Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- **Aspect-based sentiment analysis**
- Aspect-based opinion summarization
- Opinion lexicon generation
- Mining comparative opinions
- Some other problems
- Opinion spam detection
- Utility or helpfulness of reviews
- Summary

We need to go further

- Sentiment classification at both the document and sentence (or clause) levels are useful, but
  - They do not find what people liked and disliked.
- They do not identify the targets of opinions, i.e.,
  - Entities and their aspects
  - Without knowing targets, opinions are of limited use.
- We need to go to the entity and aspect level.
  - *Aspect-based opinion mining and summarization* (Hu and Liu 2004).
  - We thus need the full opinion definition.

Recall an opinion is a quintuple

- An *opinion* is a quintuple
  $(e_j, a_{jk}, so_{ijkl}, h_i, t_l)$,
  where
  - e_j is a target entity.
  - a_{jk} is an aspect/feature of the entity e_j.
  - so_{ijkl} is the sentiment value of the opinion of the opinion holder h_i on feature a_{jk} of entity e_j at time t_l. so_{ijkl} is +ve, -ve, or neu, or a more granular rating.
  - h_i is an opinion holder.
  - t_l is the time when the opinion is expressed.

Aspect-based sentiment analysis

- Much of the research is based on online reviews
- For reviews, aspect-based sentiment analysis is easier because the entity (i.e., product name) is usually known
  - Reviewers simply express positive and negative opinions on different aspects of the entity.
- For blogs, forum discussions, etc., it is harder:
  - both entity and aspects of entity are unknown,
  - there may also be many comparisons, and
  - there is also a lot of irrelevant information.

Find entities (entity set expansion)

- Although similar, it is somewhat different from the traditional named entity recognition (NER).
- E.g., one wants to study opinions on phones
  - given Motorola and Nokia, find all phone brands and models in a corpus, e.g., Samsung, Moto,
- Formulation: Given a set *Q* of seed entities of class *C*, and a set *D* of candidate entities, we wish to determine which of the entities in *D* belong to *C*.
  - A classification problem. It needs a binary decision for each entity in *D* (belonging to *C* or not)
  - But it's often solved as a ranking problem

Aspect extraction

- Goal: Given an opinion corpus, extract all aspects
- A frequency-based approach (Hu and Liu, 2004): nouns (NN) that are frequently talked about are likely to be true aspects (called frequent aspects) .
- Why the frequency based approach?
  - Different reviewers tell different stories (irrelevant)
  - When product aspects/features are discussed, the words they use converge.
  - They are the main aspects.
- Sequential/association pattern mining finds frequent nouns and noun phrases.

An example review

GREAT Camera., Jun 3, 2004
Reviewer: **jprice174** from Atlanta, Ga.

I did a lot of research last year before I bought this camera... It kinda hurt to leave behind my beloved nikon 35mm SLR, but I was going to Italy, and I needed something smaller, and digital.

The pictures coming out of this camera are amazing. The 'auto' feature takes great pictures most of the time. And with digital, you're not wasting film. …

….

Infrequent aspect extraction

- To improve recall due to loss of infrequent aspects. It uses opinions words to extract them
- Key idea: opinions have targets, i.e., opinion words are used to modify aspects and entities.
  - “The pictures are absolutely amazing.”
  - “This is an amazing software.”
- The modifying relation was approximated with the nearest noun to the opinion word.
- The idea was generalized to dependency in (Zhuang et al 2006) and double propagation in (Qiu et al 2009;2011).
  - It has been used in many papers and practical systems

Explicit and implicit aspects
(Hu and Liu 2004)

- Explicit aspects: Aspects explicitly mentioned as nouns or noun phrases in a sentence
  - The picture quality is of this phone is great.
- Implicit aspects: Aspects not explicitly mentioned in a sentence but are implied
  - "This car is so expensive."
  - "This phone will not easily fit in a pocket.
  - "Included 16MB is stingy"
- Not much work has been done on mining or mapping implicit aspects.

