Aspect-based Opinion Mining from Online Reviews

Samaneh Moghaddam & Martin Ester
Simon Fraser University

Tutorial at SIGIR 2012

August 12th 2012

SFU

Outline

- Introduction [15 min]
- General Opinion Mining Tasks [60 min]
- Aspect-based Opinion Mining [15 min]
- Frequency and Relation based Approaches [30 min]
- Model-based Approaches [30 min]
- Design Guidelines for LDA-based Models [15 min]
- Conclusion and Future Directions [15 min]

Outline

- Introduction
  - Social media
  - Online reviews
  - Opinion mining
  - Case study: Google Products
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Introduction

- Emergence of social media
- *Social media* are media for social interaction, using highly accessible and scalable communication techniques to create and exchange user-generated content. (Kaplan and Haenlein 2010)
- Collaborative projects, e.g. Wikipedia.
- Blogs and microblogs, e.g.Twitter.
- Content communities, e.g. Youtube.
- Social networking sites, e.g. Facebook.

Introduction

Social Media Landscape

Introduction

- Conventional media
  expensive to produce,
  only professional authors,
  one-way communication.
- Social media
  cheap to produce,
  large number of amateur authors,
  various forms of interaction among content producers and consumers (sharing, rating, and commenting on user-generated content).

Online Reviews

- More and more reviews online
  - Generic reviewing sites such as Epinions, Amazon and Cnet.
  - Specialized reviewing sites such as TripAdvisor.

Online Reviews

- Review
  - For a specific product or service
  - Full text review
  - Overall numerical rating
  - Optional: numerical rating of predefined aspects of the product / service
  - Optional: short phrases summarizing pros and cons

Online Reviews

Nikon D5000 Digital Camera

Great quality pictures, amazing camera!

Written: Aug 09 '11

Product Rating: ★★★★★

Ease of Use:

Durability:

Battery Life:

Photo Quality:

Shutter Lag

Pros: Great quality, easy to use, great settings, has video, good LCD

Cons: Video quality is not as good as it could be

Full Review:

I purchased this camera just over a year ago and I am in love with it. I was just starting out with photography, and this camera made it very easy and less confusing. The pre-set settings (Portrait, Landscape, etc.) take such great pictures that it was only until recently that I even bothered to learn how to use the manual setting. Before purchasing the D5000, I had used the Nikon D3000. The D5000 has a much better screen, and in my opinion has a better design.

Online Reviews

For consumers: aid in decision making

- when purchasing products or services. Seek opinions from friends and family.

For producers: source of consumer feedback.

- Benchmark products and services.
- Businesses spend a lot of money to obtain consumer opinions, using surveys, focus groups, opinion polls, consultants.

Online Reviews

- **97%** who made a purchase based on an online review found the review to be **accurate** (Comscore/The Kelsey Group, Oct. 2007)
- **92%** have ***more* confidence** in info found online than they do in anything from a salesclerk or other source (Wall Street Journal, Jan 2009)
- **75%** of people don't believe that companies **tell the truth** in advertisements (Yankelovich)
- **70%** consult reviews or ratings ***before* purchasing** (BusinessWeek, Oct. 2008)
- **51%** of consumers use the Internet even before making a purchase **in shops** (Verdict Research, May 2009)

Online Reviews

- **45%** say they are **influenced** a fair amount or a great deal by reviews on social sites from people they follow. (Harris Poll, April 2010)
- **34%** have turned to social media to air their **feelings** about a company. 26% to express dissatisfaction, 23% to share companies or products they like. (Harris Poll, April 2010)
- **46%** feel they can be **brutally honest** on the Internet. 38% aim to influence others when they express their preferences online (Harris Poll, April 2010)
- Reviews on a site can **boost conversion +20%** (Bazaarvoice.com/resources/stats 'Conversion Results')

http://www.searchenginepeople.com/blog/12-statistics-on-consumer-reviews.html#ixzz23C5f0C00

Online Reviews

- Other sources of opinions:
    - Customer feedback from emails, call centers, etc.
    - News and reports
    - Discussion forums
    - Tweets

→ Not as focused

→ Only text

Opinion Mining

- *Opinion* is a subjective belief, and is the result of emotion or interpretation of facts.

  http://en.wikipedia.org/wiki/Opinion

- *Sentiment*: often used as synonym of opinion.

Opinion Mining

SFU

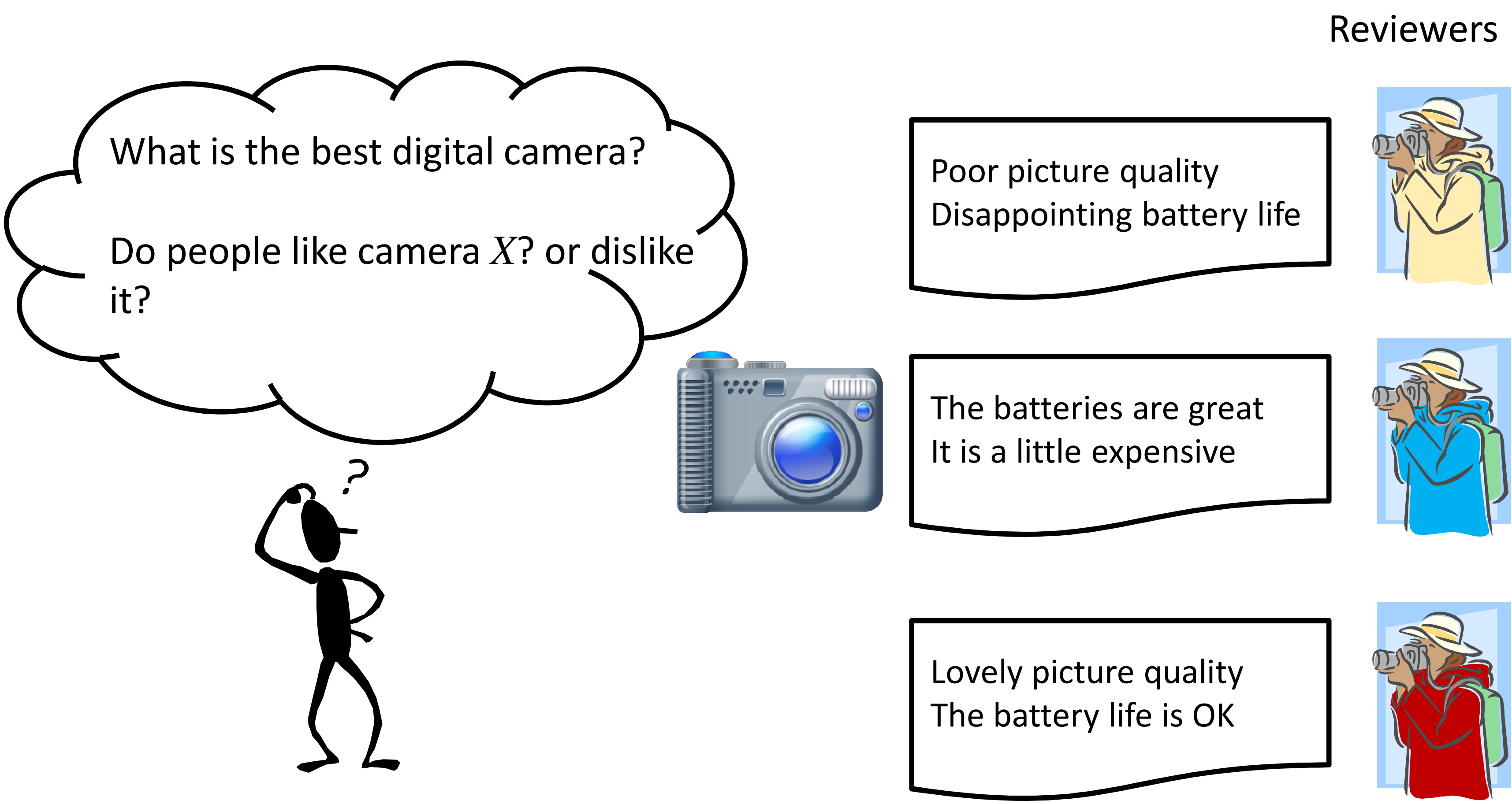
- An opinion from a single person (unless a VIP) is often not sufficient for action.
- We need to analyze opinions from many people.
- *Opinion mining*: detect patterns among opinions.

→ Opinion mining promises to have great practical impact!

→ There are already lots of practical applications.

Opinion Mining

- There are too many reviews to read.



Opinion Mining

SFU

- User-generated content is
  - mostly unstructured text,
  - lower quality,
  - noisy,
  - spam . . .
- Opinion mining is hard!

→ A thriving research area (Liu 2012)

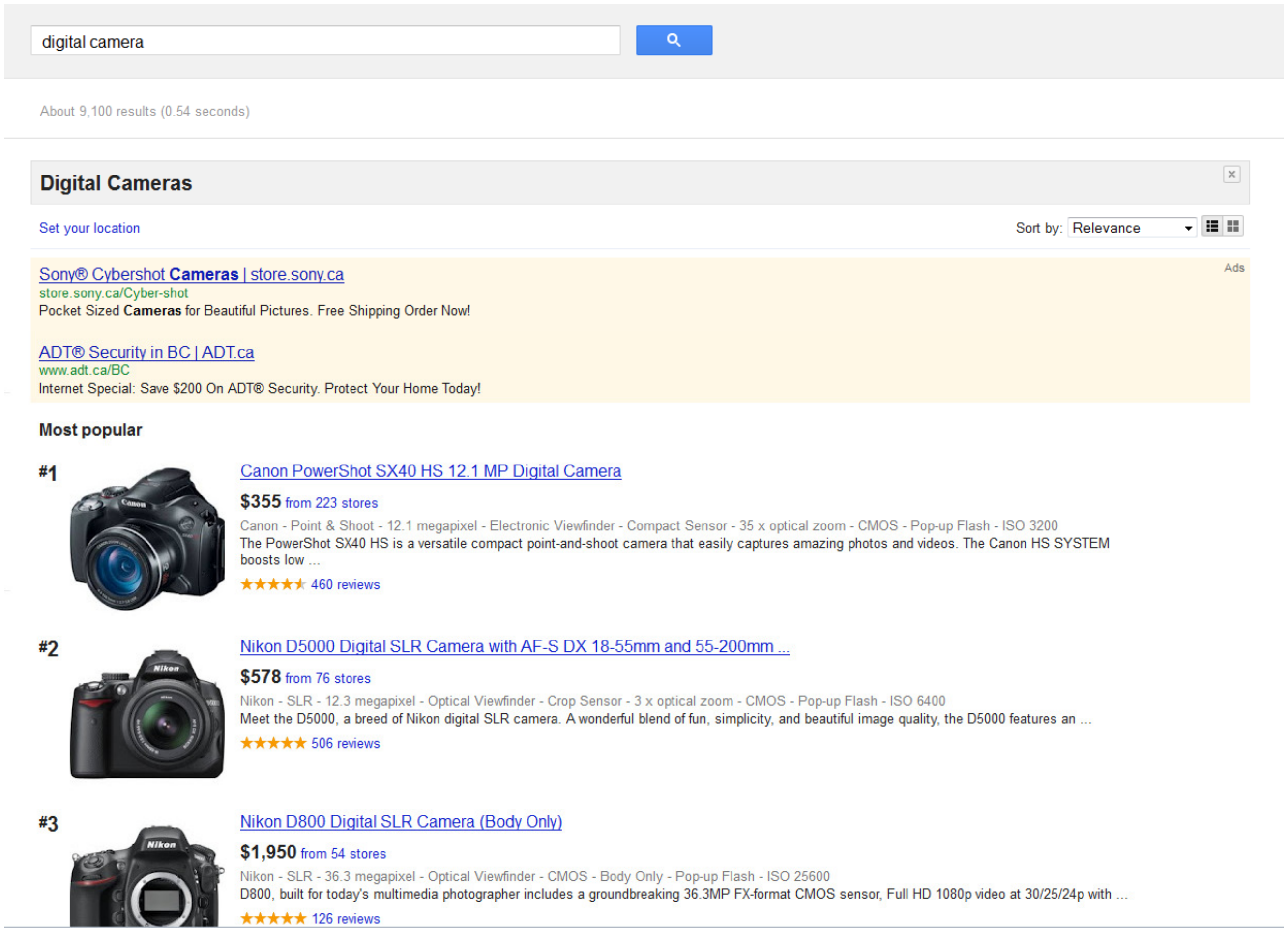
NLP, ML, data and text mining

→ Several tutorials, in particular (Liu 2011)

Case Study: Google Products

- For consumers
  - Product search and comparison
  - Online product reviews
- For producers
  - PowerReviews: structuring and analyzing user-generated content.
  - Boosts product sales, drives traffic, and increases customer engagement
  - Accelerates the impact of user-generated content throughout the entire organization.

Case Study: Google Products



Case Study: Google Products

Canon PowerShot SX40 HS 12.1 MP Digital Camera

$355 online

482 reviews +1 +25 Recommend this on Google

#1 in Digital Cameras

September 2011 - Canon - Point & Shoot - 12.1 megapixel - Electronic Viewfinder - Compact Sensor - 35 x optical zoom - CMOS - Pop-up Flash - ISO 3200

The PowerShot SX40 HS is a versatile compact point-and-shoot camera that easily captures amazing photos and videos. The Canon HS SYSTEM boosts low-light performance with the Canon DIGIC 5 Image Processor and a high-sensitivity 12.1 Megapixel CMOS sensor. This advanced Canon technology delivers stunning image quality with reduced noise and blur. Colors are more vibrant and white balance is true-to-life. The 35x Optical Zoom offers incredible reach and range, allowing you to shoot any scene, near or far. The Optical Image Stabilizer helps images come out steady and clear, and a 24mm ultra Wide-Angle lens makes it easy to take pictures of large... more »

Add to Shopping List   Browse Digital Cameras »

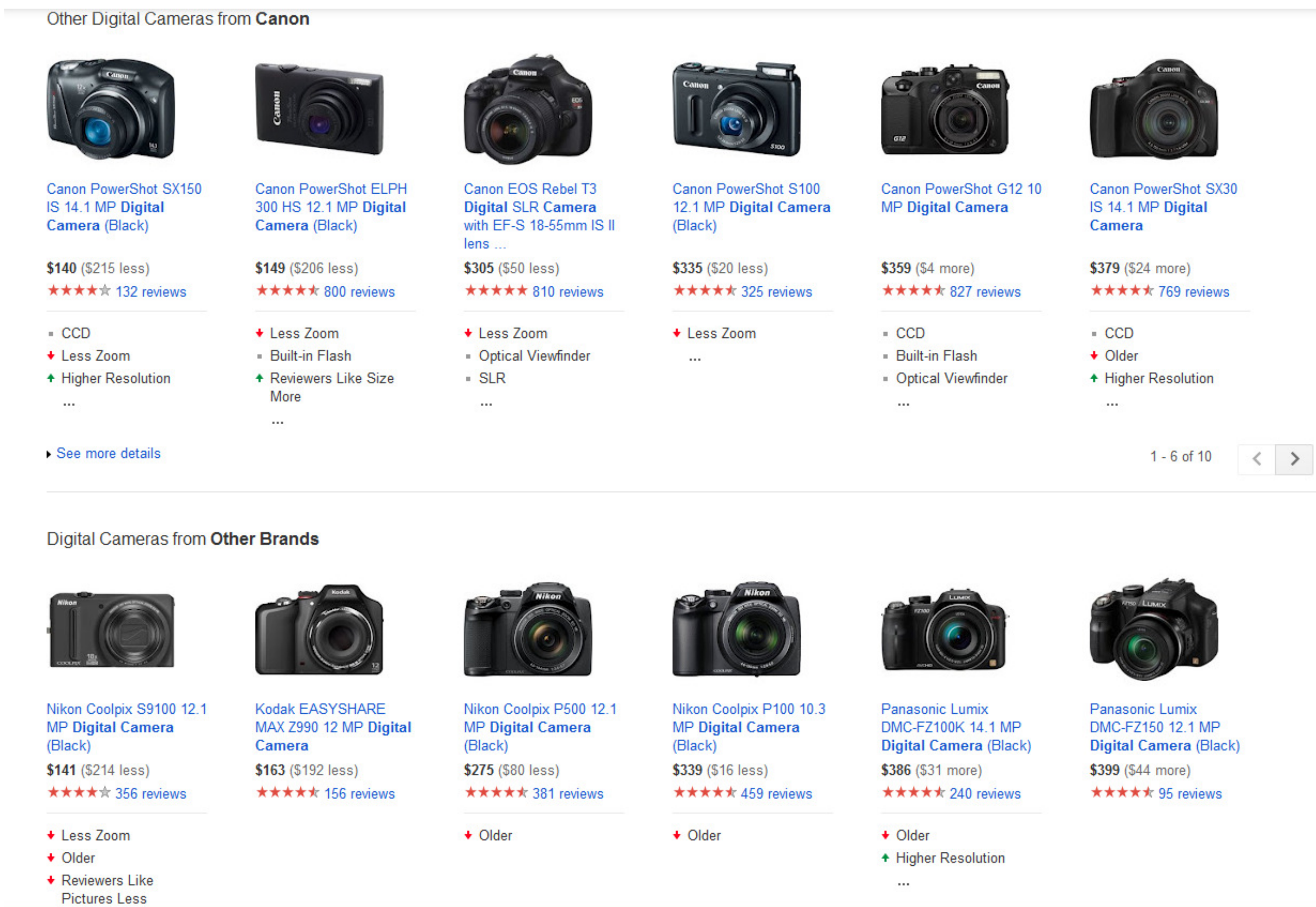
Online stores   Related items   Features   Reviews   Details   Accessories   Videos

