Opinion Mining in Online Reviews: Recent Trends

Samaneh Moghaddam & Martin Ester
Simon Fraser University
Tutorial at WWW2013
May 14th 2013
Outline

• Introduction
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
• Guidelines for Designing LDA-based Models
• Conclusion and Future Directions
Outline

• Introduction
  – Social media
  – Online reviews
  – Opinion mining

• Opinion Mining Tasks

• Applications

• Aspect-based Opinion Mining

• Frequency and Relation based Approaches

• Model-based Approaches

• Guidelines for Designing 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
  – Reviews of products
  – Products of services

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
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

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.


• *Sentiment*: often used as synonym of opinion.
Opinion Mining

• 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.

Reviewers

What is the best digital camera?
Do people like camera X? or dislike it?

- Poor picture quality
- Disappointing battery life

- The batteries are great
- It is a little expensive

- Lovely picture quality
- The battery life is OK
Opinion Mining

• 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)
Outline

• Introduction
• Opinion Mining Tasks
  – Definitions
  – Sentiment Lexicons
  – Subjectivity Classification
  – Sentiment Classification
  – Opinion Helpfulness Prediction
  – Opinion Spam Detection
  – Opinion Summarization
  – Mining Comparative Opinions
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
• Guidelines for Designing 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 \((e_j, a_{jk}, s_{ijkl}, h_i, t_l)\) where
  
  – \(e_j\) is a target entity,
  – \(a_{jk}\) is an aspect of the entity \(e_j\),
  – \(h_i\) is an opinion holder,
  – \(t_l\) is the time when the opinion is expressed, and
  – \(s_{ijkl}\) is the sentiment orientation of opinion holder \(h_i\) on feature \(a_{jk}\) of entity \(e_j\) at time \(t_l\).
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 \((e_j, f_jk, s_{ijkl}, h_i, t_l)\).
- Also simpler versions of the problem.
- Using the quintuples, unstructured text \(\rightarrow\) 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

**Input**

<table>
<thead>
<tr>
<th>Canon GL2 Mini DV Camcorder</th>
<th>Overall Rating: ★★★☆☆</th>
<th>Full Review:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of Use</td>
<td>★★☆☆☆</td>
<td>... it has great zoom feature ...</td>
</tr>
<tr>
<td>Durability</td>
<td>★☆☆☆☆</td>
<td>the sound is terrible ...</td>
</tr>
<tr>
<td>Battery Life</td>
<td>★☆☆☆☆</td>
<td>the screen is blurry ...</td>
</tr>
<tr>
<td>Movie quality</td>
<td>★☆☆☆☆</td>
<td>with the affordable price ...</td>
</tr>
</tbody>
</table>

**Output**

<table>
<thead>
<tr>
<th>Canon GL2 Mini DV Camcorder</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom</td>
<td>★☆☆☆☆</td>
</tr>
<tr>
<td>Sound</td>
<td>★☆☆☆☆</td>
</tr>
<tr>
<td>Screen</td>
<td>★☆☆☆☆</td>
</tr>
<tr>
<td>Price</td>
<td>★☆☆☆☆</td>
</tr>
<tr>
<td>Size</td>
<td>★☆☆☆☆</td>
</tr>
</tbody>
</table>
Relationships Among OM Tasks

- Subjectivity Classification
- Sentiment Analysis
- Opinion Search and Retrieval
- Opinion Question Answering
- Opinion summarization
- Opinion Spam Detection
- Aspect-based Opinion Mining
- Opinion Helpfulness Est.
- OM in Comparative sentences
Sentiment Lexicons

- Crucial resources representing linguistic knowledge for opinion mining. Entries of words with sentiment orientation.
- Public domain sentiment lexicons
  - SentiWordNet http://sentiwordnet.isti.cnr.it/
  - Bing Liu’s sentiment lexicon
    http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
  - Emotion lexicon
    http://www.umiacs.umd.edu/~saif/WebPages/ResearchInterests.html#SemanticOrientation
Sentiment Lexicons

- SentiWordNet
  for every word: probability distribution over \{positive, negative, objective\}
- Bing Liu’s sentiment lexicon
  list of positive and negative opinion / sentiment words
- Emotion lexicon
  whether a word is positive or negative, and whether
  word has associations with basic emotions (joy, sadness, anger, fear, surprise, anticipation, trust, disgust)
- Manual or automatic acquisition.