Implicit aspect mapping

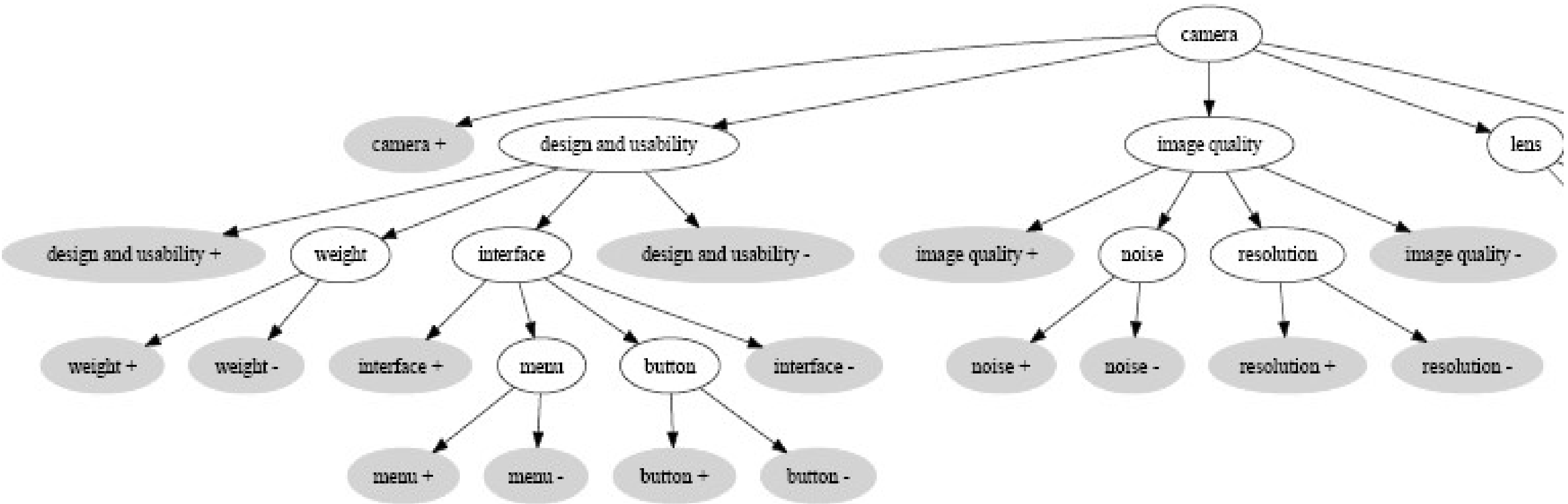
- There are many types of implicit aspect expressions. Adjectives and adverbs are perhaps the most common type.
  - Most adjectives modify or describe some specific attributes of entities.
  - “expensive” $\Rightarrow$ aspect “price,” “beautiful” $\Rightarrow$ aspect “appearance”, “heavy” $\Rightarrow$ aspect “weight”
- Although manual mapping is possible, in different contexts, the meaning can be different.
  - E.g., “The computation is expensive”.

Identify aspect synonyms (Carenini et al 2005)

- Once aspect expressions are discovered, group them into aspect categories.
  - E.g., power usage and battery life are the same.
- It proposed a method based on some similarity metrics, but it needs a taxonomy of aspects.
  - The system merges each discovered aspect to a aspect node in the taxonomy.
  - Similarity metrics: string similarity, synonyms and other distances measured using WordNet.
- Many ideas in Web information integration are applicable.

Sentiment ontology tree (Wei and Gulla, 2010)

- Recall in the definition of opinions, we simplified the tree structure to two levels (entity & aspects).
- This paper uses a full tree ontology to denote the relationships of aspects of a product.



Sentiment ontology tree (contd)

- The leaves of the tree are positive or negative sentiments.
- It then uses a hierarchical classification model to learn to assign an sentiment to each node, which is reflected as a child leaf node.
  - Hierarchical classifier is useful here because it considers parents when classifying children.
- However, the ontology for each product has to be built manually.

Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- Aspect-based sentiment analysis
- **Aspect-based opinion summarization**
- Opinion lexicon generation
- Mining comparative opinions
- Some other problems
- Opinion spam detection
- Utility or helpfulness of reviews
- Summary

Aspect-based opinion summarization

- A multi-document summarization problem.
  - An opinion from a single person is usually not sufficient for action unless from a VIP (e.g., President)
- Key Idea: Use aspects as basis for a summary
  - Not done in traditional multi-document summarization.
- We have discussed the aspect-based summary using quintuples earlier (Hu and Liu 2004; Liu, 2010).
  - Also called: *Structured Summary*
- Similar approaches are also taken in
  - (e.g., Ku et al 2006; Carenini, Ng and Paul 2006) and
  - By most topic model based methods

Text summary of opinions

- One can also generate a summary in the tradition fashion, e.g., producing a short text summary (Lerman et al 2009), by extracting some important sentences, etc.
  - Weakness: It is only qualitative but not quantitative.
- One can generate sentences based on aspects and opinions using some templates.
  - E.g., 60% of the people like the picture quality.

Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- Aspect-based sentiment analysis
- Aspect-based opinion summarization
- **Mining comparative opinions**
- Some other problems
- Opinion spam detection
- Utility or helpfulness of reviews
- Summary

Comparative Opinions
(Jindal and Liu, 2006)

- *Gradable*
  - *Non-Equal Gradable*: Relations of the type *greater* or *less than*
    - *Ex: "optics of camera A is better than that of camera B"*
  - *Equative*: Relations of the type *equal to*
    - Ex: "*camera A and camera B both come in 7MP*"
  - *Superlative*: Relations of the type *greater* or *less than all others*
    - Ex: "*camera A is the cheapest in market*"

Analyzing 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.

- Note: not positive or negative opinions.

An example

- Consider the comparative sentence
  - "*Canon's optics is better than those of Sony and Nikon*."
  - Written by John in 2010.
- The extracted comparative opinion/relation:
  - ({Canon}, {Sony, Nikon}, {optics}, *preferred*:{Canon}, John, 2010)

Common comparatives

- In English, comparatives are usually formed by adding **-***er* and superlatives are formed by adding *-est* to their base adjectives and 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*.
- Irregular comparatives and superlatives, i.e., *more most*, *less*, *least*, *better*, *best*, *worse*, *worst*, etc

Analysis of comparative opinions

- Gradable comparative sentences can be dealt with *almost* as normal opinion sentences.
  - E.g., "*optics of camera A is better than that of camera B*"
  - Positive: *"optics of camera A"*
  - Negative: *"optics of camera B"*
- Difficulty: recognize non-standard comparatives
  - E.g., "I am so happy because my new iPhone is nothing like my old slow ugly Droid."
  - ?

Identifying preferred entities

(Ganapathibhotla and Liu, 2008)

- The following rules can be applied

  Comparative Negative ::= increasing comparative N
  | decreasing comparative P

  Comparative Positive ::= increasing comparative P
  | decreasing comparative N

  - E.g., “Coke tastes better than Pepsi”
  - “Nokia phone’s battery life is longer than Moto phone”
- Context-dependent comparative opinion words
  - Using context pair: (aspect, JJ/JJR)
  - Deciding the polarity of (battery_life, longer) in a corpus

Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- Aspect-based sentiment analysis
- Aspect-based opinion summarization
- Opinion lexicon generation
- Mining comparative opinions
- **Some other problems**
- Opinion spam detection
- Utility or helpfulness of reviews
- Summary

Coreference resolution: semantic level?

- Coreference resolution (Ding and Liu, 2010)
  - "I bought the Sharp tv a month ago. The picture quality is so bad. Our other Sony tv is much better than this Sharp. ***It*** is also so expensive".
    - "it" means "Sharp"
  - "I bought the Sharp tv a month ago. The picture quality is so bad. Our other Sony tv is much better than this Sharp. ***It*** is also very reliable."
    - "it" means "Sony
- Sentiment consistency.