Online stores   Google Wallet   Free shipping   New items   Your location: ZIP or city, sta   OK

Relevance	Seller rating	Condition	Tax and shipping (estimated)	Total price	Base price
Amazon.com	8,100 seller ratings	New	Free shipping		$389.00
Walmart + Show all 2	999 seller ratings	New	Free shipping		$399.00
Target	1,575 seller ratings	New			$399.99
Best Buy	2,936 seller ratings	New	Free shipping		$399.99
Overstock.com + Show all 2	4,939 seller ratings	New	Free shipping		$389.96
Dell + Show all 2	5,831 seller ratings	New	Free shipping		$399.00
Sears	4,288 seller ratings	New	Free shipping		$399.99

Case Study: Google Products

SFU



Case Study: Google Products

SFU

Canon PowerShot SX40 HS Review

★★★★★ By ConsumerSearch - May 31, 2012 - Editorial review - ConsumerSearch

Pros: Versatile lens range, excellent CMOS image sensor, shoots full 1080p HD video, optical image stabilization, lots of manual control, improved shooting speed

Cons: No RAW mode, relatively small (2.7-inch) LCD screen

The Canon PowerShot SX40 HS is basically the same camera as the critics' former favorite extreme-zoom, the Canon PowerShot SX30 IS -- only the SX40 HS boosts the shooting speed and substitutes a superior image sensor and full 1080p HD video. That gives the SX40 better image quality, especially in low light, experts say. ...
Read full review

Full review provided by:

4 out of 4 people found this review helpful. Was this review helpful? Yes - No

Review: Canon PowerShot SX40 HS

★★★★☆ By Andrew Williams - Nov 22, 2011 - Editorial review - TrustedReviews

The Canon PowerShot SX40 HS is a brilliantly versatile bridge camera whose key feature, that 35x optical zoom, is made all the more attractive by an excellent image stabilisation system. Picture quality is good and overall speed is much improved over its series precursor, the SX30. If you can live without the picture quality perfection and improved low-light performance of a DSLR, this is a great buy. Narrowly missing out on a nomination for the TrustedReviews Awards 2011 powered by Duracell, could the Canon PowerShot SX40 make the shortlist for next year's awards? Click here Getting serious about photography is not a cheap endeavour. Buying a basic DSLR setup isn't so painful, with decent models like the Nikon D3100 now available for under £400. But once you start adding the cost of the additional lenses needed for anything approaching all-purpose flexibility, you can expect to spend at least double that. If this is beyond your budget, then the Canon SX40 HS could be worth a look. It's a super zoom bridge camera that gives you a hugely flexible focal range and plenty of potential for manual control – if not quite DSLR-rivalling image quality. Canon SX40 HS 3 The Canon SX40 HS represents a significant upgrade over its predecessor, the SX30. It offers much faster performance, full HD video recording and improved light sensitivity for improved low-light performance. The effective resolution of the SX40's sensor is lower at 12.1 megapixels (instead of 14.1), but the sensor type has changed ... Read full review

4 out of 5 people found this review helpful. Was this review helpful? Yes - No

Review: Canon Powershot SX40 HS

★★★★☆ By Amy Davies - Dec 15, 2011 - Editorial review - TechRadar UK

Canon introduced the SX40 in September, at the same time as the compact PowerShot S100. The SX40 is one of a new generation of Canon cameras to be equipped with the fast Digic 5 processor. Canon promises that this boosts the HS system and now also supports Full HD (1080p) video shooting.

On board the camera is a 35x zoom, making it the longest zoom lens on any Canon compact camera. In 35mm terms, that makes the zoom range from a wide angle 24mm, to an incredible 840mm - and all this is optical zoom, not digital.

Case Study: Google Products

Canon PowerShot SX40 HS 12.1 MP Digital Camera

$355 online

★★★★★ 482 reviews +1 +25 Recommend this on Google

#1 in Digital Cameras

September 2011 - Canon - Point & Shoot - 12.1 megapixel - Electronic Viewfinder - Compact Sensor - 35 x optical zoom - CMOS - Pop-up Flash - ISO 3200

« Back to overview

Reviews

Summary - Based on 482 reviews

1 2 3 4 stars 5 stars

What people are saying

pictures	"Picture clarity is great."
features	"Great images, great features, easy to use."
zoom/lens	"This is a great low end camera for Canon DSLR users."
design	"Another point is the overall camera speed."
video	"Great video quality."
value	"Amazing product with excellent price!!"
size	"The SX 40 is very easy to use.small learning curve."

Showing reviews that mention: Size - Show all reviews

Case Study: Google Products

SFU

Review: Canon PowerShot SX40 HS

★★★★☆ By Andrew Williams - Nov 22, 2011 - Editorial review - TrustedReviews

Of course, if you're happy to carry around your camera in a rucksack rather than a pocket, "man bag" or handbag, then this needn't be a problem. Read full review

4 out of 5 people found this review helpful. Was this review helpful? Yes - No

Awesome camera

★★★★★ By Hilly - Jun 25, 2012 - B&H

I have owned Canon power shot pocket cameras exclusively over the years. Read full review

1 out of 1 people found this review helpful. Was this review helpful? Yes - No

Have Fun, Will Traval

★★★★★ By jazzman1 - May 27, 2012 - Best Buy

I have a Canon SD1000 compact camera and though the camera performs well I find it too small and light and the controls are too small for my fingers. Read full review

Was this review helpful? Yes - No

An upgrade for the budget minded.

★★★★☆ By Steve - Feb 29, 2012 - BuyDig.com

Overall, a great camera for those seeking an upgrade from a traditional "pocket" point and shoot to a more advanced camera without the DSLR price. Read full review

Was this review helpful? Yes - No

Great Camera

★★★★★ By TheITGuru - Jan 5, 2012 - Best Buy

The case cannot hold the camera, it is half the size. Read full review

Was this review helpful? Yes - No

Outline

- Introduction
- General Opinion Mining Tasks
  – Definitions
  – Subjectivity Classification
  – Sentiment Classification
  – Opinion Helpfulness Prediction
  – Opinion Spam Detection
  – Opinion Summarization
  – Mining Comparative Opinions
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Definitions

- An *opinion* is a subjective statement, view, attitude, emotion, or appraisal about an entity or an aspect of the entity (Hu and Liu 2004; Liu 2006) from an opinion holder (Bethard et al 2004; Kim and Hovy 2004).
- Sentiment orientation of an opinion: positive, negative, or neutral (no opinion).
- Also called *opinion orientation*, *semantic orientation*, *sentiment polarity*.

Definitions

- An *entity* is a concrete or abstract object such as product, person, event, organization.
- An entity can be represented as a hierarchy of components, sub-components, and so on.
- Each node represents a component and is associated with a set of attributes of the component.
- An opinion can be expressed on any node or attribute of the node.
- For simplicity, we use the term aspects to represent both components and attributes.

Definitions

- An *opinion* is a quintuple $(e_j, a_{jk}, so_{ijkl}, h_i, t_l)$ where
  - *ej* is a target entity,
  - *ajk* is an aspect of the entity *ej*,
  - *hi* is an opinion holder,
  - *tl* is the time when the opinion is expressed, and
  - *soijkl* is the sentiment orientation of opinion holder *hi* on feature *ajk* of entity *ej* at time *tl*.

Definitions

- Definition applies not only to products, but also to services, politicians, companies etc.
- The five components in (*ej*, *ajk*, *soijkl*, *hi*, *tl*) must correspond to one another. Very hard to achieve.
- The five components are essential. Without any of them, the opinion may be of limited use.

Opinion Mining (OM)

- Goal: Given an opinionated document, discover all quintuples (*ej*, *fjk*, *soijkl*, *hi*, *tl*).
- Also simpler version of the problem.
- Using the quintuples, Unstructured text → structured data.
- Traditional data mining and visualization tools can be used to visualize and analyze the results.
- Enable qualitative and quantitative analysis.

Opinion Mining

Structure of review

- Consists of sentences
- Which consist of phrases

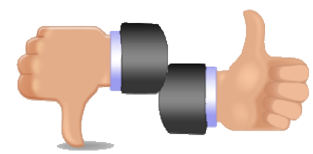
Different levels of opinion mining

- Document (review) level
- Sentence level
- Phrase level

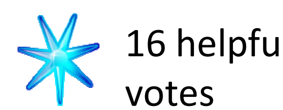
Document-level Opinion Mining

- Subjectivity Classification
  - Determines whether a given document expresses an opinion or not.
  - Can also be done at the sentence level.

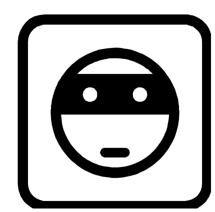
- Sentiment Classification
  - determines whether the sentiment polarity is positive or negative.



- Opinion Helpfulness Prediction
  - Estimating the helpfulness of a review.



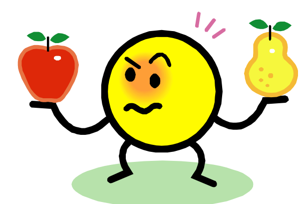
- Opinion Spam Detection
  - Identifying whether a review is spam or not.



Sentence-level Opinion Mining

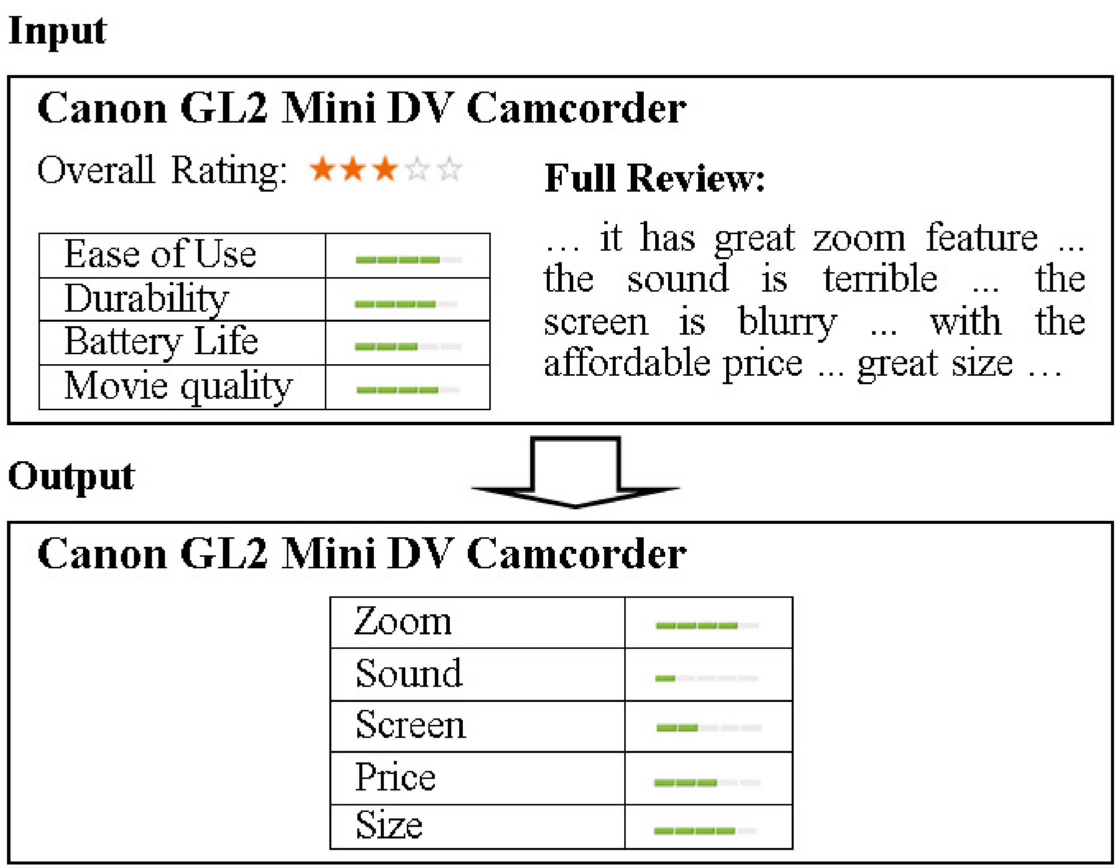
- Opinion Summarization
  - Extraction of key sentences
  - Either per product or per aspect

- Mining Comparative Opinions
  - Identification of comparative sentences
  - Extraction of comparative opinions

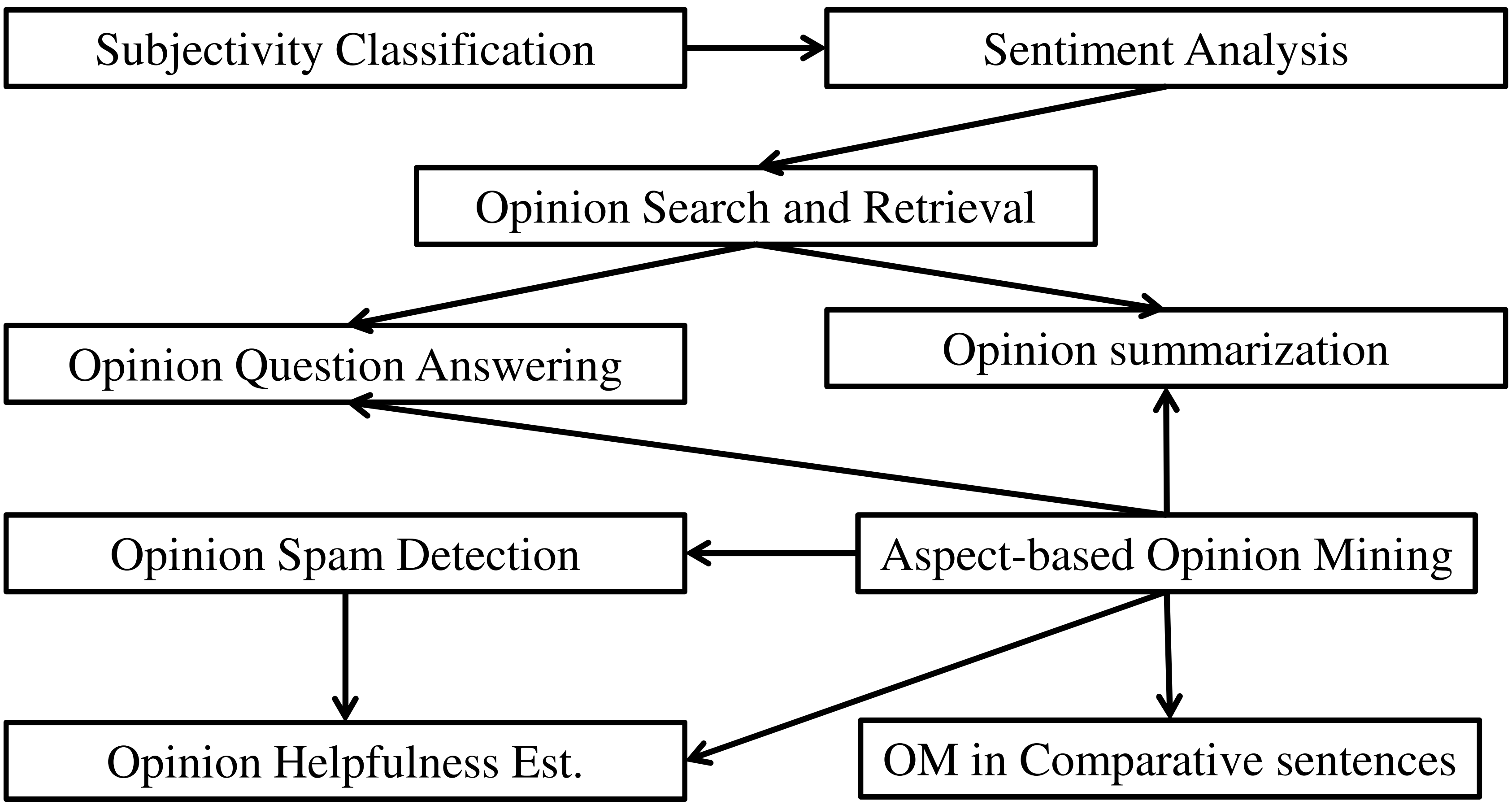


Phrase-level Opinion Mining

- Aspect-Based Opinion Mining
  - Identify aspects and their ratings from reviews



Relationships Among OM Tasks



Subjectivity Classification

- Goal is to identify subjective sentences.
- Classify a sentence into one of the two classes: objective and subjective.
- Most techniques use supervised learning.
- Assumption: Each sentence is written by a single person and expresses a single positive or negative opinion/sentiment.

Subjectivity Classification

(Rilloff and Wiebe, 2003)

- A bootstrapping approach.
- A high precision classifier is first used to automatically identify some subjective and objective sentences.
- Two high precision (but low recall) classifiers are used, a high precision subjective classifier, a high precision objective classifier.
- Based on manually collected lexical items, single words and ngrams, which are good subjective clues.