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Sentiment Lexicons
(Kamps et al., 2004)

- Dictionary-based acquisition
- Start with small set of seed sentiment words.
- Expand seed set using WordNet’s synonyms and antonyms.
- Compute semantic orientation of term \( t \):

\[
SO(t) = \frac{d(t, "bad") - d(t, "good")}{d("good", "bad")}
\]

where \( d(t1, t2) \) is the length of the shortest-path between terms \( t1 \) and \( t2 \) in WordNet.

→ Disadvantage: lexicon is domain-independent.
Sentiment Lexicons
(Hatzilvassiloglou et al., 1997)

• Corpus-based acquisition

• Define set of linguistic connectors, e.g. and, or, neither-nor, either-or.

• Construct graph of adjectives that are connected to each other in the corpus through one of these connectors.

• Connected words are assumed to have similar sentiment orientation (sentiment consistency).

→ Advantage: lexicon is domain-specific.
Sentiment Lexicons
(Velikovich et al., 2010)

• Corpus-based acquisition, alternative approach
• Define co-occurrence vector of word: number of co-occurrences in corpus for given vocabulary.
• Define similarity between two words as the cosine similarity of their co-occurrence vectors.
• Construct similarity graph of words.
• Explore graph from positive seed words to compute strength of positive sentiment, and from negative seed words to compute strength of negative sentiment.
• Aggregate overall sentiment orientation.
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.

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
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.
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 words can be obtained from sentiment lexicon.
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

\[
PMI(phrase_1, phrase_2) = \log_2 \left( \frac{P(phrase_1 \wedge phrase_2)}{P(phrase_1)P(phrase_2)} \right)
\]
Sentiment Classification
(Turney 2002)

Semantic orientation (SO):

\[
SO(\text{phrase}) = PMI(\text{phrase}, "excellent") - PMI(\text{phrase}, "poor")
\]

Use near operator of a search engine to find the number of hits to compute PMI.

• 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 (from sentiment lexicon)
  - 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

- 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

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

(Zhang and Varadarajan, 2006; Kim et al. 2006; Ghose and Ipeirotis 2007; Liu et al 2007)
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

• 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

• 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
• Opinion Mining Tasks
• Applications
  – tweetfeel
  – The Stock Sonar
  – Google Products
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
• Guidelines for Designing LDA-based Models
• Conclusion and Future Directions
Application: tweetfeel

• Twitter and Facebook
  – Target of many opinion mining applications
• Monitoring opinions on a brand, politician, etc.
  – most common application
• tweetfeel
  – real-time analysis of tweets that contain a given term
• Main opinion mining task
  – sentiment classification of collection of tweets
Application: tweetfeel

Tweetfeel

blackberry

Search

Try some Twitter trends: May the Fourth Iron Man 3 Game 7 Titanic BBQ

211 216 = 51%

Those are all the results available right now. Try again or try another term to see how people feel towards it.

Got questions? Read our FAQ.

@ShyGuy101 ohhh lol...well that's entirely up to u :) bt i dont like blackberry bt i cant live witout it!..i hate my fone it annoys me!

i dont like blackberry bc shit but i dont wanna get an iphone bc on blackberry u dont need wifi for twitter wHAT SHOULD I DO

@RamyAssem i dont like blackberry's.. :-P
Application: tweetfeel

- Cannot deal with complex sentences, e.g. irony.

- No deep linguistic analysis.
Application: The Stock Sonar

- Analysis of financial markets, in particular public companies.
- Sources: news articles, blogs, tweets, etc.
- Main opinion mining task
  - sentiment classification of all documents about a given stock
- Visualization of
  - Daily positive and negative sentiment
  - Price of the stock.
Application: The Stock Sonar

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Application: The Stock Sonar

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Application: 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

• Main opinion mining task
  – Aspect-based opinion mining
Application: Google Products

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Application: Google Products

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Application: Google Products

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 extra-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: ConsumerSearch

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 30x 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. Namely, 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 endeavor. Buying a basic DSLR setup isn't so painful, with decent models like the Nikon D3100 now available for under £500. 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-rivaling 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 low-light sensitivity for improved low-light performance. The effective resolution of the 940,000-pixel 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.
Application: Google Products

Canon PowerShot SX40 HS 12.1 MP Digital Camera

$355 online
482 reviews
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

Reviews
Summary - Based on 482 reviews

What people are saying
pictures
features
zoom/lens
design
video
value
size

"Picture clarity is great."
"Great images, great features, easy to use."
"This is a great low end camera for Canon DSLR users."
"Another point is the overall camera speed."
"Great video quality."
"Amazing product with excellent price!!"
"The SX40 is very easy to use, small learning curve."

