \sim \begin{cases} \text{Beta}(\alpha_\equiv) & \text{if } x_\ell = x_{\ell'} \\ \text{Beta}(\alpha_\neq) & \text{otherwise} \end{cases} \]
\[ \eta_d = [\eta_{d,1} \ldots \eta_{d,K}]^T \]
\[ \text{where } \eta_{d,k} \propto \begin{cases} 1 & \text{if } x_{h_d,i} = k \text{ for any } i \\ 0 & \text{otherwise} \end{cases} \]
\[ \lambda \sim \text{Beta}(\lambda_0) \]
\[ c_{d,n} \sim \text{Bernoulli}(\lambda) \]
\[ \phi \sim \text{Dirichlet}(\phi_0) \]
\[ z_{d,n} \sim \begin{cases} \text{Multinomial}(\eta_d) & \text{if } c_{d,n} = 1 \\ \text{Multinomial}(\phi) & \text{otherwise} \end{cases} \]
\[ \theta_k \sim \text{Dirichlet}(\theta_0) \]
\[ w_{d,n} \sim \text{Multinomial}(\theta_{z_{d,n}}) \]
I've been very impressed by the sound quality of the Motorola T720. It's extremely clear and has none of that echo or tinny sound prevalent in so many mobile phones. Signal strength/reception has been fantastic compared to what I had before. This phone includes a web browser (extra fee for Internet/Data access), a phone book with multiple fields for entry (kind of like Outlook or any other actual phone book application), datebook, calculator, media center (for customizing tones & displaying photos on the phone), voice notes (for recording voice memos/notes), and functions for customizing call & ringer settings.
I've been very impressed by the sound quality of the Motorola T720. It's extremely clear and has none of that echo or tinny sound prevalent in so many mobile phones. Signal strength/reception has been fantastic compared to what I had before. This phone includes a web browser (extra fee for Internet/Data access), a phonebook with multiple fields for entry (kind of like Outlook or any other actual phonebook application), datebook, calculator, media center (for customizing tones & displaying photos on the phone), voice notes (for recording voice memos/notes), and functions for customizing call & ringer settings.
I've been very impressed by the sound quality of the Motorola T720. It's extremely clear and has none of that echo or tinny sound prevalent in so many mobile phones. Signal strength/reception has been fantastic compared to what I had before. This phone includes a web browser (extra fee for Internet/Data access), a phonebook with multiple fields for entry (kind of like Outlook or any other actual phonebook application), datebook, calculator, media center (for customizing tones & displaying photos on the phone), voice notes (for recording voice memos/notes), and functions for customizing call & ringer settings.
Predicting Document Properties

\[ p(\text{topic} \mid \text{doc}) \]  - probability of topic given document

\[ p(\text{topic} \mid \text{word}) \]  - probability of topic given word

\[ p(\text{word} \mid \text{topic}) \]  - probability of word given topic

\[ p(\text{topic}) \]  - prior probability of topic

\[ c \]  - threshold tuned on a development set

\[ p(\text{topic} \mid \text{doc}) \propto \sum_{\text{words}} p(\text{topic} \mid \text{word}) \]

\[ \propto \sum_{\text{words}} p(\text{word} \mid \text{topic}) p(\text{topic}) \]

- A topic is predicted for a document if \[ p(\text{topic} \mid \text{doc}) > c \]

- Each topic determines a unique keyphrase cluster
Learning to Predict Document Properties

- \( p(\text{topic} \mid \text{doc}) \) - probability of topic given document
- \( p(\text{topic} \mid \text{word}) \) - probability of topic given word
- \( p(\text{word} \mid \text{topic}) \) - probability of word given topic
- \( p(\text{topic}) \) - prior probability of topic
- \( c \) - threshold tuned on a development set

\[
p(\text{topic} \mid \text{doc}) \propto \sum_{\text{words}} p(\text{topic} \mid \text{word}) \\
\propto \sum_{\text{words}} p(\text{word} \mid \text{topic}) p(\text{topic})
\]

- A topic is predicted for a document if \( p(\text{topic} \mid \text{doc}) > c \)
- Each topic determines a unique keyphrase cluster
Estimating Model Parameters

Goal: Find values for parameters that maximize probability of the observed data.