Subjectivity Classification

(Rilloff and Wiebe, 2003)

- A set of patterns is then learned from these identified subjective and objective sentences.
- Syntactic templates are provided to restrict the kinds of patterns to be discovered, e.g., <subj> passive-verb.
- The learned patterns are then used to extract more subjective and objective sentences.
- Repeat until convergence.

Sentiment Classification

- Classify a whole opinion (subjective) document based on the overall sentiment of the opinion holder (Pang et al 2002; Turney 2002).
- Classes: Positive, negative (possibly neutral).
- Neutral or no opinion is hard. Most papers ignore it.
- *"I bought an iPhone a few days ago. It is such a nice phone, although a little large. The touch screen is cool. The voice quality is clear too. I simply love it!"*

Sentiment Classification

- Different from topic-based text classification.
- In topic-based text classification (e.g., computer, sport, science), topic words are important.
- But in sentiment classification, opinion/sentiment words are more important, e.g., great, excellent, horrible, bad, worst, etc.

Sentiment Classification

(Turney 2002)

- Unsupervised approach
- Step 1: Part-of-speech (POS) tagging
  Extracting two consecutive words (two-word phrases) from reviews if their tags conform to some given patterns, e.g., (1) JJ, (2) NN.
- Step 2: Estimate the sentiment orientation (SO) of the extracted phrases

$$PMI(word_1, word_2) = \log_2\left(\frac{P(word_1 \wedge word_2)}{P(word_1)P(word_2)}\right)$$

Sentiment Classification

(Turney 2002)

Semantic orientation (SO):

SO(phrase) = PMI(phrase, “excellent”) - PMI(phrase, “poor”)

- Using near operator of a search engine to find the number of hits to compute PMI and SO.
- Step 3: Compute the average SO of all phrases Classify the review as positive if average SO is positive, negative otherwise.

Sentiment Classification

- Supervised approaches
- Key: feature engineering. A large set of features have been tried by researchers, e.g.,
  - Term frequency and different IR weighting schemes
  - Part of speech (POS) tags
  - Opinion words and phrases
  - Negations
  - Syntactic dependency

Sentiment Classification

- Dasgupta and Ng (2009) used semi-supervised learning.
- Kim et al. (2009) and Paltoglou and Thelwall (2010) studied different IR term weighting schemes.
- Mullen and Collier (2004) used PMI, syntactic relations and other features with SVM.
- Yessenalina et al. (2010) found subjective sentences and used them for model building.

Opinion Helpfulness Prediction

SFU

- Goal: Determine the usefulness, helpfulness, or utility of a review.
- Many review sites have been collecting and presenting user feedback, e.g., amazon.com.

  "x of y people found the following review helpful."
- But a review takes a long time to gather enough user feedback.

Opinion Helpfulness Prediction

- Usually formulated as a regression problem.
- A set of features is engineered for model building.
- The learned model assigns an utility score to each review.
- Ground truth for both training and testing from user helpfulness feedback.

Opinion Helpfulness Prediction

SFU

- Example features include
  - review length,
  - counts of some POS tags,
  - opinion words,
  - product aspect mentions,
  - comparison with product specifications,
  - timeliness.

(Zhang et al 2006; Kim et al. 2006; Ghose et al 2007; Liu et al 2007; Moghaddam et al 2012(a))

Opinion Spam Detection

(Jindal and Liu 2007, 2008)

- Opinion spam refers to fake or untruthful opinions, e.g.,
  - Write undeserving positive reviews for some target entities in order to promote them.
  - Write unfair or malicious negative reviews for some target entities in order to damage their reputations.
- Opinion spamming has become a business in recent years.
- Increasing number of customers are wary of spam reviews.

Opinion Spam Detection

(Jindal and Liu 2007, 2008)

- Manual labeling of training/test dataset is extremely hard.
- Propose to use duplicate and near-duplicate reviews as positive training data.
- Use non-duplicate reviews as negative training data.

Opinion Spam Detection

(Jindal and Liu 2007, 2008)

- Use the following features:
- Review centric features (content)

  n-grams, ratings, etc
- Reviewer centric features

  different unusual behaviors, etc
- Product centric features

  sales rank, etc

Opinion Summarization

- An opinion from a single person is usually not sufficient, unless from a VIP.

→ multi-document summarization

- Traditional approach: produce a short text summary by extracting some important sentences.
  E.g. (Lerman et al 2009)
- Weakness: It is only qualitative but not quantitative.

Opinion Summarization

- Novel approach for review summarization: Use aspects as basis for a summary.
- We have discussed the aspect-based summary using quintuples earlier (Hu and Liu 2004; Liu, 2010).
- Also called: *Structured Summary*
- Similar approaches taken by most topic model-based methods.

Mining Comparative Opinions

- Objective: Given an opinionated document *d*, extract *comparative opinions*: (*E*1, *E*2, *A*, *po, h, t*), where *E*1 and *E*2 are the entity sets being compared based on their shared aspects *A*, *po* is the preferred entity set of the opinion holder *h*, and *t* is the time when the comparative opinion is expressed.

(Ganapathibhotla and Liu, 2008)

Mining Comparative Opinions

SFU

- In English, comparatives are usually formed by adding *-er* and superlatives are formed by adding *-est* to their base adjectives or adverbs.
- Adjectives and adverbs with two syllables or more and not ending in *y* do not form comparatives or superlatives by adding *-er* or *-est*.
- Instead, *more*, *most*, *less*, and *least* are used before such words, e.g., *more beautiful*.

Mining Comparative Opinions

SFU

- Challenge: recognize non-standard comparatives

  E.g., “I am so happy because my new iPhone is nothing like my old slow ugly Droid.”

- Comparative opinions can be easily converted into regular opinions. E.g.,

“optics of camera A is better than that of camera B”

Positive opinion: *“optics of camera A”*

Negative opinion: *“optics of camera B”*

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
  - Problem Definition
  - Challenges
  - Evaluation Metrics
  - Benchmark Datasets
- Frequency and Relation based Approaches
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Aspect-based Opinion Mining

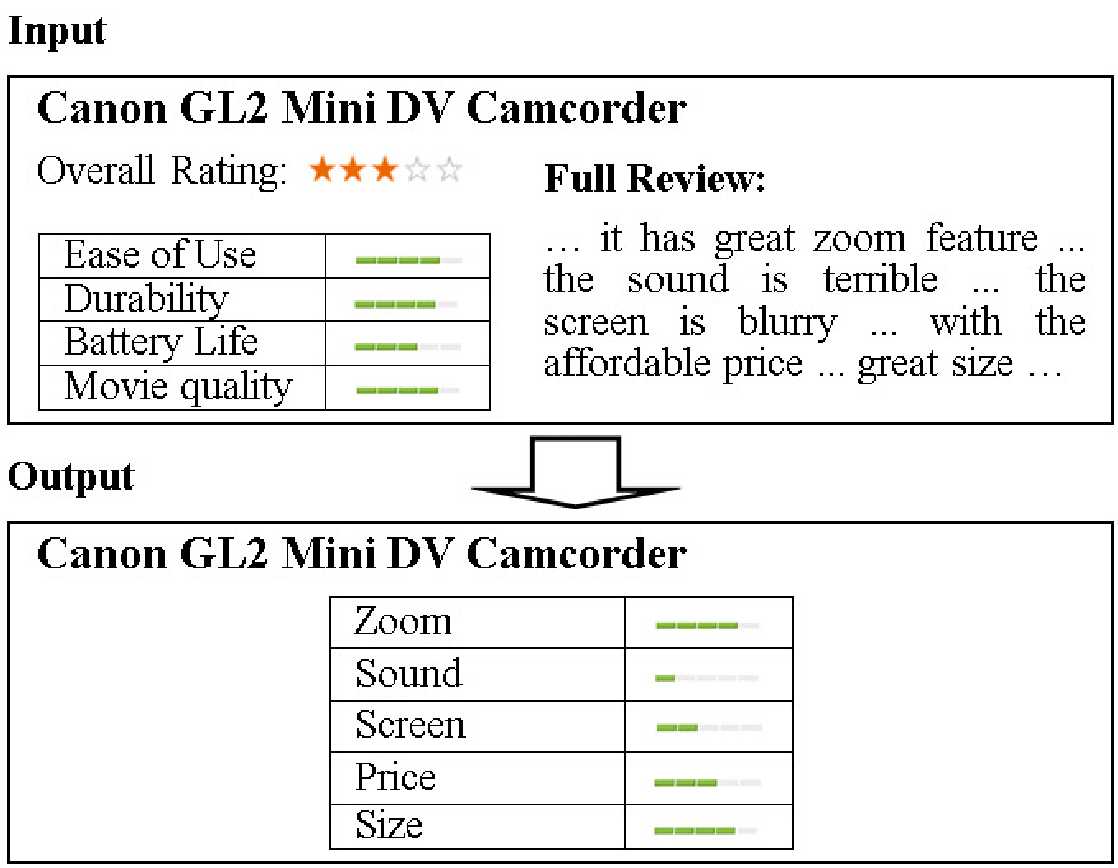
SFU

- Opinion $(e_j, aj_k, soi_{jkl}, hi, tl)$
- In online reviews: entity, opinion holder, and time explicitly provided.
- In blogs, forum discussions, etc.:
  both entity and aspects of entity are unknown,
  there may also be many comparisons, and
  there is also a lot of irrelevant information.
- Much of the research addresses online reviews

Aspect-based Opinion Mining

SFU

- Problem: identify aspects and their sentiment orientation (ratings) from a set of reviews



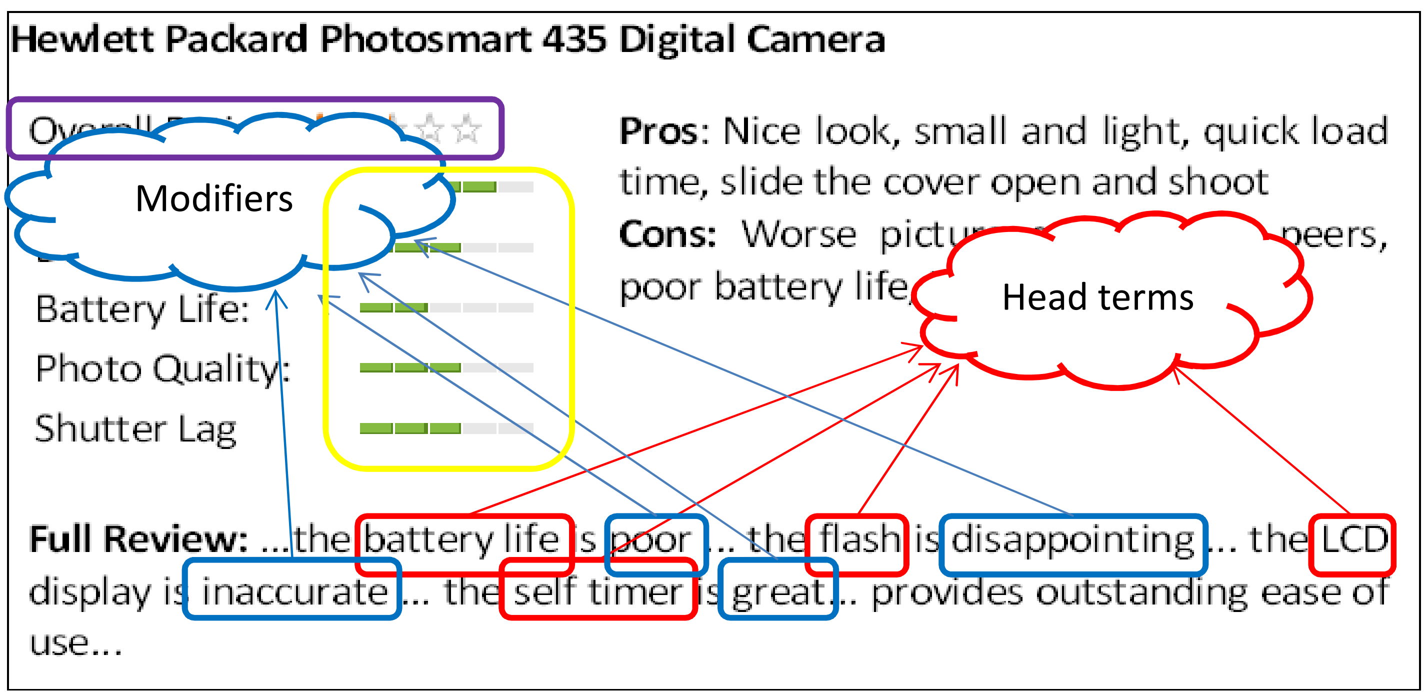
Aspect-based Opinion Mining

Typically, three step-approach:

1. Extract opinion phrases: <head term, modifier>

   e.g., <LCD, blurry>,<screen, inaccurate>, <display, poor>

2. Cluster head terms referring to same aspect and modifiers referring to same rating
3. Choose “names” for aspects and ratings



Aspect-based Opinion Mining

SFU

Input

Canon GL2 Mini DVD Camcorder
… excellent zoom ... blurry lcd … great picture quality … accurate zooming ... poor battery … inaccurate screen … good quality … affordable price … poor display … inadequate battery life … fantastic zoom … great price …

Aspect Extraction

Rating Prediction

Output

Aspect	Rating
zoom	5
price	4
picture quality	4
battery life	2
screen	1
…	…

Challenges

- Different head terms for same aspect

Photo quality is a little better than most of the cameras in this class.

That gives the SX40 better image quality, especially in low light, experts say.

These images are recorded in full resolution, making it particularly useful for shooting fast moving subjects.

- Different modifiers for same rating

For a camera of this price, the picture quality is amazing.

I am going on a trip to France and wanted something that could take stunning pictures with, but didn't cost a small fortune.

Challenges

- "Noise"

Canon is a company that never rests on its laurels, instead choosing to make continuous refinements and upgrades to its cameras.

I have owned Canon power shot pocket cameras exclusively over the years.

First of all, I am an amateur photographer and love to take close-ups.

It's been three months since I bought this camera and I can definitely say that I made a right decision.

I have fat hands but short fingers.

PS ordered Saturday arrived Tuesday (bank holiday Monday). Amazing service.

Was humilated by management trying to get a price match so I nearly left the camera and walked out. I wish I would have!

Challenges

- Relatively easy: explicit aspects and ratings

That gives the SX40 better image quality, especially in low light, experts say.

Cons: a bit bulky in size.

- Challenging: implicit aspects or ratings

After a twenty-one mile bike ride a four mile backpacking river hike, the size, weight, and performance of this camera has been the answer to my needs.

The grip and weight make it easy to handle and the mid zoom pictures have exceeded expectation.

Challenges

- Comparative opinions

This camera is everything the SX30 should have been and was not.

The SX40 HS significantly improves the low light performance of its predecessors.

Yes, I have used comparable Nikon and Olympus products. The SX40 HS is the best for me...

I bought this camera as an upgrade to my Panasonic Fz28 which I have had for a few years and found it to be inferior in almost every aspect.

Evaluation

- Which aspect-based OM mining method is best (for a given application)?
- Need to have appropriate evaluation metrics.
- Need benchmark datasets.

Evaluation Metrics

- For aspect extraction
  - Based on ground truth, i.e. “true” groupings of head terms into aspects.
  - Match extracted aspects to gold standard aspects.
  - For each matched pair of aspects, compare sets of head terms.
  - Compute precision, recall, F-measure, KL-divergence.

Evaluation Metrics

SFU

- For aspect extraction
  - Using gold standard groupings

$$Precision = \frac{|ExtractedAspects \bigcap GoldStandardAspects|}{|ExtractedAspects|}$$

$$Recall = \frac{|ExtractedAspects \bigcap GoldStandardAspects|}{|GoldStandardAspects|}$$

$$F-measure = \frac{2 \times Recall \times Precision}{(Recall + Precision)}$$

Evaluation Metrics

SFU

- For aspect extraction
  - Using ground truth word distribution

$$D_{KL}(P\|Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

$P(x)$: ground truth word distribution
$Q(x)$: word distribution learned by model

Evaluation Metrics

- For rating prediction
  - Based on ground truth, i.e. “true” groupings of modifiers into sentiment (orientations) and assignment of sentiments to “true” ratings.
  - Match extracted sentiments to gold standard sentiments.
  - Compare predicted ratings against “true” ratings.
  - Compute Mean Absolute Error (MAE), Mean Squared Error (MSE), and Ranking Loss.

Evaluation Metrics

- For rating prediction

$$MAE = \frac{1}{k}\sum_{i=1}^{k}|\hat{r_i} - r_i|$$

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(\hat{r_i} - r_i)^2$$

$\hat{r}_i$ estimated rating of the ith aspect
r_i true rating of the ith aspect
k total number of aspects

$$RankingLoss = \sum_{n}\frac{ActualRanking_n - PredictedRanking_n}{N}$$

Evaluation Metrics

- These metrics are meaningful from an application point of view.
- But ground truth is expensive to obtain, since it typically requires manual labeling.
- Sometimes, background knowledge can be exploited to obtain ground truth, e.g.

reviewers provide ratings for predefined aspects or reviewing sites provide rating guidelines.

Evaluation Metrics

- For latent variable models (probabilistic graphical models), perform cross-validation and compute the likelihood of withheld test data given different models

$$perplexity(D_{test}) = exp\{-\frac{\sum_{d=1}^{N} \log P(\boldsymbol{v}_d)}{\sum_{d=1}^{N} M_d}\}$$

N : #reviews M : #observed variables in each review v_d

- No need for ground truth, but metric less meaningful from an application point of view.