Showing reviews that mention: Size - Show all reviews
Outline

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

- Opinion \((e_j, a_j^k, s_o i_{jkl}, h_i, t_l)\)
- 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

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

**Input**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
<th>Full Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of Use</td>
<td>★★★★☆☆</td>
<td>... it has great zoom feature ...</td>
</tr>
<tr>
<td>Durability</td>
<td>★★★★★★</td>
<td>the sound is terrible ...</td>
</tr>
<tr>
<td>Battery Life</td>
<td>★★★★★★</td>
<td>the screen is blurry ...</td>
</tr>
<tr>
<td>Movie quality</td>
<td>★★★★★★</td>
<td>with the affordable price ...</td>
</tr>
</tbody>
</table>

**Output**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom</td>
<td>★★★★★★</td>
</tr>
<tr>
<td>Sound</td>
<td>★★★★★★</td>
</tr>
<tr>
<td>Screen</td>
<td>★★★★★★</td>
</tr>
<tr>
<td>Price</td>
<td>★★★★★★</td>
</tr>
<tr>
<td>Size</td>
<td>★★★★★★</td>
</tr>
</tbody>
</table>
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

---

**Hewlett Packard Photosmart 435 Digital Camera**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Rating</td>
<td>🌟🌟🌟🌟</td>
<td>Nice look, small and light, quick load time, slide the cover open and shoot</td>
<td>Worse picture quality than peers, poor battery life, inaccurate LCD</td>
</tr>
<tr>
<td>Ease of Use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Battery Life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photo Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shutter Lag</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Full Review:** ...the battery life is poor ... the flash is disappointing ... the LCD display is inaccurate ... the self timer is great... provides outstanding ease of use...
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

Hewlett Packard Photosmart 435 Digital Camera

Overall Rating: ★★★★☆
Ease of Use:                        Pros: Nice look, small and light, quick load time, slide the cover open and shoot
Durability:                         Cons: Worse picture quality than peers, poor battery life
Battery Life:                      
Photo Quality:                      Head terms
Shutter Lag

Full Review: ...the battery life is poor... the flash is disappointing... the LCD display is inaccurate... the self timer is great... provides outstanding ease of use...
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

Hewlett Packard Photosmart 435 Digital Camera

Pros: Nice look, small and light, quick load time, slide the cover open and shoot
Cons: Worse picture quality than peers, poor battery life, inaccurate LCD

Full Review: …the battery life is poor … the flash is disappointing … the LCD display is inaccurate … the self timer is great … provides outstanding ease of use…
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

Hewlett Packard Photosmart 435 Digital Camera

**Overall Rating:** ★★★★★

- **Ease of Use:**
- **Durability:**
- **Battery Life:**
- **Photo Quality:**
- **Shutter Lag**

**Pros:** Nice look, small and light, quick load time, slide the cover open and shoot

**Cons:** Worse picture quality than peers, poor battery life, inaccurate LCD

**Full Review:** ...the battery life is poor ... the flash is disappointing ... the LCD display is inaccurate ... the self timer is great... provides outstanding ease of use...
Aspect-based Opinion Mining

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 … |

Output

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Aspect-based Opinion Mining

Input

<table>
<thead>
<tr>
<th>Canon GL2 Mini DVD Camcorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>… 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 …</td>
</tr>
</tbody>
</table>

Aspect Extraction

Output

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Aspect-based Opinion Mining

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 …

Output

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

Rating Prediction

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
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 humiliated 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