Coreference resolution (contd)

- “The picture quality of this Canon camera is very good. ***It*** is not expensive either.”
  - Does “it” mean “Canon camera” or “Picture Quality”?
    - Clearly it is Canon camera because picture quality cannot be expensive.
    - Commonsense knowledge, but can be discovered.
- For coreference resolution, we actually need to
  - do sentiment analysis first, and
  - mine adjective-noun associations using dependency
- Finally, use supervised learning

Some interesting sentences

- “Trying out Google chrome because Firefox keeps crashing.”
  - The opinion about Firefox is clearly negative, but for Google chrome, there is no opinion.
  - We need to segment the sentence into clauses to decide that “crashing” only applies to Firefox.
  - “Trying out” also indicates no opinion.
- How about this
  - “I changed to Audi because BMW is so expensive.”

Some interesting sentences (contd)

- Conditional sentences are hard to deal with (Narayanan et al. 2009)
  - "If I can find a good camera, I will buy it."
  - But conditional sentences can have opinions
    - "If you are looking for a good phone, buy Nokia"
- Questions may or may not have opinions
  - No sentiment
    - "Are there any great perks for employees?"
  - With sentiment
    - "Any idea how to repair this lousy Sony camera?"

Some interesting sentences (contd)

- Sarcastic sentences
  - "What a great car, it stopped working in the second day."
- Sarcastic sentences are very common in political blogs, comments and discussions.
  - They make political blogs difficult to handle
  - Many political aspects can also be quite complex and hard to extract because they cannot be described using one or two words.
- Some initial work by (Tsur, Davidov, Rappoport 2010)

Some interesting sentences (contd)

- The following two sentences are from reviews in the paint domain.
  - “For paint_X, one coat can cover the wood color.”
  - “For paint_Y, we need three coats to cover the wood color.
- We know that paint_X is good and Paint_Y is not, but how by a system.
  - Do we need commonsense knowledge and understanding of the text?

Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- Aspect-based sentiment analysis
- Aspect-based opinion summarization
- Opinion lexicon generation
- Mining comparative opinions
- Some other problems
- **Opinion spam detection**
- Utility or helpfulness of reviews
- Summary

Opinion spam detection
(Jindal and Liu 2007, 2008)

- Opinion spamming refers to people giving 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 fake reviews (biased reviews, paid reviews)

Problem is wide-spread

Professional Fake Review Writing Services

- Post positive reviews
- Fake review writer
- Product review writer for hire
- Hire a content writer

Manipulating Social Media (sock puppets - fake identities - fake personas)

- Revealed: US spy operation that manipulates social media, Guardian.co.uk, Thursday 17 March 2011.
- America's absurd stab at systematising sock puppetry, Guardian.co.uk, Thursday 17 March 2011.

China's Internet "Water Army" (Shuijun) - Opinion Spammers

- You can hire people to write and post fake reviews or comments, and even bribe staff at review, forum
- 'Water Army' Whistleblower Threatened, January 7, 2011, People's Daily.
- The Chinese Online "Water Army", June 25, 2010, Wired.com.
- If you read Chinese, see this description from Baidu Baike at baidu.com.