Benchmark Data Sets

- Hotel reviews dataset from TripAdvisor (http://sifaka.cs.uiuc.edu/~wang296/Data/LARA/TripAdvisor/)
- #hotels 2,232, #reviews 37,181, #reviewers 34,187, avg length 96.5
- In addition to the overall ratings, reviewers are also asked to provide ratings on 7 pre-defined aspects in each review (*value, room, location, cleanliness, check in/front desk, service, business service*) ranging from 1 star to 5 stars.

Benchmark Data Sets

- MP3 reviews data set from Amazon (http://sifaka.cs.uiuc.edu/~wang296/Data/LARA/Amazon/mp3/)
- #MP3s 686, #reviews 16,680, #reviewers 15,004, avg length 87.3
- There is only one overall rating in each review, ranging from 1 star to 5 stars.

Benchmark Data Sets

SFU

- Review dataset from Amazon
  http://liu.cs.uic.edu/download/data/
- Contains 5.8 million reviews from 2.14 million reviewers.
- Each review consists of 8 parts

<Product ID> <Reviewer ID> <Rating> <Date> <Review Title> <Review Body> <Number of Helpful Feedbacks> <Number of Feedbacks>

Benchmark Data Sets

- Review dataset from Epinions
  http://www.sfu.ca/~sam39/Datasets/EpinionsReviews/
- #Reviews 1,560,144, #Products 200,953,
  #Reviewers 326,983, #Raters 120,492,
  #Rated Reviews 755,760, #Ratings 13,668,320
- Contains not only text of reviews, but also ratings of reviews assigned by different users (raters).

Aspect-based Opinion Mining from Online Reviews

Samaneh Moghaddam & Martin Ester

Tutorial at SIGIR 2012

Part 2

Aspect-Based Opinion Mining

Input

Canon GL2 Mini DVD Camcorder
… excellent zoom ... blurry lcd … great picture quality … accurate zoom ... poor battery … inaccurate screen … good quality … affordable price … poor display … inadequate battery life… fantastic zoom … great price …

Output

Aspect	Rating
zoom	5
price	4
picture quality	4
battery life	2
screen	1
…	…

Aspect-Based Opinion Mining

- Tasks
  - Aspect extraction
  - Rating (polarity) prediction
  - Aspect grouping
  - Coreference resolution
  - Entity, opinion holder, time extraction

Aspect-Based Opinion Mining

- Aspect extraction (Liu 2012)
  - Frequency-based
  - Relation-based
  - Supervised learning
  - Topic modeling

- Rating (polarity) prediction
  - Supervised learning
  - Lexicon-based

Rating (polarity) Prediction

- Supervised learning approach
  - Applying classification techniques (e.g., Snyder et al. 2007, Wei et al. 2010).
- Limitations
  - Needs training data
  - Dependent on the domain

Rating (polarity) Prediction

- Lexicon-based approach
  - They use sentiment lexicon, e.g., GI, MPQA, SentiWordNet, etc. (e.g., Hu and Liu 2004, Ding et al. 2008).

- Strength
  – Typically unsupervised
  – Perform quite well in a large number of domains

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
  - Frequency Methods
    - Feature-based Summarization
    - OPINE
  - Relation-based Methods
  - Hybrid Methods
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Frequency-based Methods

- Applying constraints on high-frequency noun phrases to identify product aspects

- Why
  - An aspect can be expressed by a noun, adjective, verb or adverb.
  - Recent research (Liu 2007) shows that 60-70% of the aspects are explicit nouns.
  - In reviews people more likely to talk about aspects which suggests that aspects should be frequent nouns.
  - Not all frequent nouns are aspects.

Frequency-based Methods

- Feature-based Summarization (Hu et al. 2004)
  - Find freq. noun phrases
  - Filter them (compactness and redundancy)
  - Extract nearby adjectives as sentiment
  - Identify polarity of aspects using seed adjectives with known polarity
  - Find infrequent aspects using extracted sentiments

Frequency-based Methods

SFU

- OPINE (Popescu et al. 2005)
  - Find freq noun phrases
  - Use KnowItAll to extract generic patterns
    - e.g., "great X", "has X", "comes with X" (X is a potential aspect).
  - Score candidates using patterns

$$PMI(f,p) = \frac{Hits(f+p)}{Hits(p) \times Hits(f)}$$

  - Apply predefined syntactic patterns to extract sentiment
  - Identify polarity using a learned classifier

Frequency-based Methods

- Other methods
  - Ku et al. (2006)
  - Scaffidi et al. (2007)
  - Zhu et al. (2009)
  - Raju et al. (2009)
  - Long et al. (2010)

Frequency-based Methods

- Strength
  - Although these methods are very simple, they are actually quite effective.
- Limitations
  - Produce too many non-aspects and miss low-frequency aspects.
  - Require the manual tuning of various parameters which makes them hard to port to another dataset.

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
  - Frequency Methods
  - Relation-based Methods
    - Opinion Observer
    - Multi-Facet Rating
    - Tree Kernel approach
  - Hybrid Methods
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Relation-based Methods

- Exploit aspect-sentiment relationships to extract new aspects and sentiments

- Why
  - Sentiments are often known or easy-to-find
  - Each sentiment expresses an opinion on a target
  - Their relationship can be used

Relation-based Methods

- Opinion Observer (Liu et al. 2005)
  - POS tag a training set of reviews
  - Manually replace aspects by a specific tag [aspect]
    - e.g., 'zoom is great' → 'zoom_NN is _VB great_JJ' → '[aspect]_NN is_VB great_JJ' → '[aspect]_NN _VB _JJ'.
  - Apply association rule mining to find POS patterns
  - Extract aspects by applying patterns
  - Identify polarity using appearance in Pros/Cons

Relation-based Methods

- Multi-Facet Rating [Baccianella et al. 2009]
  - Predefined POS patterns
  - Polarity → General Inquirer
  - Filtering → variance of distribution across polarity

PATTERN ::= A *| B | C*
A ::= [AT] ADJ NOUN
B ::= NOUN VERB ADJ
C ::= Hv A
NOUN ::= [AT] [NN$] NN
ADJ ::= [CONG] ADV ADJ
ADV ::= RB ADV *| QL ADV | JJ | AP ADV |*
CONG ::= CC *| CS*
VERB ::= V *| Be*

Relation-based Methods

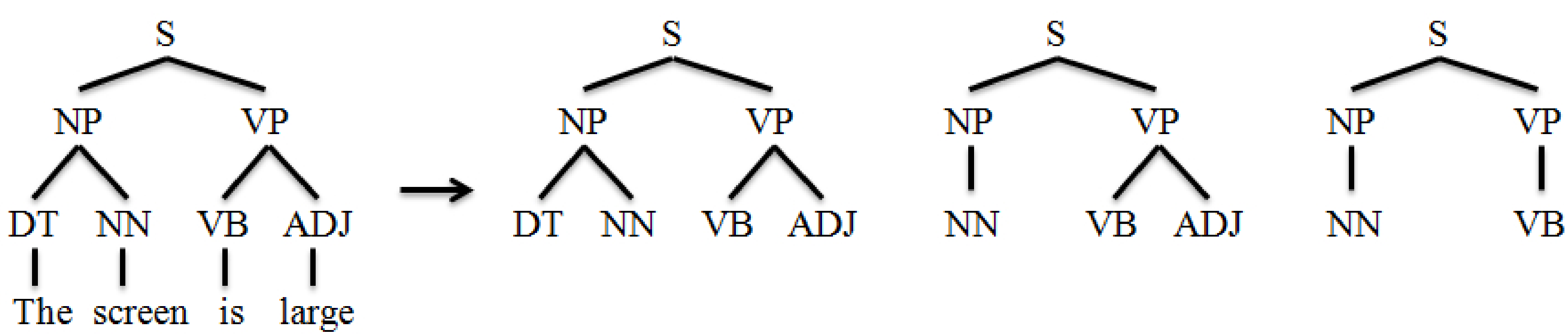
- Multi-Facet Rating [Baccianella et al. 2009]
  - Predefined POS patterns
  - Polarity → General Inquirer
  - Filtering → variance of distribution across polarity

Extracted Opinion Phrases	Enriched GI Expressions
great location	[Strong] [Positive] location
great hotel	[Strong] [Positive] hotel
very friendly staff	very [Emot] [positive] staff
good location	[Positive] location

Relation-based Methods

SFU

- Tree Kernel Approach (Jiang et al. 2010)
  - Using tree kernels to improve the limitation of exact matching
  - Exploring the substructure of the syntactic structure



Relation-based Methods

- Other methods
  - Zhuang et al. (2006)
  - Du et al. (2009)
  - Zhang et al. (2010)
  - Hai et al. (2011)
  - Qiu et al. (2011)

Relation-based Methods

- Strength
  - Can find low-frequency aspects
- Limitation
  - Produce many non-aspects matching with the patterns

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Methods
  - Frequency Methods
  - Relation-based Methods
  - Hybrid Methods
    - Sentiment Summarizer
    - Opinion Digger
- Model-based Approaches
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Hybrid Methods

- Using aspect-sentiment relations for filtering frequent noun phrases

- Why
  - Aspects are mostly frequent nouns
  - Aspects and sentiments have some relationship

Hybrid Methods

- Sentiment Summarizer (Blair-Goldensohn et al. 2008)
  - Classify sentences as positive/negative/neutral
  - Extract frequent noun phrases
  - Filter
    - Predefined syntactic patterns
    - Appearance in sentiment-bearing sentences
  - Rating of an aspect is the aggregation of sentence polarity

Hybrid Methods

SFU

- Opinion Digger (Moghaddam et al. 2010)
  - Aspect extraction
    - Find freq. noun phrases
    - Mine opinion patterns using known aspects
    - Filter based on the number of matching patterns.

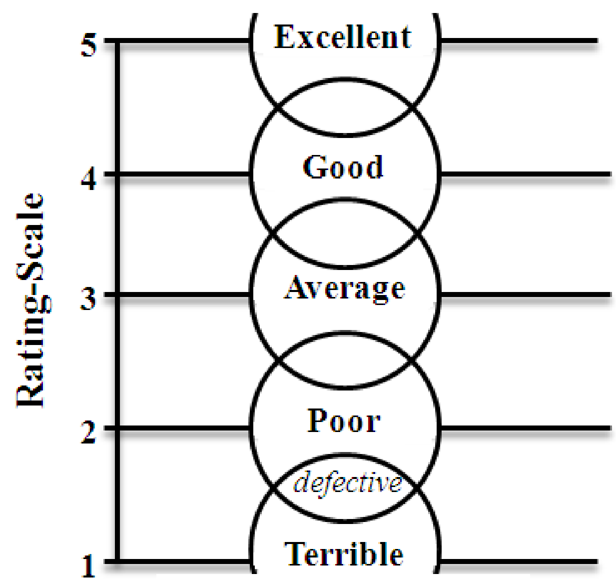
Known Aspect	**Matching segments**	**Mined Patterns**
photo quality	disappointing photo quality	_JJ_ASP
battery life	battery life is great	_ASP_VB_JJ
photo quality	lovely feature is photo quality	_JJ_NP_VB_ASP

SFU

Hybrid Methods

- Opinion Digger (Moghaddam et al. 2010)
  - Rating Prediction
    - Extract the nearest adjectives as sentiment
    - Estimate rating [1, 5] using the rating guideline
      - Use KNN algorithm
      - Use Wordnet to compute similarity.
    - Aggregate rating of sentiments

    to estimate rating of aspects.



Hybrid Methods

- Other methods
  - Li et al. (2009)
  - Zhao et al. (2010)
  - Yu et al. (2011)

Hybrid Methods

- Strength
  – Limit the number of non-aspect
- Limitations
  – Miss low-frequency aspects
  – Require the manual tuning of various parameters

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
  - Supervised learning techniques
    - OpinionMiner
    - Skip-Tree CRF
    - CFACTS
    - Wong's Model
  - Topic modeling techniques
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

Supervised Learning Techniques

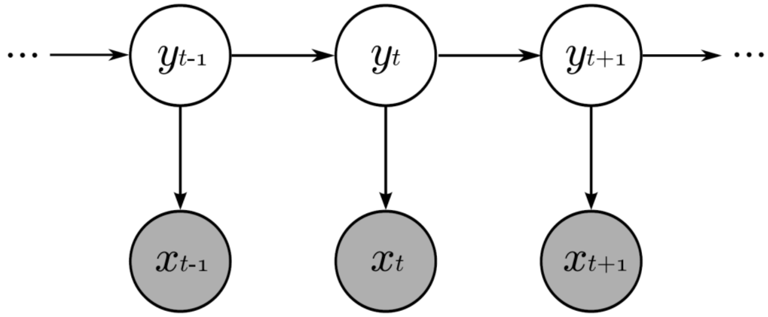
- Inferring a function from labeled (supervised) training data to apply for unlabeled data.

- Why
  - Identifying aspects, sentiments, and their polarity can be seen as a labelling problem.

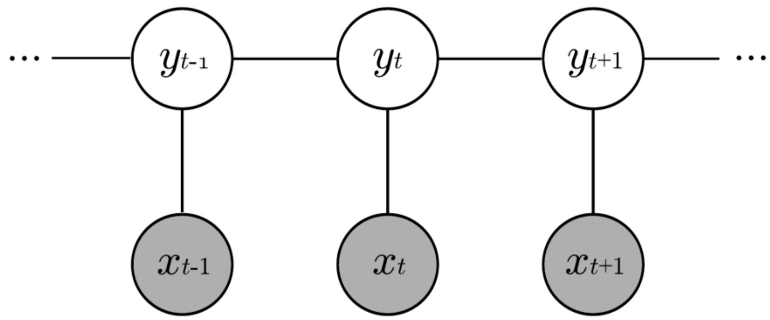
Supervised Learning Techniques

- HMM

$$p(x, y) = \prod_{t=1}^{T} p(y_t \mid y_{t-1}) \, p(x_t \mid y_t)$$



- CRF



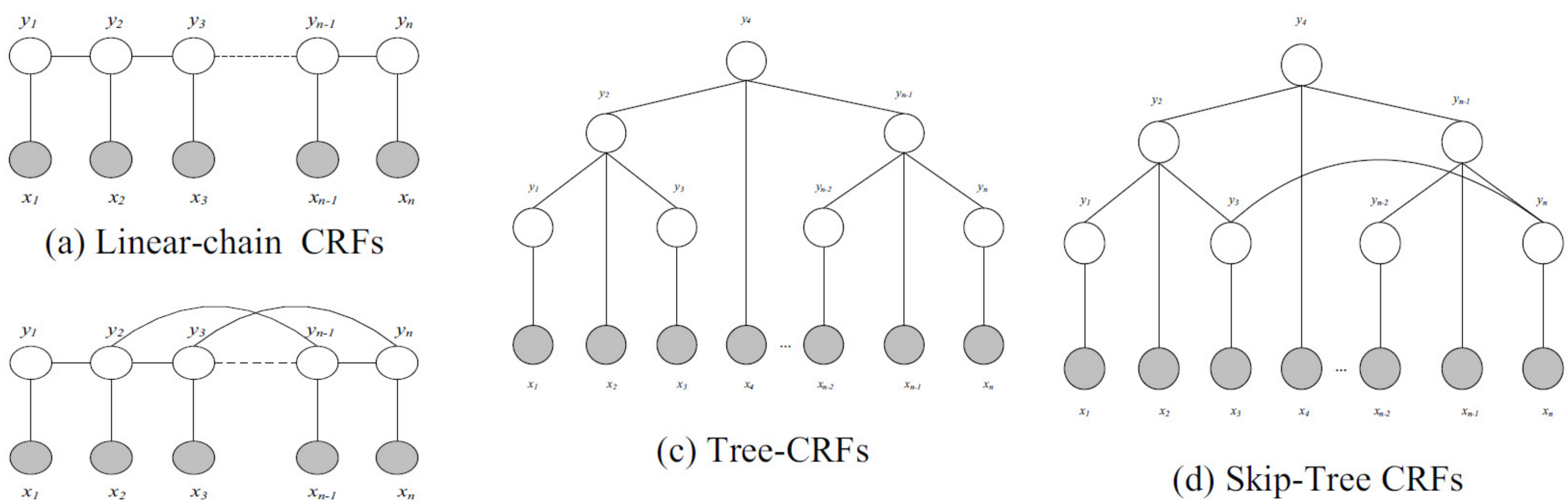
$$p(y \mid x) = \frac{1}{Z(x)} \exp\{ \sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t) \}$$

Supervised Learning Techniques

- OpinionMiner [Jin et al. 2009]
  - Base model: HMM
  - Task: identifying aspects, sentiments, their polarity
  - Novelty: integrating POS information with the lexicalization technique.