• For aspect extraction
  – Using gold standard groupings

\[
Precision = \frac{|ExtractedAspects \cap GoldStandardAspects|}{|ExtractedAspects|}
\]

\[
Recall = \frac{|ExtractedAspects \cap GoldStandardAspects|}{|GoldStandardAspects|}
\]

\[
F\text{-measure} = \frac{2 \times Recall \times Precision}{Recall + Precision}
\]
Evaluation Metrics

• For aspect extraction
  – Using ground truth word distribution

\[ D_{KL}(P \parallel Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \]

\( P(x) \): 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 \( i \)th aspect
\( r_i \) true rating of the \( i \)th 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

\[
\text{perplexity}(D_{\text{test}}) = \exp\left\{ -\frac{\sum_{d=1}^{N} \log P(v_d)}{\sum_{d=1}^{N} M_d} \right\}
\]

\(N:\) #reviews \hspace{1cm} \(M:\) #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

- 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).
Opinion Mining in Online Reviews: Recent Trends

Samaneh Moghaddam & Martin Ester
Tutorial at WWW2013
Part 2
Aspect-Based Opinion Mining

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 ...

Output

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Aspect-Based Opinion Mining

**Input**

<table>
<thead>
<tr>
<th>Canon GL2 Mini DVD Camcorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>… 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 …</td>
</tr>
</tbody>
</table>

**Aspect Extraction**

**Output**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
Aspect-Based Opinion Mining

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 …

Output

Rating Prediction

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoom</td>
<td>5</td>
</tr>
<tr>
<td>price</td>
<td>4</td>
</tr>
<tr>
<td>picture quality</td>
<td>4</td>
</tr>
<tr>
<td>battery life</td>
<td>2</td>
</tr>
<tr>
<td>screen</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
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**
  - Frequency- and relation-based approaches
  - Model-based approaches

- **Rating (polarity) prediction**
  - Supervised learning methods
  - Lexicon-based methods
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
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
  – Frequency-based Methods
    • Feature-based Summarization
    • OPINE
  – Relation-based Methods
  – Hybrid Methods
• Model-based Approaches
• Guidelines for Designing 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

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Frequency-based Methods

• 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
    \[
    PMI(f, p) = \frac{\text{Hits}(f + p)}{\text{Hits}(p) \times \text{Hits}(f)}
    \]
  - 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
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
  – Frequency-based Methods
  – Relation-based Methods
    • Opinion Observer
    • Multi-Facet Rating
    • Tree Kernel approach
  – Hybrid Methods
• Model-based Approaches
• Guidelines for Designing 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 (Liu 2012)
  – 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 $\rightarrow$ General Inquirer
  – Filtering $\rightarrow$ variance of distribution across polarity

\[
\text{PATTERN ::= } A / B / C \\
A ::= [AT] \text{ADJ NOUN} \\
B ::= \text{NOUN VERB ADJ} \\
C ::= Hv A \\
\text{NOUN ::= } [AT] [NN$] \text{NN} \\
\text{ADJ ::= } [\text{CONG}] \text{ADV ADJ} \\
\text{ADV ::= } \text{RB ADV} / \text{QL ADV} / \text{JJ} / \text{AP ADV} / \\
\text{CONG ::= } \text{CC} / \text{CS} \\
\text{VERB ::= } V / Be
\]
Relation-based Methods

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

<table>
<thead>
<tr>
<th>Extracted Opinion Phrases</th>
<th>Enriched GI Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>great location</td>
<td>[Strong] [Positive] location</td>
</tr>
<tr>
<td>great hotel</td>
<td>[Strong] [Positive] hotel</td>
</tr>
<tr>
<td>very friendly staff</td>
<td>very [Emot] [positive] staff</td>
</tr>
<tr>
<td>good location</td>
<td>[Positive] location</td>
</tr>
</tbody>
</table>
Relation-based Methods

• 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
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
  – Frequency-based Methods
  – Relation-based Methods
  – Hybrid Methods
    • Sentiment Summarizer
    • Opinion Digger
• Model-based Approaches
• Guidelines for Designing 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

• Opinion Digger (Moghaddam et al. 2010)
  – Find freq. noun phrases
  – Mine opinion patterns using known aspects
  – Extract the nearest adjectives as sentiment
  – Estimate rating [1, 5] using the rating guideline