An example practice of review spam

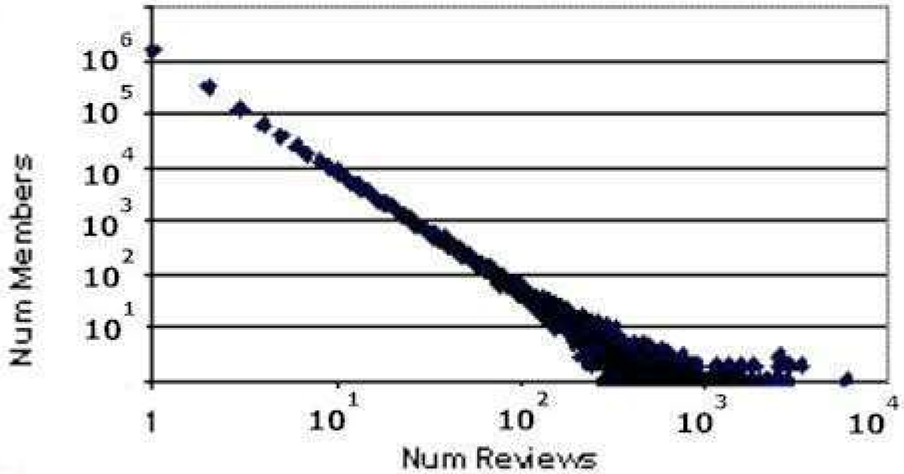
Belkin International, Inc

- Top networking and peripherals manufacturer | Sales ~ $500 million in 2008
- Posted an ad for writing fake reviews on amazon.com (65 cents per review)

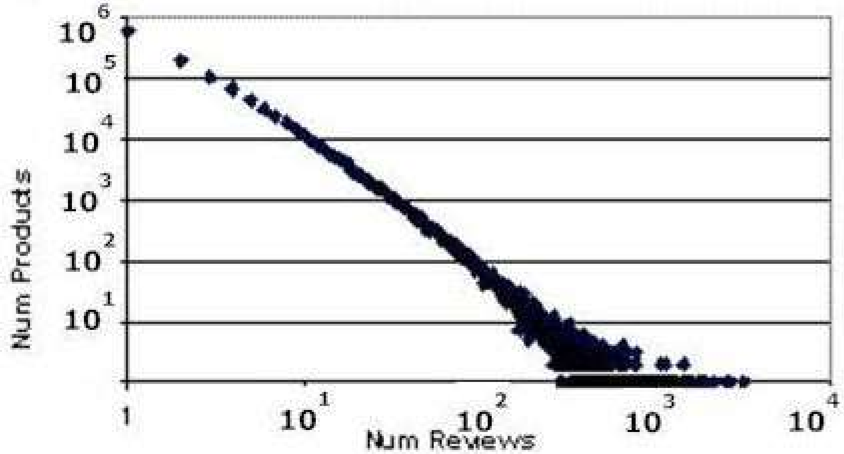
Jan 2009

Log-log plot

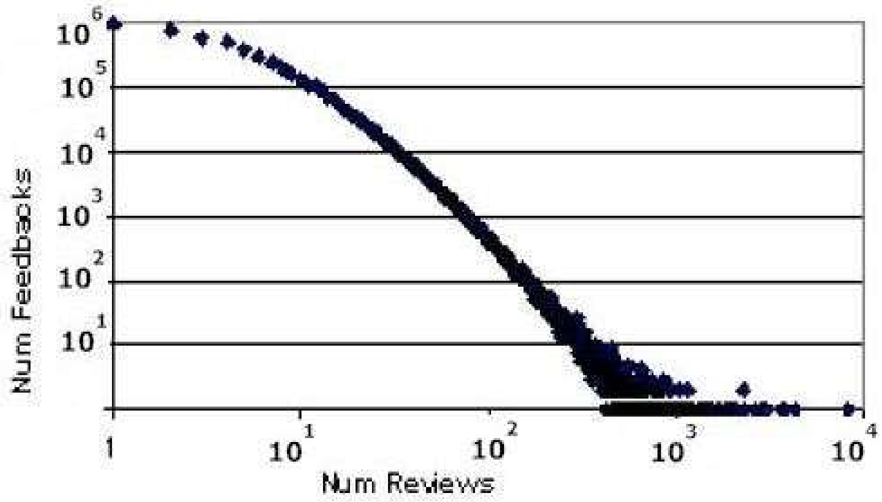
Amazon reviews, reviewers and products



- Fig. 1 reviews and reviewers



- Fig. 2 reviews and products



- Fig. 3 reviews and feedback s

Categorization of opinion spam

(Jindal and Liu 2008)