Supervised Learning Techniques

- Skip-Tree CRF (Li et al. 2010)
  - Base Model: CRF
  - Task: extracting aspects, sentiments, their polarity
  - Novelty: utilize the conjunction structure and syntactic tree structure



(a) Linear-chain CRFs

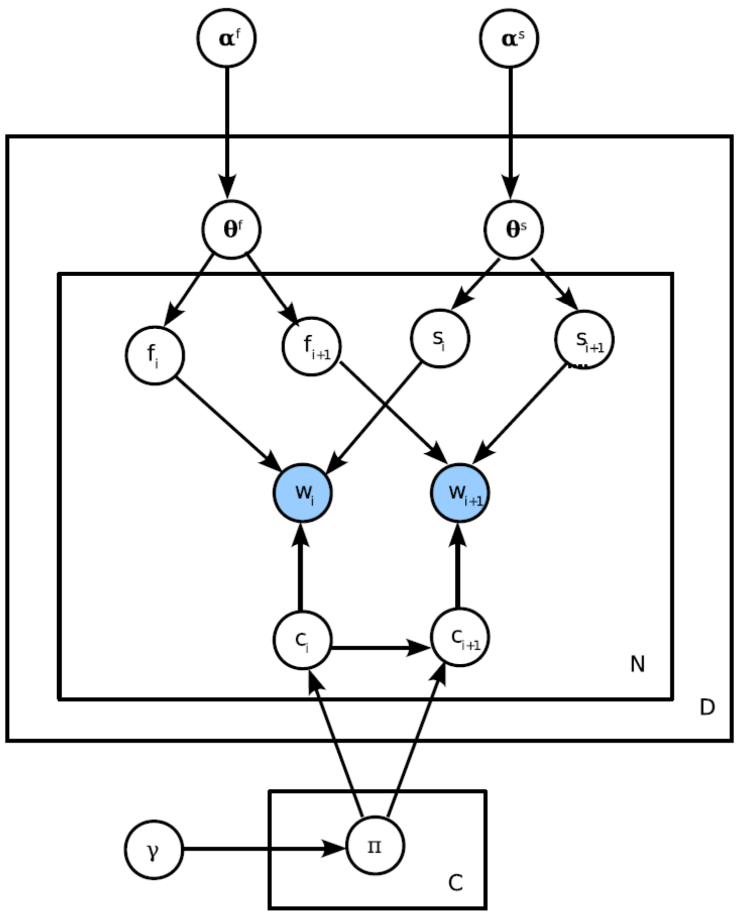
(b) Skip-chain CRFs

(c) Tree-CRFs

(d) Skip-Tree CRFs

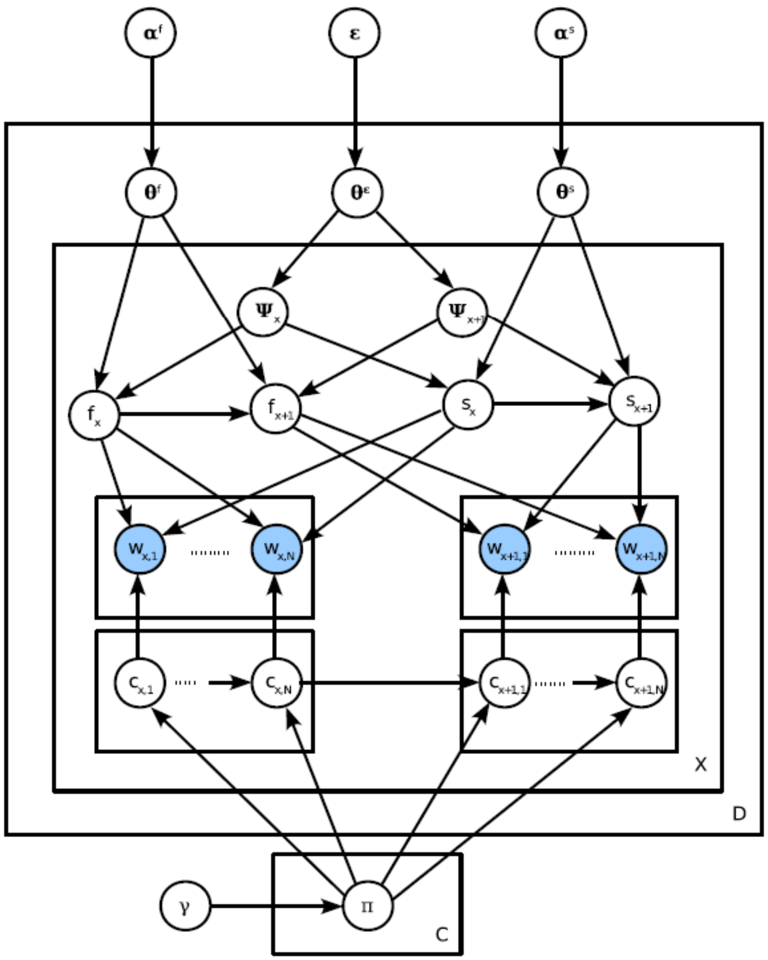
Supervised Learning Techniques

- CFACTS (Lakkaraju et al. 2011)
  - Base model: HMM
  - Task: discovering aspects, sentiments, their rating.
  - Novelty: capturing the syntactic dependencies between aspects and sentiments
  - Improvement: incorporating coherence in reviews
  - Rating of aspects are computed using a normal linear model



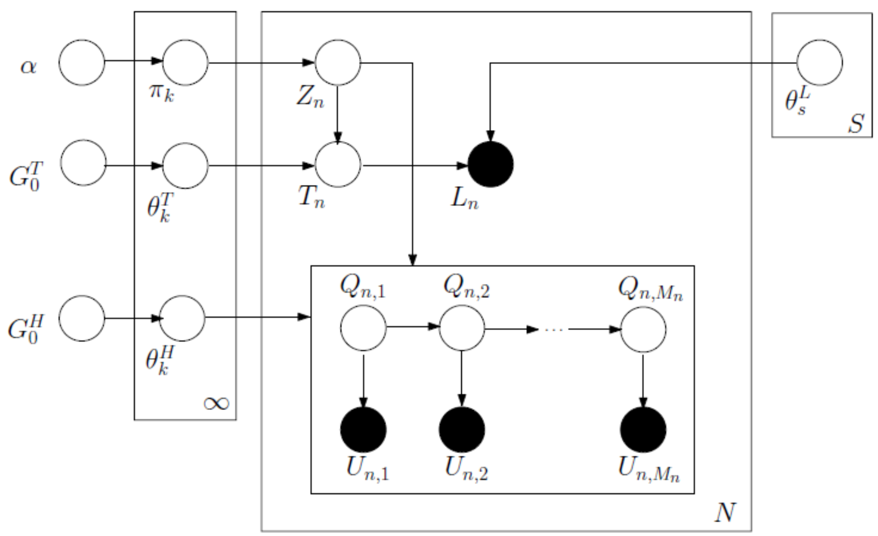
Supervised Learning Techniques

- CFACTS (Lakkaraju et al. 2011)
  - Base model: HMM
  - Task: discovering aspects, sentiments, their rating.
  - Novelty: capturing the syntactic dependencies between aspects and sentiments
  - Improvement: incorporating coherence in reviews
  - Rating of aspects are computed using a normal linear model



Supervised Learning Techniques

- Wong's model (Wong et al. 2008)
  - Base model: HMM
  - Task: extracting and grouping aspects from multiple websites
  - Novelty: using content and layout information



Supervised Learning Techniques

- Other methods
  - Kobayashi et al. (2007)
  - Jakob et al. (2010)
  - Choi et al. (2010)
  - Yu et al. (2011)
  - Sauper et al. (2011)

Supervised Learning Techniques

- Strength
  - Overcome frequency-based limitations by learning the model parameters from the data.
- Limitation
  - Need manually labeled data for training.

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
  - Supervised learning techniques
  - Topic modeling techniques
    - TSM
    - MG-LDA
    - JST
    - Sentiment-LDA
    - MaxEnt-LDA
    - ASUM
    - SDWP
    - LARAM
    - STM
    - ILDA
- Design Guidelines for LDA-based Models
- Conclusion and Future Directions

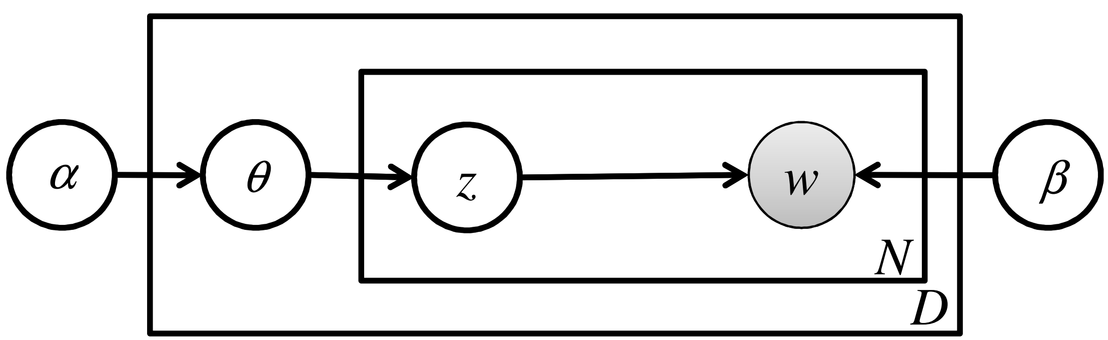
Topic-Modeling Techniques

- Extending the basic topic models to jointly model both aspects and sentiments

- Why
  - Intuitively, topics from topic models cover aspects in reviews
  - However, topics cover both aspects and sentiments

Topic-Modeling Techniques

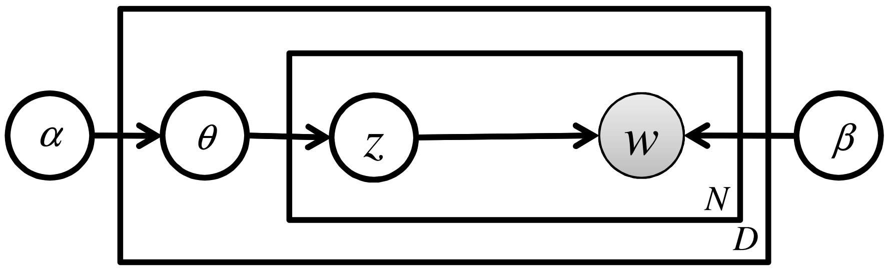
- Latent Dirichlet Allocation (LDA)(Blei et al. 2003)
  - Documents are represented as mixtures over latent topics where topics are associated with a distribution over the words of the vocabulary.



Topic-Modeling Techniques

- LDA generative process

1. Sample $\theta \sim Dir(\alpha)$
2. For each word $w_n,\ n \in \{1,2,\ldots,N\}$
   (a) Sample a topic $z_n \sim Mult(\theta)$
   (b) Sample a word $w_n \sim P(w_n \mid z_n, \beta)$



- Joint probability distribution

$$P(z,\mathbf{w},\theta \mid \alpha,\beta) = P(\theta \mid \alpha)\prod_{n=1}^{N}[P(z_n \mid \theta)P(w_n \mid z_n,\beta)]$$

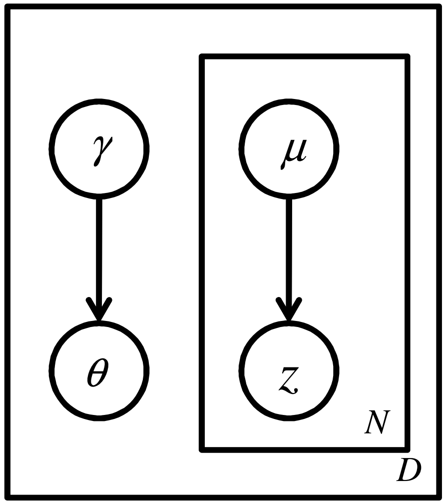
$$P(z,\theta \mid \mathbf{w},\alpha,\beta) = \frac{P(z,\mathbf{w},\theta \mid \alpha,\beta)}{P(\mathbf{w} \mid \alpha,\beta)}$$

Topic-Modeling Techniques

- Variational Inference

$$Q(z,\theta \mid \gamma,\mu) = Q(\theta \mid \gamma)\prod_{n=1}^{N} Q(z_n \mid \mu_n)$$

$$(\gamma^*,\mu^*) = \arg\min_{(\gamma,\mu)} KL(Q(z,\theta \mid \gamma,\mu) \parallel P(z,\theta \mid \mathbf{w},\alpha,\beta))$$



- Parameter Estimation

$$\ell(\alpha,\beta) = \sum_{d=1}^{D} \log P(\mathbf{w}_d \mid \alpha,\beta)$$

Topic-Modeling Techniques

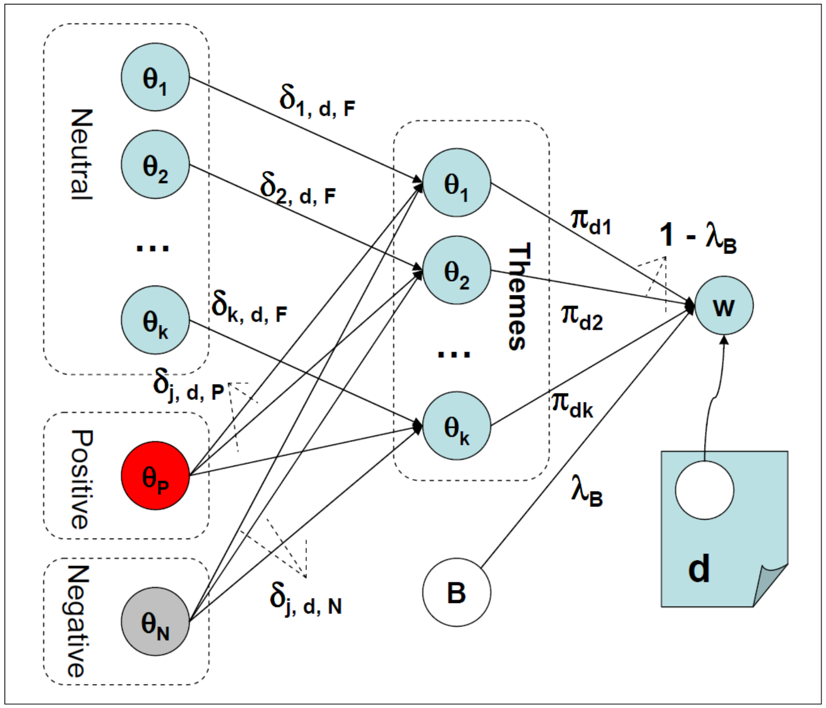
- Computational Complexity
  - Each iteration of variational inference for the basic LDA requires O(Nk) operations.

- Avoiding over-fitting
  - Smoothing (assigning positive probability to all vocabulary terms whether or not they are observed in the training set).

Topic-Modeling Techniques

SFU

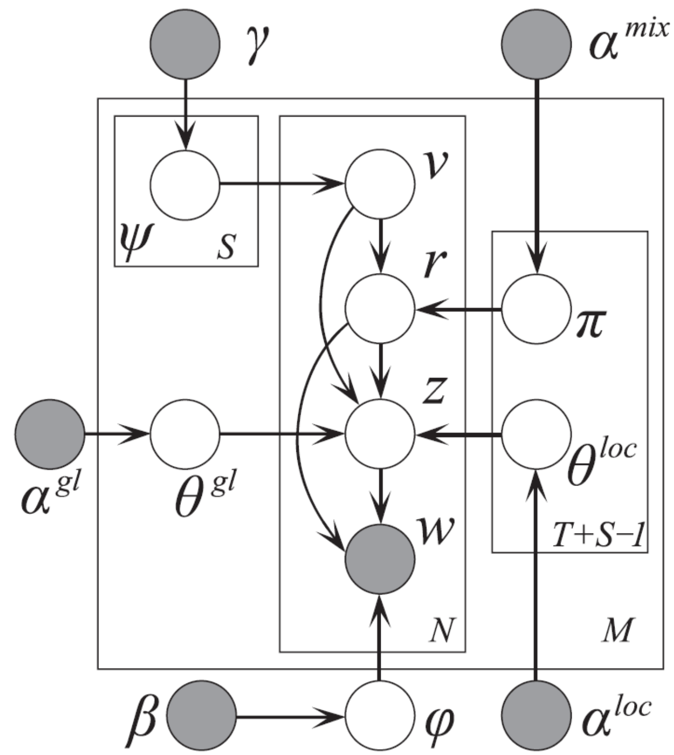
- TSM (Mei et al. 2007)
  - Task: identifying aspects and their polarity
  - Novelty: distribution of background words



Topic-Modeling Techniques

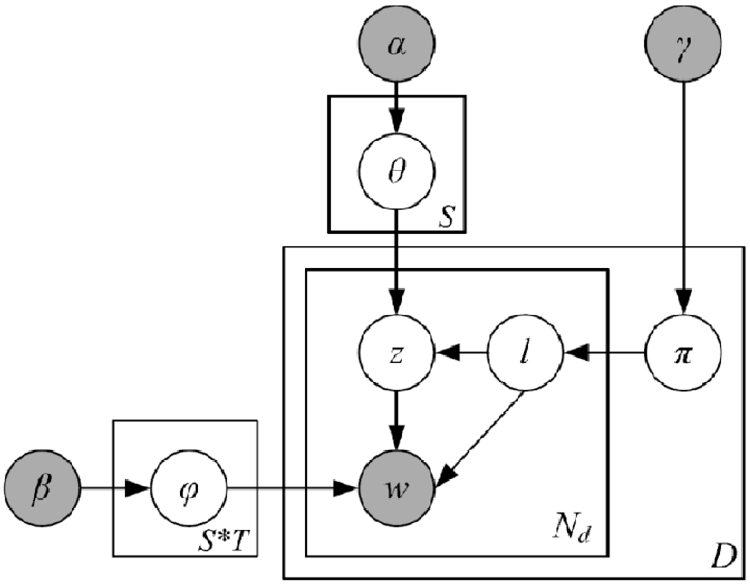
SFU

- MG-LDA (Titov et al. 2008a)
  - Task: extracting aspects
  - Novelty: considering global and local topics
  - Extension (Titov et al. 2008b)
    - Find correspondence between topics and aspects.