<table>
<thead>
<tr>
<th>Known Aspect</th>
<th>Matching segments</th>
<th>Mined Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>photo quality</td>
<td>disappointment photo quality</td>
<td><em>JJ</em> ASP</td>
</tr>
<tr>
<td>battery life</td>
<td>battery life is great</td>
<td>_ASP_VB_JJ</td>
</tr>
<tr>
<td>photo quality</td>
<td>lovely feature is photo quality</td>
<td>_JJ_NP_VB_ASP</td>
</tr>
</tbody>
</table>

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
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
• Opinion Mining Tasks
• Applications
• 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
• Guidelines for Designing 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

\[ p(y \mid x) = \frac{1}{Z(x)} \exp \left\{ \sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t) \right\} \]
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
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
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
  – Supervised learning techniques
  – Topic modeling techniques
    • TSM
    • MG-LDA
    • JST
    • ...
• Guidelines for Designing 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

- **PLSI**

\[
P(z, w, d | \beta) = P(d) \prod_{n=1}^{N} [P(z_n | d) P(w_n | z_n, \beta)]
\]

- **LDA**

\[
P(z, w, \theta | \alpha, \beta) = P(\theta | \alpha) \prod_{n=1}^{N} [P(z_n | \theta) P(w_n | z_n, \beta)]
\]
Topic-Modeling Techniques

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

![Diagram of TSM model](image.png)
Topic-Modeling Techniques

- 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.

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
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.

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Topic-Modeling Techniques

- 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

• 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

- **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**

- **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

- 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

• 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

- **SAS (Mukherjee et al. 2012a)**
  - Task: identifying aspects and sentiments
  - Novelty: using user provided seed words for clustering aspects
Topic-Modeling Techniques

- Other models
  - Branavan et al. (2008)
  - Lu et al. (2009)
  - Brody et al. (2010)
  - Mukherjee et al. (2012b)
  - Luo 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
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
• Guidelines for Designing LDA-based Models
  – Abstract LDA-based Models
  – Comparison
  – Addressing the cold start problem
• Conclusion and Future Directions
Design Guidelines for LDA-based Models

• 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 (Moghaddam et al. 2012(b))
Design Guidelines for LDA-based Models

LDA

\[ \alpha \rightarrow \theta \rightarrow z \rightarrow w \rightarrow \beta \]

PLDA

\[ \alpha \rightarrow \theta \rightarrow z \rightarrow h \rightarrow m \rightarrow \beta \rightarrow \pi \]

S-LDA

\[ \alpha \rightarrow \theta \rightarrow a \rightarrow r \rightarrow w \rightarrow \beta \]

S-PLDA

\[ \alpha \rightarrow \theta \rightarrow a \rightarrow h \rightarrow r \rightarrow m \rightarrow \beta \rightarrow \pi \]

D-LDA

\[ \alpha \rightarrow \theta \rightarrow a \rightarrow r \rightarrow w \rightarrow \beta \]

D-PLDA

\[ \alpha \rightarrow \theta \rightarrow a \rightarrow h \rightarrow r \rightarrow m \rightarrow \beta \rightarrow \pi \]
Design Guidelines for LDA-based Models

• Extraction of opinion phrases
  – Frequent noun technique (Moghaddam et al. 2011)
    • Frequency of phrases
  – POS patterns (Zhan et al. 2011)
    • Syntactic relations, e.g., NN_VB_JJ → <NN, JJ>
  – Dependency patterns (Moghaddam et al. 2012)
    • Semantic relations, e.g., amod(NN, JJ) → <NN, JJ>

• Data set (Epinions 500K reviews)
  – Define 5 subsets of items with different #reviews
Design Guidelines for LDA-based Models

• Extraction of opinion phrases
  - Frequency
    • Frequency of noun phrases (Moghaddam et al. 2011)
    • Frequency of POS patterns (Zhan et al. 2011)
  - Syntactic relations, e.g., \text{NN} \_\_ \text{VB} \_\_ \text{JJ} \rightarrow \text{<NN, JJ>}
    • Dependency patterns (Moghaddam et al. 2012)
  - Semantic relations, e.g., \text{amod(NN, JJ)} \rightarrow \text{<NN, JJ>}

• Data set (Epinions 500K reviews)
  - Define 5 subsets of items with different #reviews