- Type 1 (fake reviews)
  Ex:
- Type 2 (Reviews on Brands Only) (?)
  Ex: "*I don't trust HP and never bought anything from them*"
- Type 3 (Non-reviews)
  - Advertisements
    Ex: "*Detailed product specs: 802.11g, IMR compliant, …*"
    "*…buy this product at: compuplus.com*"
  - Other non-reviews
    Ex: "*What port is it for*"
    "*The other review is too funny*"
    "*Go Eagles go*"

Type 1 Spam Reviews

- Hype spam – promote one's own products
- Defaming spam – defame one's competitors' products

Table 4. Spam reviews vs. product quality

	Positive spam review	Negative spam review
Good quality product	1	**2**
Bad quality product	**3**	4
Average quality product	**5**	**6**

Harmful Regions

- Very hard to detect manually

Harmful spam are outlier reviews?

- Assumption: Most reviewers and reviews are honest,
  - Not true when a group of people spam on a product (called group spam, discussed later).
- Outliers reviews: Reviews which deviate a great deal from the average product rating
- Harmful spam reviews:
  - Outliers are necessary but not sufficient condition for harmful spam reviews.
  - This idea helps us identify learning features.

Types of spam reviews

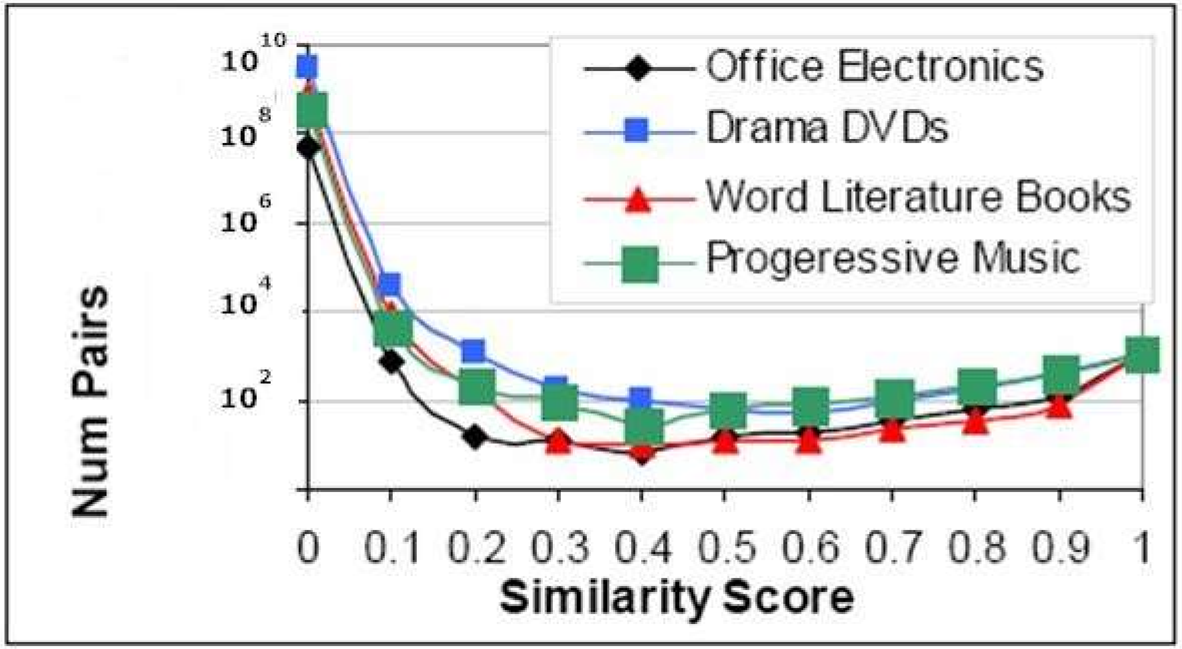
- Type 1 (fake review):These are reviews that give undeserving positive or negative opinions to some target entities.
- Type 2 (review on brand only): These reviews do not comment on the specific products that they are supposed to review, but only comment on the brands, the manufacturers, or the sellers of the products. Example: “I hate HP. I never buy any of their products”.
- Type 3 (non-review): These are not reviews or opinionated although they appear as reviews. There are two main sub-types:
  - Advertisements
  - Other irrelevant texts containing no opinions (e.g., questions, answers, and random texts).