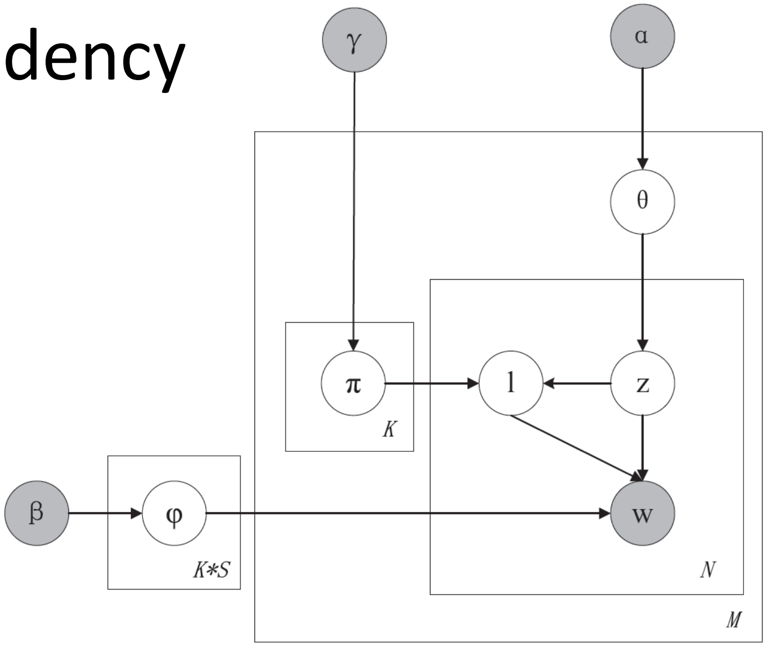
Topic-Modeling Techniques

- JST (Lin et al. 2009)
  - Task: identify aspect and their polarity
  - Novelty: considering different aspect distributions for each polarity
  - Extension (He et al. 2011)
    - Using word prior polarity for cross-domain extraction



Topic-Modeling Techniques

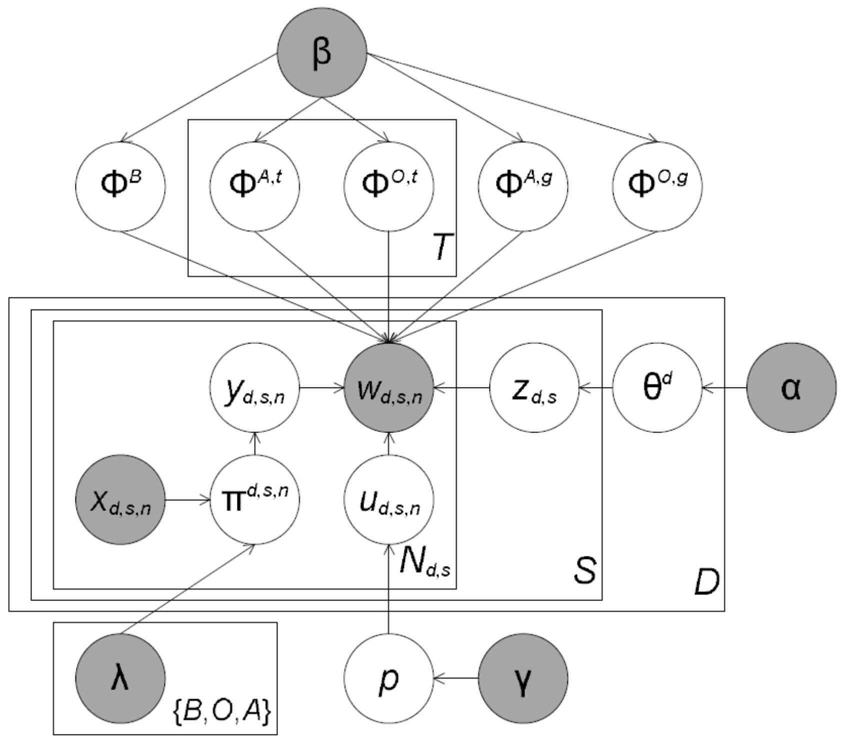
- Sentiment-LDA (Li et al. 2010)
  - Task: identifying aspects and their polarity
  - Novelty: considering different polarity distributions for each aspect
  - Improvement: consider dependency of polarity to local context.



Topic-Modeling Techniques

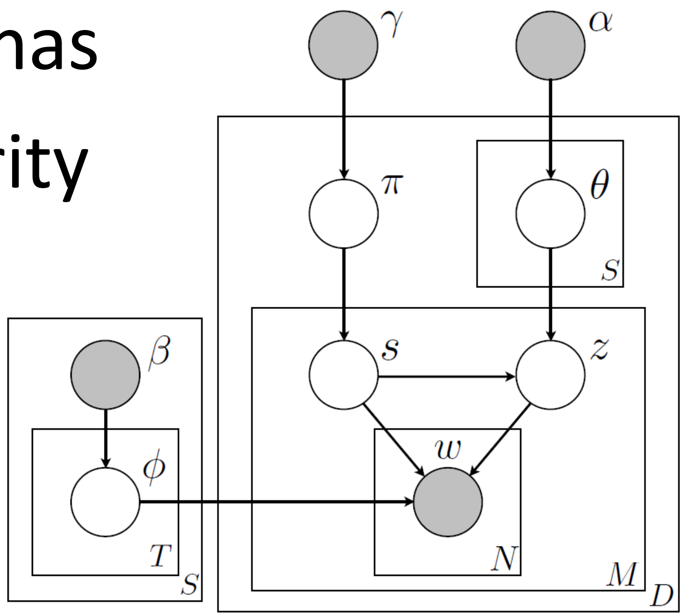
SFU

- MaxEnt-LDA (Zhao et al. 2010)
  - Task: identifying aspects and sentiments
  - Novelty: leverage POS tags of words to separate aspects, sentiments, and background words



Topic-Modeling Techniques

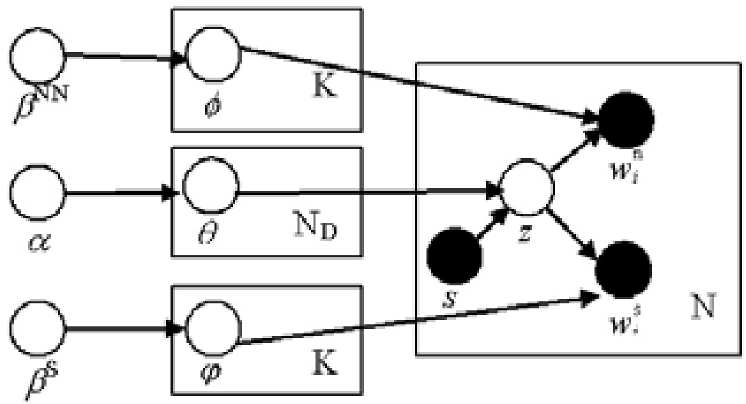
- ASUM (Jo et al. 2011)
  - Task: identifying aspects and their polarity
  - Novelty: extension of JST by a further assumption
  - Assumption: each sentence has one aspect and related polarity



Topic-Modeling Techniques

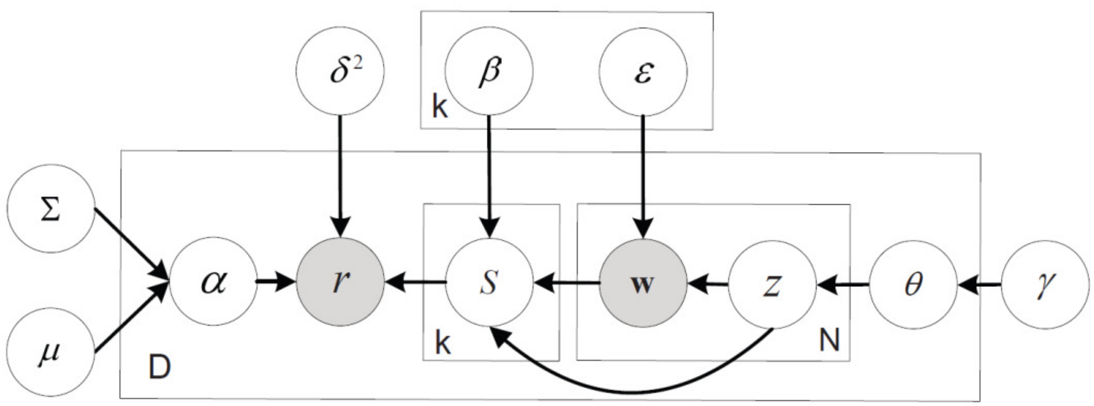
SFU

- SDWP (Zhan et al. 2011)
  - Task: identifying aspect and sentiments
  - Preprocessing: chunking reviews into opinion phrases
  - Novelty: modeling opinion phrase rather than all words



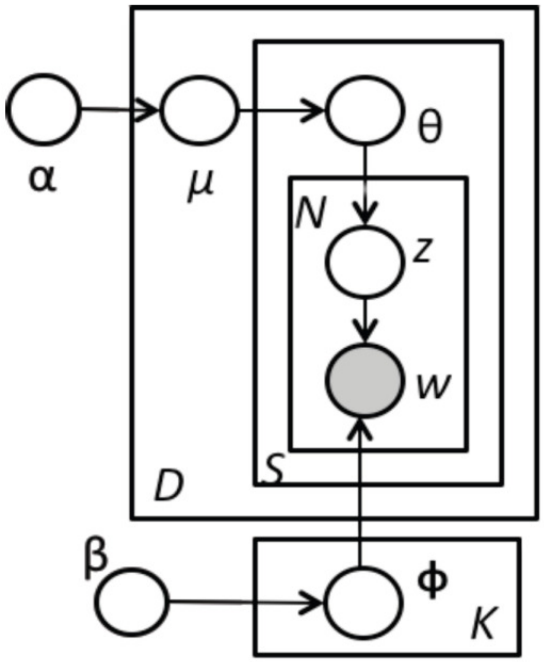
Topic-Modeling Techniques

- LARA (Wang et al. 2011)
  - Task: identifying aspects, ratings, and the weight placed on each aspect by the reviewer
  - Novelty: estimating the emphasis weight of aspects



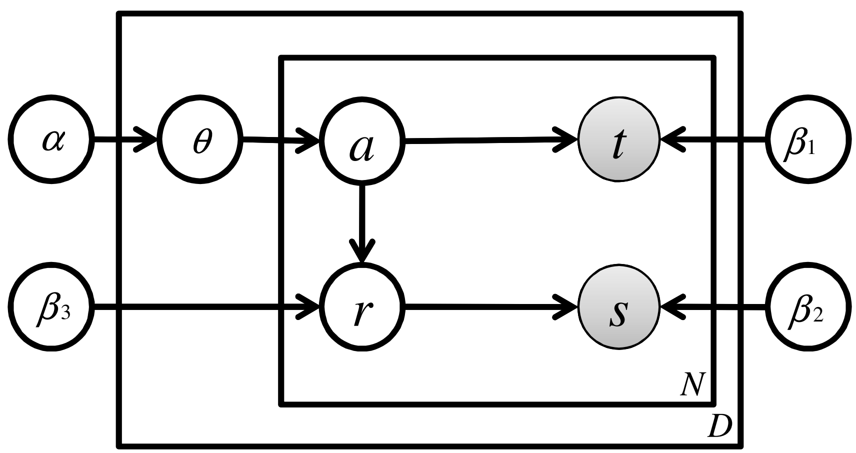
Topic-Modeling Techniques

- STM (Lu et al. 2011)
  - Task: identifying aspects and their rating
  - Novelty: jointly modeling document- and sentence-level topics
  - Rating prediction by training a regression model on overall rating



Topic-Modeling Techniques

- ILDA (Moghaddam et al. 2011)
  - Task: identifying aspects and rating simultaneously
  - Preprocessing: chunking reviews to opinion phrases
  - Novelty: rating of sentiments depends on aspects



Topic-Modeling Techniques

- Other models
  - Branavan et al. (2008)
  - Lu et al. (2009)
  - Brody et al. (2010)
  - Mukherjee et al. (2012)

Topic-Modeling Techniques

- Strengths
  - No need for manually labeled data
  - Perform both aspect extraction and grouping at the same time in an unsupervised manner
- Limitation
  - Need a large volume of data

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
- Design Guidelines for LDA-based Models
  - Abstract LDA-based Models
  - Comparison
- Conclusion and Future Directions

Design Guidelines for LDA-based Models

SFU

- State-of-the-art LDA models
  - A lot in common
  - Some differences correspond to design decisions

- Questions
  - Does a new model always outperform the existing ones?
  - Is there a "one-size-fits-all" model?

- Answer
  - Comparing state-of-the-art models
  - Present design guidelines for LDA-based models (Moghaddam et al. 2012(b))

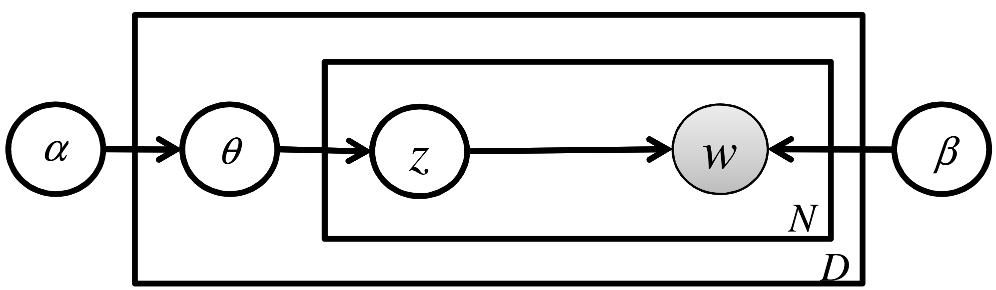
Design Guidelines for LDA-based Models

- Distinguishing Characteristics
  - Modeling words using one latent variable vs. having separate latent variables for aspects and ratings.
  - Modeling all words of the reviews vs. modeling only opinion phrases.
  - Modeling the dependency between aspects and ratings vs. modeling them independently.
  - Using only review texts as input vs. also using additional input data, e.g. a review’s overall rating.

Design Guidelines for LDA-based Models

- LDA: The basic LDA model which learns general topics of reviews using all words of the training reviews.
- Similar models: (Brody et al. 2010), (Zhao et al. 2010), (Titov et al. 2008)

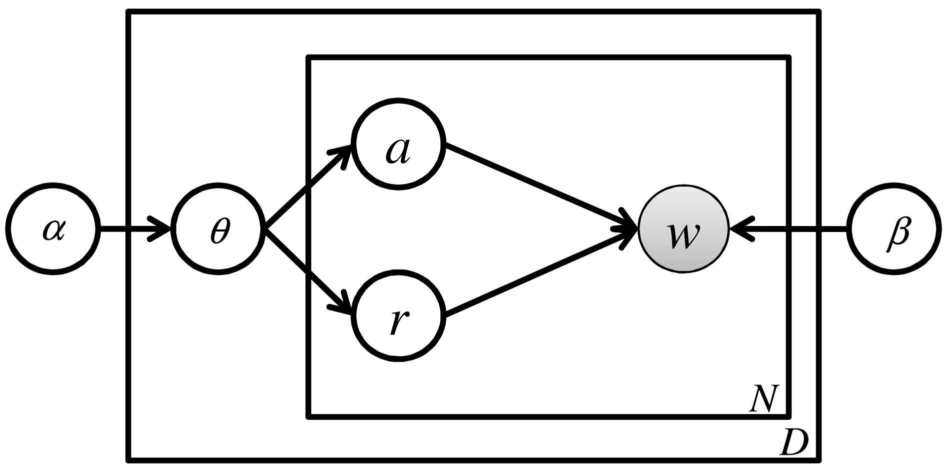
$$P(z, w, \theta \mid \alpha, \beta) = P(\theta \mid \alpha) \prod_{n=1}^{N} [P(a_n \mid \theta) P(w_n \mid z_n, \beta)]$$



Design Guidelines for LDA-based Models

- S-LDA: Extension of LDA where the model learns both aspects and ratings from reviews.
- Similar model: (Lakkaraju et al. 2011)