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Subset & \#Products & \#Rev./Product \\
\hline
1 < \#Rev. \leq 10 & 36,166 & 3 \\
10 < \#Rev. \leq 50 & 7,886 & 19 \\
50 < \#Rev. \leq 100 & 869 & 67 \\
100 < \#Rev. \leq 200 & 368 & 137 \\
200 < \#Rev. & 179 & 341 \\
\hline
\end{tabular}
\end{table}
Design Guidelines for LDA-based Models

- Perplexity comparison

\[ \text{perplexity}(D_{\text{test}}) = \exp\left\{ -\frac{\sum_{d=1}^{D} \log P(h_d, m_d)}{\sum_{d=1}^{D} N_d} \right\} \]
Design Guidelines for LDA-based Models

- Perplexity comparison

- Evaluation on a labeled dataset (Hu et al. 2004)

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>MSE</th>
<th>Perplexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-LDA</td>
<td>0.54</td>
<td>0.51</td>
<td>1.25</td>
<td>813.11</td>
</tr>
<tr>
<td>LDA</td>
<td>0.54</td>
<td>0.52</td>
<td>1.22</td>
<td>795.72</td>
</tr>
<tr>
<td>D-LDA</td>
<td>0.58</td>
<td>0.55</td>
<td>1.18</td>
<td>748.26</td>
</tr>
<tr>
<td>PLDA</td>
<td>0.81</td>
<td>0.73</td>
<td>0.96</td>
<td>587.82</td>
</tr>
<tr>
<td>S-PLDA</td>
<td>0.83</td>
<td>0.73</td>
<td>0.93</td>
<td>335.02</td>
</tr>
<tr>
<td>D-PLDA</td>
<td>0.87</td>
<td>0.78</td>
<td>0.85</td>
<td>131.80</td>
</tr>
</tbody>
</table>
Design Guidelines for LDA-based Models

- When learning from bag-of-phrases, having separate latent variables can improve the performance.

LDA vs. S-LDA

PLDA vs. S-PLDA
Design Guidelines for LDA-based Models

- When learning from bag-of-phrases, having separate latent variables can improve the performance.
- When learning from phrases and having separate latent variables, assuming dependency improves performance.

S-LDA vs. D-LDA

S-PLDA vs. D-PLDA
Design Guidelines for LDA-based Models

- When learning from bag-of-phrases, having separate latent variables can improve the performance.
- When learning from phrases and having separate latent variables, assuming dependency improves performance.
- When separate latent variables are assumed, using preprocessing techniques can improve the performance.
Design Guidelines for LDA-based Models

- When learning from bag-of-phrases, having separate latent variables can improve the performance.
- When learning from phrases and having separate latent variables, assuming dependency improves performance.
- When separate latent variables are assumed, using preprocessing techniques can improve the performance.
- Using dependency patterns consistently achieves the best performance for extracting opinion phrases.

<table>
<thead>
<tr>
<th>Preprocessing</th>
<th>Frequent nouns</th>
<th>POS patterns</th>
<th>Dependency patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset of Products</td>
<td>PLDA</td>
<td>S-PLDA</td>
<td>D-PLDA</td>
</tr>
<tr>
<td>1 &lt; #Rev. &lt;= 10</td>
<td>11834.99</td>
<td>11824.35</td>
<td>11740.95</td>
</tr>
<tr>
<td>10 &lt; #Rev. &lt;= 50</td>
<td>6724.61</td>
<td>6129.57</td>
<td>5829.33</td>
</tr>
<tr>
<td>50 &lt; #Rev. &lt;= 100</td>
<td>3024.59</td>
<td>2243.22</td>
<td>1882.91</td>
</tr>
<tr>
<td>100 &lt; #Rev. &lt;= 200</td>
<td>1885.52</td>
<td>1511.47</td>
<td>660.96</td>
</tr>
<tr>
<td>200 &lt; #Rev.</td>
<td>1337.93</td>
<td>1165.88</td>
<td>406.27</td>
</tr>
</tbody>
</table>
Design Guidelines for LDA-based Models

• When learning from bag-of-phrases, having separate latent variables can improve the performance.
• When learning from phrases and having separate latent variables, assuming dependency improves performance.
• When separate latent variables are assumed, using preprocessing techniques can improve the performance.
• Using dependency patterns consistently achieves the best performance for extracting opinion phrases.
• For items with few reviews, LDA model outperforms the more complex models. For others, D-PLDA performs best.