Spam detection

- Type 2 and Type 3 spam reviews are relatively easy to detect
  - Supervised learning, e.g., logistic regression
  - It performs quite well, and not discuss it further.
- Type 1 spam (fake) reviews
  - Manual labeling is extremely hard
  - Propose to use duplicate and near-duplicate reviews as positive training data

Duplicate reviews

Two reviews which have similar contents are called duplicates



Four types of duplicates

1. Same userid, same product
2. Different userid, same product
3. Same userid, different products
4. Different userid, different products

- The last three types are very likely to be fake!

Supervised model building

- Logistic regression
  - Training: duplicates as spam reviews (positive) and the rest as non-spam reviews (negative)
- Use the follow data attributes
  - Review centric features (content)
    - About reviews (contents (n-gram), ratings, etc)
  - Reviewer centric features
    - About reviewers (different unusual behaviors, etc)
  - Product centric features
    - Features about products reviewed (sale rank, etc)

Predictive power of duplicates

- Representative of all kinds of spam
- Only 3% duplicates accidental
- Duplicates as positive examples, rest of the reviews as negative examples

Table 5. AUC values on duplicate spam reviews.

Features used	AUC
All features	78%
Only review features	75%
Only reviewer features	72.5%
Without feedback features	77%
Only text features	63%

- reasonable predictive power
- Maybe we can use duplicates as type 1 spam reviews(?)

Tentative classification results

- Negative outlier reviews tend to be heavily spammed
- Those reviews that are the only reviews of products are likely to be spammed
- Top-ranked reviewers are more likely to be spammers
- Spam reviews can get good helpful feedbacks and non-spam reviews can get bad feedbacks

Roadmap

- Opinion Mining Problem
- Document sentiment classification
- Sentence subjectivity & sentiment classification
- Aspect-based sentiment analysis
- Aspect-based opinion summarization
- Opinion lexicon generation
- Mining comparative opinions
- Some other problems
- Opinion spam detection
- **Utility or helpfulness of reviews**
- Summary

Utility or quality of reviews

- Goal: Determining the usefulness, helpfulness, or utility of each review.
  - It is desirable to rank reviews based on utilities or qualities when showing them to users, with the highest quality review first.
- Many review aggregation sites have been practicing this, e.g., amazon.com.
  - "*x of y people found the following review helpful*."
  - Voted by user - "*Was the review helpful to you?*"

Application motivations

- Although review sites use helpfulness feedback to rank,
  - A review takes a long time to gather enough feedback.
    - New reviews will not be read.
  - Some sites do not provide feedback information.
- It is thus beneficial to score each review once it is submitted to a site.

Regression formulation
(Zhang and Varadarajan, 2006; Kim et al. 2006)

- Formulation: Determining the utility of reviews is usually treated as a regression problem.
  - A set of features is engineered for model building
  - The learned model assigns an utility score to each review, which can be used in review ranking.
- Unlike fake reviews, the ground truth data used for both training and testing are available
  - Usually the user-helpfulness feedback given to each review.

Features for regression learning

- Example features include
  - review length, review rating, counts of some POS tags, opinion words, tf-idf scores, wh-words, product aspect mentions, comparison with product specifications, timeliness, etc (Zhang and Varadarajan, 2006; Kim et al. 2006; Ghose and Ipeirotis 2007; Liu et al 2007)
- Subjectivity classification was applied in (Ghose and Ipeirotis 2007).
- Social context was used in (O'Mahony and Smyth 2009; Lu et al. 2010).