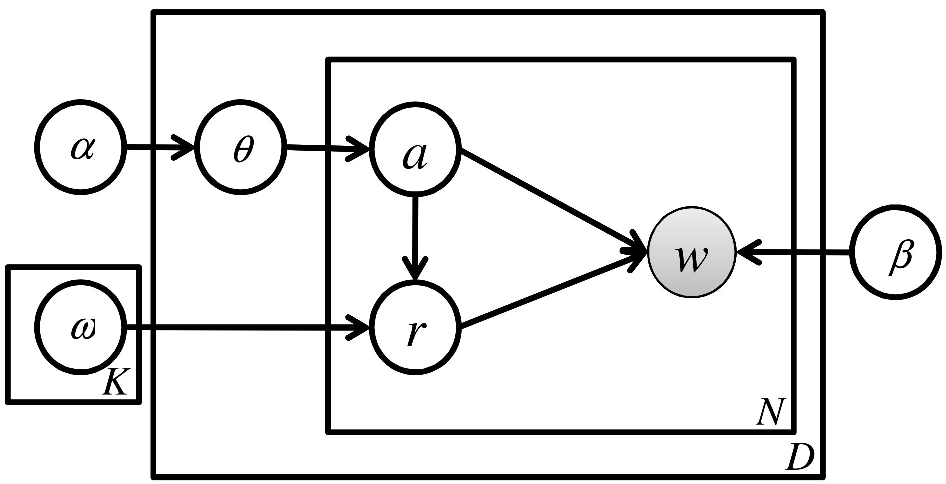
$$P(a,r,w,\theta \mid \alpha,\beta) = P(\theta \mid \alpha)\prod_{n=1}^{N}[P(a_n \mid \theta)P(r_n \mid \theta)P(w_n \mid a_n, r_n, \beta)]$$



Design Guidelines for LDA-based Models

- D-LDA: Extension of S-LDA considering the dependency between aspects and their ratings.
- Similar models: (Li et al. 2010), (Lin et al. 2009), (Wang et al. 2011), (Jo et al. 2011)

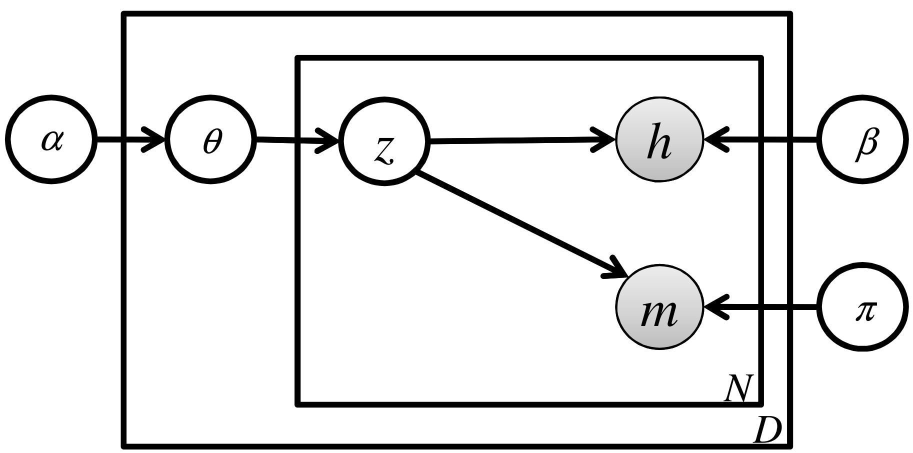
$$P(a, r, w, \theta \mid \alpha, \beta, \omega) = P(\theta \mid \alpha) \prod_{n=1}^{N} [P(a_n \mid \theta) P(r_n \mid a_n, \omega) P(w_n \mid a_n, r_n, \beta)]$$



Design Guidelines for LDA-based Models

- PLDA: The basic LDA model which learns general topics of reviews from opinion phrases.
- Similar model: (Zhan et al. 2011)

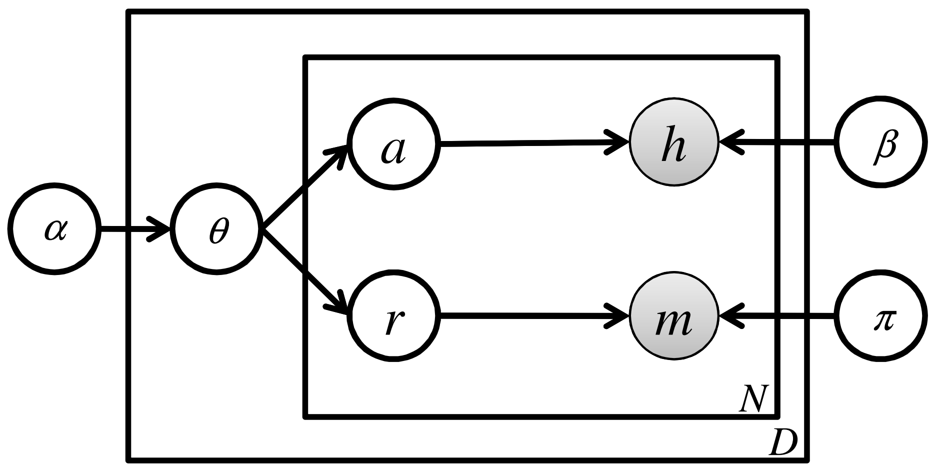
$$P(z,h,m,\theta \mid \alpha,\beta,\pi) = P(\theta \mid \alpha)\prod_{n=1}^{N}[P(z_n \mid \theta)P(h_n,m_n \mid z_n,\beta,\pi)]$$



Design Guidelines for LDA-based Models

- S-PLDA: An extension of PLDA where the model learns both aspects and ratings from phrases.
- Similar model: (Moghadam et al. 2011)

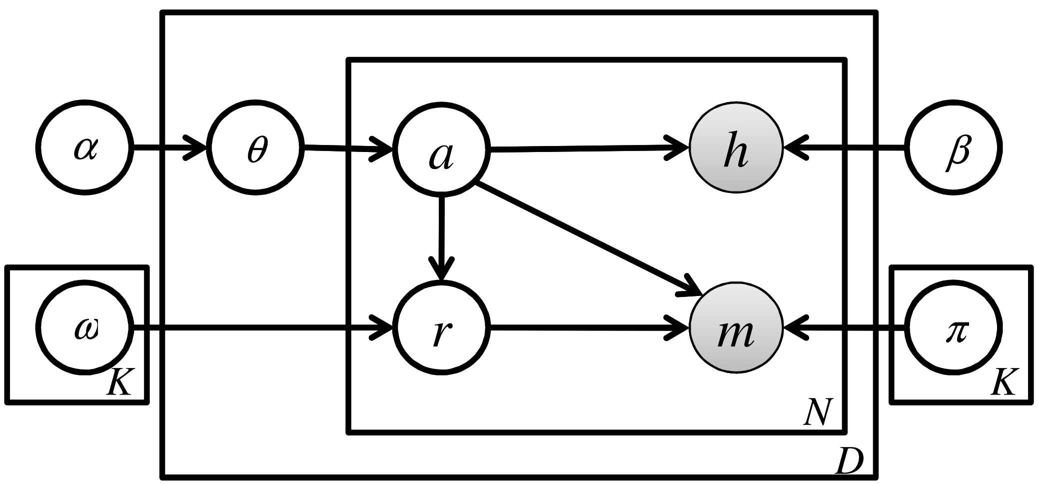
$$P(a,r,h,m,\theta \mid \alpha,\beta,\pi) = P(\theta \mid \alpha)\prod_{n=1}^{N}[P(a_n \mid \theta)P(r_n \mid \theta)P(h_n \mid a_n,\beta)P(m_n \mid r_n,\pi)]$$



Design Guidelines for LDA-based Models

- D-PLDA: An LDA-based model learning aspects and their corresponding ratings from opinion phrases while considering the dependency between the aspects and ratings.
- Similar model: (Moghaddam et al. 2011)

$$P(a,r,h,m,\theta \mid \alpha,\omega,\beta,\pi) = P(\theta \mid \alpha)\prod_{n=1}^{N}[P(a_n \mid \theta)P(r_n \mid a_n,\omega)P(h_n \mid a_n,\beta)P(m_n \mid a_n,r_n,\pi)]$$



Design Guidelines for LDA-based Models

SFU

- Extraction of opinion phrases
  - Frequency technique (e.g, Moghaddam et al. 2011)
    - Frequency of phrases
  - POS patterns (e.g., Zhao et al. 2010)
    - Syntactic relations, e.g., NN_VB_JJ → <NN, JJ>
  - Dependency patterns (e.g., Moghaddam et al. 2012(b))
    - Semantic relations, e.g., amod(NN, JJ) → <N, JJ>

Design Guidelines for LDA-based Models

- Dataset
  - Crawl Epinions.com
  - Publicly available at http://www.sfu.ca/~sam39/Datasets/

Subset	#Products	#Rev./Product
1<#Rev.<=10	36,166	3
10<#Rev.<=50	7,886	19
50<#Rev.<=100	869	67
100<#Rev.<=200	368	137
200<#Rev.	179	341

Design Guidelines for LDA-based Models

- Qualitative evaluation

Prep.	Model	Top words/phrases (stemmed)
N/A	LDA	good, pictur, digit, resolut, set, disk, great, time, shot, featur
	S-LDA	featur, zoom, disk, good, pictur, set, shot, resolut, camera, floppi
	D-LDA	bright, time, resolut, good, disk, set, great, digit, shot, camera
Freq. nouns	PLDA	<pictur,mani>,<resolut,high>,<time,hard>,<camera,digit>,<disk,floppi>
	S-PLDA	<time,hard>,<drive,floppi>,<resolut,mani>,<camera,digit>,<featur,good>
	D-PLDA	<resolut,high>,<camera,digit>,<drive,floppi>,<usb,easi>,<disk,hard>
POS patrn	PLDA	<usag,normal>,<price,high>,<drive,floppi>,<pictur,good>,<featur,sever>
	S-PLDA	<pictur,good>,<effect,special>,<displai,nice>,<life,long>,<printer,great>
	D-PLDA	<batteri,dead>,<resolut,mani>,<camera,digit>,<pictur,good>,<featur,offer>
Dep. Patrn	PLDA	<displai,nice>,<zoom,optic>,<effect,mani>,<pictur,good>,<reolut,high>
	S-PLDA	<resolut,high>,<qualiti,amaz>,<printer,compat>,<displai,nice>,<zoom,optic>
	D-PLDA	<photo qualiti,high><zoom,optic><storag capac,unlimit><printer,compat><viewfind,love>

Design Guidelines for LDA-based Models

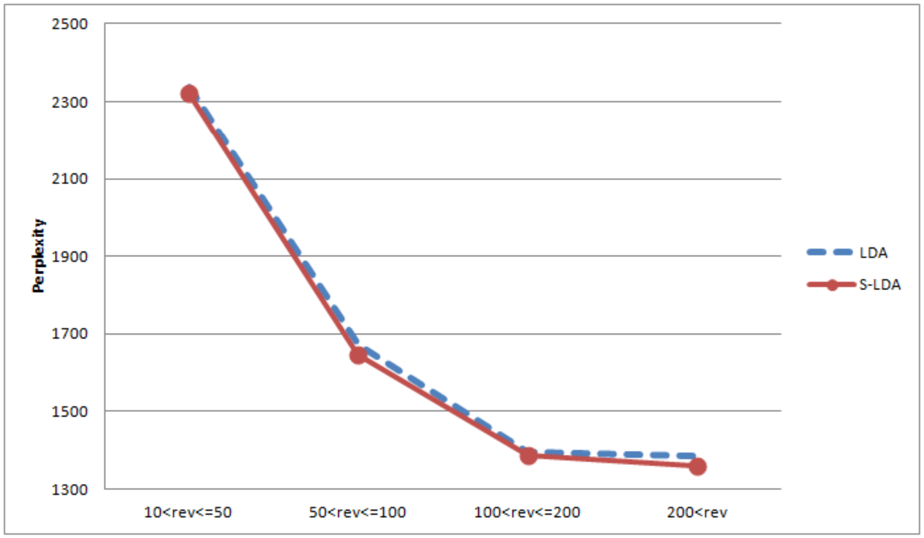
- Quantitative evaluation
  - Precision, Recall and MSE
    - Public data set with ground truth (Hu et al. 2004)
  - Perplexity

Design Guidelines for LDA-based Models

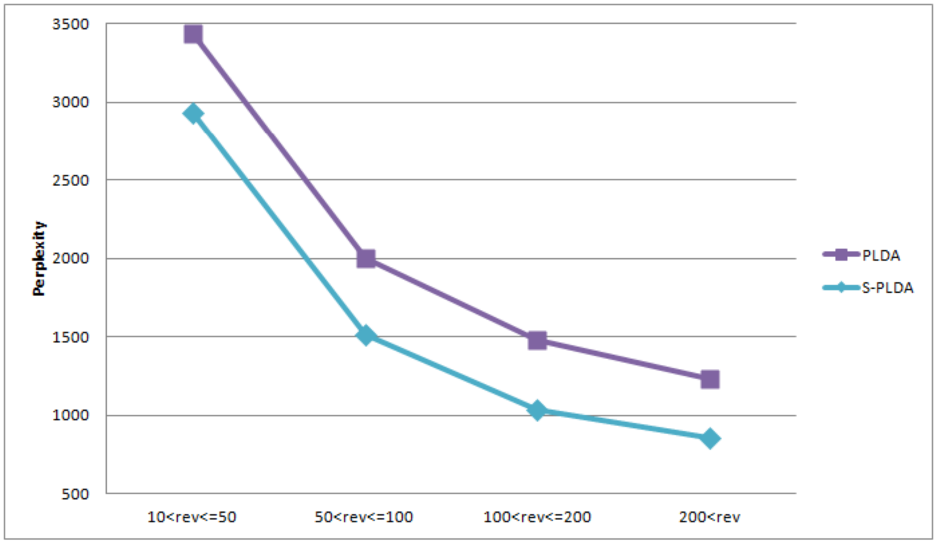
SFU

- Is it better to have separate latent variables for aspects and ratings?

LDA vs. S-LDA



PLDA vs. S-PLDA



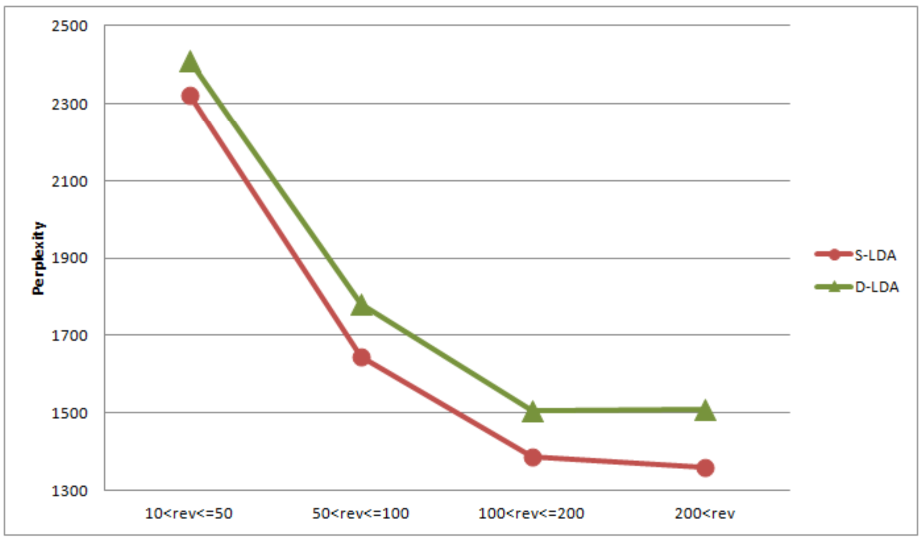
- Having separate aspect and rating helps when learning from opinion phrases.

Design Guidelines for LDA-based Models

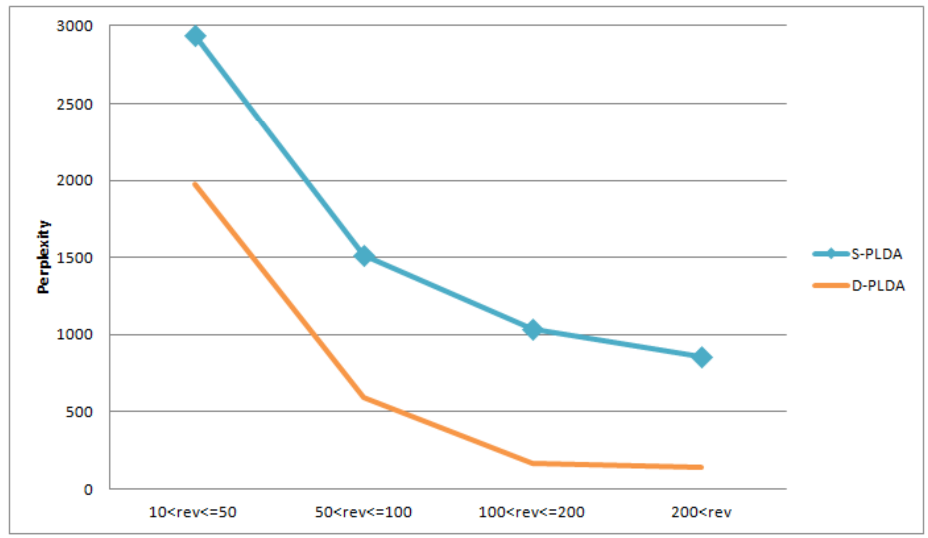
SFU

- Is it better to assume dependency between ratings and aspects?