<table>
<thead>
<tr>
<th></th>
<th>LDA</th>
<th>S-LDA</th>
<th>D-LDA</th>
<th>PLDA</th>
<th>S-PLDA</th>
<th>D-PLDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>4413.6</td>
<td>4567.2</td>
<td>4729.3</td>
<td>5463.8</td>
<td>5422.2</td>
<td>5413.6</td>
</tr>
</tbody>
</table>

Moghaddam & Ester: Opinion Mining in Online Reviews: Recent Trends, Tutorial at WWW 2013
Addressing the Cold Start Problem

- FLDA (Moghaddam et al. 2013)
  - Task: Addressing the cold start problem
  - Novelty: Learns at the category level and consider reviewer factor
Outline

• Introduction
• Opinion Mining Tasks
• Applications
• Aspect-based Opinion Mining
• Frequency and Relation based Approaches
• Model-based Approaches
• Guidelines for Designing LDA-based Models
• Conclusion and Future Directions
  – Summary
  – Future Research Directions
Summary

• Introduction to Opinion Mining
  – What is opinion mining?
  – Lots of challenging research issues
  – Case study: Google Products

• General opinion mining tasks
  – Document-level
  – Sentence-level
  – Phrase-level
Summary

• Aspect-based opinion mining
  – Challenges
  – Applications
  – Benchmark Datasets

• Frequency and relation based approaches
  – Frequency-based
  – Relation-based
  – Hybrid methods
Summary

• Model-based approaches
  – Supervised learning
  – Topic modeling

• Design guidelines for LDA-based models
  – Comparison of abstract LDA-based models
  – Addressing the cold start problem
Future Research Directions

• Better evaluation techniques

• Dealing with noun and verb sentiments

• Discovery of implicit aspects

• Coreference resolution
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

- Exploring the impact of different additional input sources

- Using aspect-based OM in other applications
Thank You!
References

- Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. Multi‐facet rating of product reviews. In ECIR ’09.
- 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.
References

• Du, Weifu, and Songbo Tan. An iterative reinforcement approach for ne-grained opinion mining. In NAACL '09.
• Hatzilvassiloglou V., McKeown K.: Predicting the semantic orientation of adjectives, ACL 1997.
• HongningWang, Yue Lu, and ChengXiang Zhai. Latent aspect rating analysis without aspect keyword supervision. In KDD '11.
• He, Yulan, Chenghua Lin, and Harith Alani. Automatically extracting polarity-bearing topics for cross-domain sentiment classification. In HLT '11.
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.
- Jo, Yohan, and Alice H. Oh. Aspect and sentiment unification model for online review analysis. In WSDM '11.
• Kobayashi, Nozomi, Kentaro Inui, and Yuji Matsumoto. Extracting aspect-evaluation and aspect-of
• Kobayashi, Nozomi, Ryu Iida, Kentaro Inui, and Yuji Matsumoto. Opinion mining on the Web by
• Ku, Lun-Wei, Yu-Ting Liang, and Hsin-Hsi Chen. Opinion extraction, summarization and tracking in
• 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
• Li, Fangtao, Jialin Pan, Sinno, Jin, Ou, Yang, Qiang and Zhu, Xiaoyan. 2012. Cross-domain co-
extraction of sentiment and topic lexicons. In ACL’12.
• 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
References


• Lu, Yue, ChengXiang Zhai, and Neel Sundaesran. Rated aspect summarization of short comments. In WWW ’09.


• 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.
• Moghaddam, samaneh and M. Ester. The FLDA model for aspect-based opinion mining: Addressing the cold start problem. In WWW'13.
• Mukherjee, Arjun and Liu, Bing. Aspect extraction through semi-supervised modeling. In ACL'12.
References

- 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.
References

- Velikovich, Leonid; Sasha Blair-Goldensohn; Kerry Hannan; and Ryan McDonald: The viability of web-derived polarity lexicons, ACL 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.
- Zhan, Tian-Jie, and Chun-Hung Li. Semantic dependent word pairs generative model for ne-grained product feature mining. In PAKDD ’11.
References

• 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.