Challenge: Analytically intractable.
Number of model parameters is large: $O(\text{topics} \times \text{vocabulary})$

Solution: Approximate inference methods.
Approximate Inference of Parameters

- Estimating all parameters at once is intractable.
- Estimating a single parameter if others are fixed is easy.
- Idea: successively estimate a new value for each parameter fixing all the others.
- This method is known as **Gibbs sampling**.
Gibbs Sampling
Data Sets

- Product reviews downloaded from Epinions.com
- Pros / Cons phrases used as keyphrase annotations
- Tested on two product categories:

<table>
<thead>
<tr>
<th></th>
<th>Restaurants</th>
<th>Cell Phones</th>
</tr>
</thead>
<tbody>
<tr>
<td># of reviews</td>
<td>3883</td>
<td>1112</td>
</tr>
<tr>
<td>review length</td>
<td>916.9</td>
<td>1056.9</td>
</tr>
<tr>
<td>keyphrases / review</td>
<td>3.42</td>
<td>4.91</td>
</tr>
</tbody>
</table>

- Compared against two types of annotations:
  - Free-text annotations from review
  - Gold standard annotations
Baseline methods

1. Random baseline
2. Phrase in document
3. Cluster in document
   - Any phrase from cluster present in document
4. Phrase classifier
   - Positive examples: document annotated with phrase
5. Cluster classifier
   - Positive examples: document annotated with any phrase from cluster
Examples of Automatically Computed Clusters

- **Size:**
  - size
  - small size
  - compact size
  - great size
  - good size
  - tiny size
  - nice size
  - sleek

- **Battery Life:**
  - battery life
  - short battery life
  - poor battery life
  - low battery life
  - bad battery life
  - so so battery life
  - battery life could be better
  - terrible battery life

- **Looks:**
  - looks
  - style
  - cute
  - nice design
  - nice looking appearance
  - looks cool
  - looks great
  - looks good
  - good looks
  - styling
  - **functionality** *
  - **clear calls** *

* Indicates additional attributes or functionalities.
Impact of Joint Learning on Clustering

Independent Clustering

Joint Learning

Keyphrases grouped by semantic class

Training Iterations
## Results

<table>
<thead>
<tr>
<th></th>
<th>Restaurants gold annotation</th>
<th></th>
<th>Restaurants free-text annotation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
<td>Precis.</td>
<td>F-Score</td>
<td>Recall</td>
</tr>
<tr>
<td>Random baseline</td>
<td>0.5000</td>
<td>0.3000</td>
<td>* 0.3750</td>
<td>0.5000</td>
</tr>
<tr>
<td>Phrase in text</td>
<td>0.0443</td>
<td>0.4828</td>
<td>* 0.0812</td>
<td>0.0779</td>
</tr>
<tr>
<td>Cluster in text</td>
<td>0.2880</td>
<td>0.3583</td>
<td>◇ 0.3193</td>
<td>0.5247</td>
</tr>
<tr>
<td>Phrase classifier</td>
<td>0.0222</td>
<td>0.6364</td>
<td>* 0.0428</td>
<td>0.0675</td>
</tr>
<tr>
<td>Cluster classifier</td>
<td>0.0981</td>
<td>0.4769</td>
<td>* 0.1627</td>
<td>0.2286</td>
</tr>
<tr>
<td>Our model</td>
<td>0.6076</td>
<td>0.3879</td>
<td>0.4735</td>
<td>0.7439</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Cell phones free-text annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
</tr>
<tr>
<td>Random baseline</td>
<td>0.5000</td>
</tr>
<tr>
<td>Phrase in text</td>
<td>0.1524</td>
</tr>
<tr>
<td>Cluster in text</td>
<td>0.6952</td>
</tr>
<tr>
<td>Phrase classifier</td>
<td>0.0190</td>
</tr>
<tr>
<td>Cluster classifier</td>
<td>0.1714</td>
</tr>
<tr>
<td>Our model</td>
<td>0.6762</td>
</tr>
</tbody>
</table>
## Results

Cluster quality assessment using Rand Index:

- Restaurant dataset: 0.9441
- Cellphone dataset: 0.9086

<table>
<thead>
<tr>
<th>Clustering</th>
<th>Restaurants</th>
<th>Cell phones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
<td>Precision</td>
</tr>
<tr>
<td>Cluster in text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>automatic</td>
<td>0.5247</td>
<td>0.6433</td>
</tr>
<tr>
<td>gold</td>
<td>0.5429</td>
<td>0.6076</td>
</tr>
<tr>
<td>Cluster classifier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>automatic</td>
<td>0.2286</td>
<td>0.8980</td>
</tr>
<tr>
<td>gold</td>
<td>0.2208</td>
<td>0.9043</td>
</tr>
<tr>
<td>Our model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>automatic</td>
<td>0.7439</td>
<td>0.7073</td>
</tr>
<tr>
<td>gold</td>
<td>0.7195</td>
<td>0.7084</td>
</tr>
</tbody>
</table>
Review Summarization System

- This method is the backbone of the review summarization system.
  - 476 product categories
  - 48,000 products
  - 490,000 reviews

- Training: 10 minutes approx for 2000 documents, on a 2GHz dual core CPU.
Conclusion

- New method for mapping text into a structured representation.
- The method is fully automatic, doesn't require manual annotation.
- Can be applied to analyze other types of textual documents
  - Wikipedia (keyphrases = section titles)
References


URL of the system:

http://projects.csail.mit.edu/reviews