S-LDA vs. D-LDA



S-PLDA vs. D-PLDA



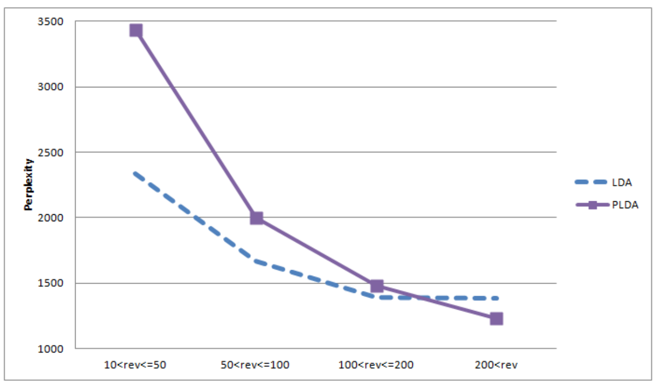
- Assuming this dependency helps when learning from opinion phrases and having separate aspect and rating.

Design Guidelines for LDA-based Models

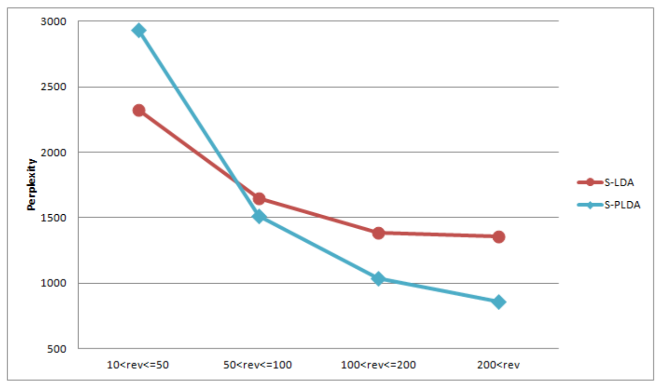
SFU

- Is it better to learn from bag-of-words or preprocess the reviews and learn from opinion phrases?

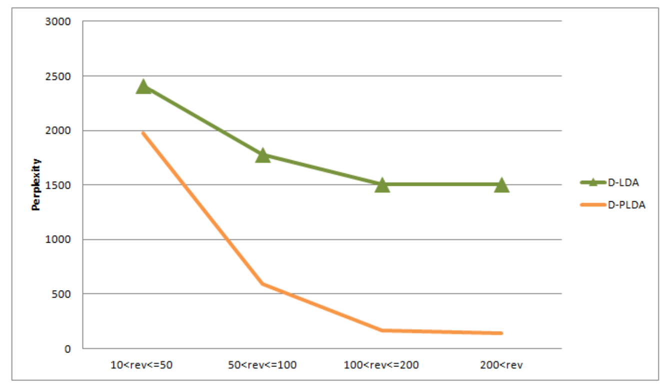
LDA vs. PLDA



S-LDA vs. S-PLDA



D-LDA vs. D-PLDA



- Using a preprocessing technique helps when separate latent variables are assumed for aspects and ratings.

Design Guidelines for LDA-based Models

- Which preprocessing technique for extracting opinion phrases works best?
  - POS technique is more effective than the frequency approach.
  - Dependency technique outperforms the other preprocessing techniques for all subsets of products.

- Using dependency patterns consistently achieves the best performance for extracting opinion phrases.

Design Guidelines for LDA-based Models

SFU

- Does the answer to the above questions differ for products with few reviews and products with many reviews?
  - For product with few reviews neither preprocessing nor model complexity could help

- For products with few reviews, the basic LDA model outperforms the other models. For other products the model learning aspects and ratings from opinion phrases with dependency assumption performs best.

Outline

- Introduction
- General Opinion Mining Tasks
- Aspect-based Opinion Mining
- Frequency and Relation based Approaches
- Model-based Approaches
- Design Guideline for LDA-based Models
- Conclusion and Future Directions
  - Summary
  - Future Research Direction

Summary

- Introduction to Opinion Mining
  - Opinion mining is already being successfully applied in industry
  - Lots of challenging research issues
  - Case study: Google Products

Summary

- General opinion mining tasks
  - Subjectivity Classification
  - Sentiment Classification
  - Opinion Helpfulness Prediction
  - Opinion Spam Detection
  - Opinion Summarization
  - Mining Comparative Opinions

Summary

- Aspect-based opinion mining
  - Problem Definition
  - Challenges
  - Evaluation Metrics
  - Benchmark Datasets

- Frequency and relation based approaches
  - Frequency-based
  - Relation-based
  - Hybrid methods

Summary

- Model-based approaches
  - Supervised learning
    - CRF & HMM
  - Topic modeling
    - PLSA and LDA

- Design guidelines for LDA-based models
  - Abstract LDA-based models
  - Comparison

Future Research Directions

- Better evaluation techniques
- Dealing with noun and verb sentiments
- Discovery of implicit aspects
- Coreference resolution for sentiment identification

Future Research Directions

- Entity, opinion holder, and time extraction
- Discovery of comparative opinions
- Dealing with "noisy" input texts such as discussion forums
- Better methods for extracting opinion phrases

Future Research Directions

- Opinion mining for entities with few reviews
- Exploring the impact of different additional input sources
- Using aspect-based OM in other applications

SFU

Thank You!

References

- Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. Multi-facet rating of product reviews. In ECIR '09.
- Bethard, S., H. Yu, A. Thornton, V. Hatzivassiloglou, and D. Jurafsky. Automatic extraction of opinion propositions and their holders. In AAAI 2004.
- Bhattarai, Archana, Nobal Niraula, Vasile Rus, and King-Ip Lin. A domain independent framework to extract and aggregate analogous features in online reviews. In CICLing'12.
- Blei, David, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. J. Mach. Learn. Res., 2003.
- Blair-Goldensohn, Sasha, Kerry Hannan, Ryan McDonald, Tyler Neylon, George A. Reis, and Jeff Reynar. Building a sentiment summarizer for local service reviews. In Proceedings of WWW-2008 workshop on NLP in the Information Explosion Era. 2008.
- Branavan, S. R. K., Harr Chen, Jacob Eisenstein, and Regina Barzilay. Learning document-level semantic properties from free-text annotations. In ACL-2008.
- Brody, Samuel, and Noemie Elhadad. An unsupervised aspect-sentiment model for online reviews. In HLT '10.
- Choi, Yejin and Claire Cardie. Hierarchical sequential learning for extracting opinions and their attributes. In ACL-2010.
- Dasgupta, S. and V. Ng. Mine the easy, classify the hard: a semi-supervised approach to automatic sentiment classification. In ACL-2009, 2009.
- Ding, Xiaowen, Bing Liu, and Philip S. Yu. A holistic lexicon-based approach to opinion mining. In Proceedings of the Conference on Web Search and Web Data Mining (WSDM-2008). 2008.

References

- Du, Weifu, and Songbo Tan. An iterative reinforcement approach for ne-grained opinion mining. In NAACL '09.
- Ganapathibhotla, M. and B. Liu. Mining opinions in comparative sentences. In Coling-2008, 2008.
- Ghani, Rayid, Katharina Probst, Yan Liu, Marko Krema, and Andrew Fano. Text mining for product attribute extraction. ACM SIGKDD Explorations Newsletter, 2006. 8(1): pp. 41–48.
- Ghose, A. and P. Ipeirotis. Designing novel review ranking systems: predicting the usefulness and impact of reviews. In Proceedings of the International Conference on Electronic Commerce 2007.
- Guo, Honglei, Huijia Zhu, Zhili Guo, XiaoXun Zhang, and Zhong Su. Product feature categorization with multilevel latent semantic association. In CIKM '09.
- HongningWang, Yue Lu, and ChengXiang Zhai. Latent aspect rating analysis without aspect keyword supervision. In KDD '11.
- Hai, Zhen, Kuiyu Chang, and Jung-jae Kim. Implicit feature identication via cooccurrence association rule mining. In CICLing '11.
- He, Yulan, Chenghua Lin, and Harith Alani. Automatically extracting polarity-bearing topics for cross-domain sentiment classiffication. In HLT '11.
- Hu, Minqing and Bing Liu. Mining and summarizing customer reviews. In KDD-2004.
- Jakob, Niklas and Iryna Gurevych. Extracting Opinion Targets in a Single-and Cross-Domain Setting with Conditional Random Fields. In EMNLP-2010.
- Jiang, Long, Mo Yu, Ming Zhou, Xiaohua Liu, and Tiejun Zhao. Target-dependent twitter sentiment classification. In ACL-2011.

References

- Jiang, Peng, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang. An approach based on tree kernels for opinion mining of online product reviews. In ICDM '10.
- Jin, Wei, Hung Hay Ho, and Rohini K. Srihari. Opinionminer: a novel machine learning system for web opinion mining and extraction. In KDD '09.
- Jindal, N. and B. Liu. Opinion spam and analysis. In WSDM-2008, 2008.
- Jindal, N. and B. Liu. Review spam detection. In Proceedings of WWW 2007.
- Jo, Yohan,and Alice H. Oh. Aspect and sentiment unffication model for online review analysis. In WSDM '11.
- Joshi, Mahesh, and Carolyn Penstein-Ros\&\#233;. 2009. Generalizing dependency features for opinion mining. In Proceedings of the ACL-IJCNLP 2009 Conference Short Papers (ACLShort '09).
- Kaplan A., Michael Haenlein. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53(1): 59–68, 2010.
- Kim, S., P. Pantel, T. Chklovski, and M. Pennacchiotti. Automatically assessing review helpfulness. In EMNLP-2006, 2006.
- Kim, J., J.J. Li, and J.H. Lee. Discovering the discriminative views: Measuring term weights for sentiment analysis. In ACL-2009, 2009.
- Kim, S. and E. Hovy. Determining the sentiment of opinions. In COLING-2004, 2004.
- Kobayashi, Nozomi, Kentaro Inui, and Yuji Matsumoto. Extracting aspect-evaluation and aspect-of relations in opinion mining. In EMNLP-CoNLL 2007.
- Kobayashi, Nozomi, Ryu Iida, Kentaro Inui, and Yuji Matsumoto. Opinion mining on the Web by extracting subject-attribute-value relations. In Proceedings of AAAI-CAA W'06. 2006.

References

- Ku, Lun-Wei, Yu-Ting Liang, and Hsin-Hsi Chen. Opinion extraction, summarization and tracking in news and blog corpora. In Proceedings of AAAI-CAA W'06. 2006.
- Lakkaraju, Himabindu, Chiranjib Bhattacharyya, Indrajit Bhattacharya, and Srujana Merugu. Exploiting coherence for the simultaneous discovery of latent facets and associated sentiments. In SDM '11.
- Li, Fangtao, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang, and Hao Yu. Structure-aware review mining and summarization. In COLING '10.
- Li, Fangtao, Minlie Huang, and Xiaoyan Zhu. Sentiment analysis with global topics and local dependency. In AAAI '10.
- Li, Zhichao, Min Zhang, Shaoping Ma, Bo Zhou, and Yu Sun. 2009. Automatic Extraction for Product Feature Words from Comments on the Web. In AIRS '09.
- Lin, Chenghua, and Yulan He. Joint sentiment/topic model for sentiment analysis. In CIKM '09.
- Liu, Bing. Sentiment Analysis and Opinion Mining, Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers, 2012.
- Liu, Bing. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, chapter 11. Springer, 2007.
- Liu, Bing, Minqing Hu, and Junsheng Cheng. Opinion observer: Analyzing and comparing opinions on the web. In Proceedings of International Conference on World Wide Web (WWW-2005). 2005.
- Liu, Bing. Sentiment Analysis, tutorial at AAAI-2011. Lerman, K., S. Blair-Goldensohn, and R. McDonald. Sentiment summarization: Evaluating and learning user preferences, 2009. EACL-2009.

References

- Liu, J., Y. Cao, C. Lin, Y. Huang, and M. Zhou. Low-quality product review detection in opinion summarization. In EMNLP-CoNLL-2007, 2007.
- Liu, J., S. Seneff, and V. Zue. Dialogue-oriented review summary generation for spoken dialogue recommendation systems. In Proceedings of HAACL-2010, 2010.
- Liu, Yang, Xiangji Huang, Aijun An, and Xiaohui Yu. ARSA: a sentiment-aware model for predicting sales performance using blogs. In SIGIR-2007.
- Long, Chong, Jie Zhang, and Xiaoyan Zhu. A review selection approach for accurate feature rating estimation. In Proceedings of Coling 2010: Poster Volume. 2010.
- Lu, Yue, ChengXiang Zhai, and Neel Sundaresan. Rated aspect summarization of short comments. In WWW '09.
- Lu, Bin, Myle Ott, Claire Cardie, and Benjamin K. Tsou. 2011. Multi-aspect Sentiment Analysis with Topic Models. In ICDMW '11.
- Lu, Yue and ChengXiang Zhai. Opinion integration through semi-supervised topic modeling. In WWW-2008.
- Mei, Qiaozhu Xu Ling, Matthew Wondra, Hang Su, and ChengXiang Zhai. Topic sentiment mixture: modeling facets and opinions in weblogs. In WWW '07.
- Moghaddam, samaneh and Martin Ester. Opinion digger: an unsupervised opinion miner from unstructured product reviews. In CIKM '10.
- Moghaddam, samaneh, and M. Ester. ILDA: Interdependent LDA model for learning latent aspects and their ratings from online product reviews. In SIGIR '11.

References

- Moghaddam, samaneh, Mohsen Jamali and Martin Ester. ETF: Extended Tensor Factorization Model for Personalizing Prediction of Review Helpfulness. In WSDM '12.
- Moghaddam, samaneh and M. Ester. On the Design of LDA Models for Aspect-based Opinion Mining. In CIKM '12.
- Mullen, T. and N. Collier. Sentiment analysis using support vector machines with diverse information sources. In EMNLP-2004, 2004.
- Paltoglou, G. and M. Thelwall. A study of information retrieval weighting schemes for sentiment analysis. ACL ’10.
- Pang, B., L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In EMNLP-2002, 2002.
- Popescu, Ana-Maria and Oren Etzioni. Extracting product features and opinions from reviews. In EMNLP-2005.
- Qiu, Guang, Bing Liu, Jiajun Bu, and Chun Chen. Opinion Word Expansion and Target Extraction through Double Propagation. Computational Linguistics, Vol. 37, No. 1: 9.27, 2011.
- Raju, Santosh, Prasad Pingali, and Vasudeva Varma. 2009. An Unsupervised Approach to Product Attribute Extraction. In ECIR '09.
- Riloff, E. and J. Wiebe. Learning extraction patterns for subjective expressions. In EMNLP-2003, 2003.
- Sauper, Christina, Aria Haghighi, and Regina Barzilay. Content models with attitude. In ACL-2011.

References

- Scaffidi, Christopher, Kevin Bierhoff, Eric Chang, Mikhael Felker, Herman Ng, and Chun Jin. Red Opal: product-feature scoring from reviews. In Proceedings of 12th ACM Conference on Electronic Commerce (EC), 2007.
- Somasundaran, Swapna and Janyce Wiebe. Recognizing stances in online debates. In Proceedings of the 47th Annual Meeting of the AC L and the 4th IJCNLP of the AFNLP (AC L-IJCNLP-2009). 2009.
- Titov, Ivan, and Ryan McDonald. Modeling online reviews with multi-grain topic models. In WWW '08.
- Titov, Ivan, and Ryan McDonald. A joint model of text and aspect ratings for sentiment summarization. In ACL-HLT '08.
- Turney, P. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In ACL-2002, 2002.
- Wang, Hongning, Yue Lu, and Chengxiang Zhai. Latent aspect rating analysis on review text data: a rating regression approach. In KDD '10.
- Wei, Wei and Jon Atle Gulla. Sentiment learning on product reviews via sentiment ontology tree. In Proceedings of Annual Meeting of the Association for Computational Linguistics (AC L-2010). 2010.
- Wong, Tak-Lam, Wai Lam, and Tik-Shun Wong. An unsupervised framework for extracting and normalizing product attributes from multiple web sites. In SIGIR '08.
- Wong, Tak-Lam, Lidong Bing, and Wai Lam. Normalizing web product attributes and discovering domain ontology with minimal effort. In WSDM '11.
- Yessenalina, A., Y. Yue, and C. Cardie. Multi-level Structured Models for Document-level Sentiment Classification. In EMNLP, 2010.

References

- Yu, Jianxing, Zheng-Jun Zha, Meng Wang, and Tat-Seng Chua. Aspect ranking: identifying important product aspects from online consumer reviews. In HLT '11.
- Yu, Jianxing, Zheng-Jun Zha, Meng Wang, Kai Wang, and Tat-Seng Chua. Domain-Assisted Product Aspect Hierarchy Generation: Towards Hierarchical Organization of Unstructured Consumer Reviews. In EMNLP-2011.
- Zhan, Tian-Jie, and Chun-Hung Li. Semantic dependent word pairs generative model for ne-grained product feature mining. In PAKDD '11.
- Zhang, Z. and B. Varadarajan. Utility scoring of product reviews. In CIKM-2006, 2006.
- Zhao, Wayne Xin Jing Jiang, Hongfei Yan, and Xiaoming Li. Jointly modeling aspects and opinions with a maxent-lda hybrid. In EMNLP '10.
- Zhao, Yanyan, Bing Qin, Shen Hu, and Ting Liu. Generalizing syntactic structures for product attribute candidate extraction. In HLT '10.
- Zhu, Jingbo, Huizhen Wang, Benjamin K. Tsou, and Muhua Zhu. Multi-aspect opinion polling from textual reviews. In CIKM-2009.
- Zhuang, Li, Feng Jing, and Xiaoyan Zhu. Movie review mining and summarization. In Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2006). 2006.