商業智慧
Business Intelligence

意見分析
(Opinion Mining)

1002BI09
IM EMBA
Fri 12,13,14 (19:20-22:10) D502

Min-Yuh Day
戴敏育
Assistant Professor
專任助理教授
Dept. of Information Management, Tamkang University
淡江大學 資訊管理學系

http://mail. tku.edu.tw/myday/
2012-06-01

課程大綱 (Syllabus)

週次　日期　內容（Subject/Topics）　備註

1　101/02/17　商業智慧導論 (Introduction to Business Intelligence)

2　101/02/24　管理決策支援系統與商業智慧
　　　　　　　(Management Decision Support System and Business Intelligence)

3　101/03/02　企業績效管理 (Business Performance Management)

4　101/03/09　資料倉儲 (Data Warehousing)

5　101/03/16　商業智慧的資料探勘 (Data Mining for Business Intelligence)

6　101/03/24　商業智慧的資料探勘 (Data Mining for Business Intelligence)

7　101/03/30　個案分析一 (分群分析)：Banking Segmentation
　　　　　　　(Cluster Analysis – KMeans)

8　101/04/06　教學行政觀摩日 (--No Class--)

9　101/04/13　個案分析二 (關連分析)：Web Site Usage Associations
　　　　　　　(Association Analysis)

課程大綱 (Syllabus)

週次　日期　　　內容（Subject/Topics）　備註

10　101/04/20　期中報告 (Midterm Presentation)

11　101/04/27　個案分析三 (決策樹、模型評估)：
Enrollment Management Case Study
(Decision Tree, Model Evaluation)

12　101/05/04　個案分析四 (迴歸分析、類神經網路)：Credit Risk Case Study
(Regression Analysis, Artificial Neural Network)

13　101/05/11　文字探勘與網頁探勘 (Text and Web Mining)

14　101/05/18　智慧系統 (Intelligent Systems)

15　101/05/25　社會網路分析 (Social Network Analysis)

16　101/06/01　意見分析 (Opinion Mining)

17　101/06/08　期末報告1 (Project Presentation 1)

18　101/06/15　期末報告2 (Project Presentation 2)

Outline

- Opinion Mining
- Sentiment Analysis

Opinion Mining and Sentiment Analysis

- Mining opinions which indicate positive or negative sentiments

- Analyzes people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics, and their attributes.

Opinion Mining and Sentiment Analysis

- Computational study of
opinions,
sentiments,
subjectivity,
evaluations,
attitudes,
appraisal,
affects,
views,
emotions,
ets., expressed in text.
  - Reviews, blogs, discussions, news, comments, feedback, or any other documents

Terminology

- Sentiment Analysis
  is more widely used in industry

- Opinion mining / Sentiment Analysis
  are widely used in academia

- Opinion mining / Sentiment Analysis
  can be used interchangeably

Source: Bing Liu (2011) , "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data," Springer, 2nd Edition,

Example of Opinion:
review segment on iPhone

"I bought an iPhone a few days ago.

It was such a nice phone.

The touch screen was really cool.

The voice quality was clear too.

However, my mother was mad with me as I did not tell her before I bought it.

She also thought the phone was too expensive, and wanted me to return it to the shop. ... "

Example of Opinion: review segment on iPhone

"(1) I bought an iPhone a few days ago.

(2) It was such a **nice** phone.

(3) The touch screen was really **cool**.

+Positive Opinion

(4) The voice quality was **clear** too.

(5) However, my mother was mad with me as I did not tell her before I bought it.

(6) She also thought the phone was too **expensive**, and wanted me to return it to the shop. … "

-Negative Opinion

Why are opinions important?

- "Opinions" are key influencers of our behaviors.

- Our beliefs and perceptions of reality are conditioned on how others see the world.

- Whenever we need to make a decision, we often seek out the opinion of others.
  In the past,
  - Individuals
    - Seek opinions from friends and family
  - Organizations
    - Use surveys, focus groups, opinion pools, consultants

SOCIAL LUMAscape

Word-of-mouth
on the Social media

- Personal experiences and opinions about anything in reviews, forums, blogs, micro-blog, Twitter.

- Posting at social networking sites, e.g., Facebook

- Comments about articles, issues, topics, reviews.

Social media + beyond

- Global scale
  - No longer – one's circle of friends.
- Organization internal data
  - Customer feedback from emails, call center
- News and reports
  - Opinions in news articles and commentaries

Applications of Opinion Mining

- **Businesses and organizations**
  - Benchmark products and services
  - Market intelligence
    - Business spend a huge amount of money to find consumer opinions using consultants, surveys, and focus groups, etc.
- **Individual**
  - Make decision to buy products or to use services
  - Find public opinions about political candidates and issues
- **Ads placements**: Place ads in the social media content
  - Place an ad if one praises a product
  - Place an ad from a competitor if one criticizes a product
- **Opinion retrieval**: provide general search for opinions.

Research Area of Opinion Mining

- Many names and tasks with difference objective and models
  - Sentiment analysis
  - Opinion mining
  - Sentiment mining
  - Subjectivity analysis
  - Affect analysis
  - Emotion detection
  - Opinion spam detection

Existing Tools
("Social Media Monitoring/Analysis")

- Radian 6
- Social Mention
- Overtone OpenMic
- Microsoft Dynamics Social Networking Accelerator
- SAS Social Media Analytics
- Lithium Social Media Monitoring
- RightNow Cloud Monitor

Existing Tools
("Social Media Monitoring/Analysis")

- Radian 6
- Social Mention
- Overtone OpenMic
- Microsoft Dynamics Social Networking Accelerator
- SAS Social Media Analytics
- Lithium Social Media Monitoring
- RightNow Cloud Monitor

http://www.tweetfeel.com

http://tweetsentiments.com/

Problem statement of Opinion Mining

- Two aspects of abstraction
  - Opinion definition
    - What is an opinion?
    - What is the structured definition of opinion?
  - Opinion summarization
    - Opinion are subjective
      - An opinion from a single person (unless a VIP) is often not sufficient for action
    - We need opinions from many people, and thus opinion summarization.

Abstraction (1) : what is an opinion?

- Id: **Abc123** on **5-1-2008** *"I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, ..."*

- One can look at this review/blog at the
  - Document level
    - Is this review + or -?
  - Sentence level
    - Is each sentence + or -?
  - Entity and feature/aspect level

Entity and aspect/feature level

- Id: **Abc123** on **5-1-2008** "*I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, …*"

- What do we see?
  - Opinion targets: entities and their features/aspects
  - Sentiments: positive and negative
  - Opinion holders: persons who hold the opinions
  - Time: when opinion are expressed

Two main types of opinions

- Regular opinions: Sentiment/Opinion expressions on some target entities
  - Direct opinions: sentiment expressions on one object:
    - "The touch screen is really cool."
    - "The picture quality of this camera is great"
  - Indirect opinions: comparisons, relations expressing similarities or differences (objective or subjective) of more than one object
    - "phone X is cheaper than phone Y." (objective)
    - "phone X is better than phone Y." (subjective)
- Comparative opinions: comparisons of more than one entity.
  - "iPhone is better than Blackberry."

Subjective and Objective

- Objective
  - An objective sentence expresses some factual information about the world.
  - "I returned the phone yesterday."
  - Objective sentences can implicitly indicate opinions
    - "The earphone broke in two days."
- Subjective
  - A subjective sentence expresses some personal feelings or beliefs.
  - "The voice on my phone was not so clear"
  - Not every subjective sentence contains an opinion
    - "I wanted a phone with good voice quality"
- ➔ Subjective analysis

A (regular) opinion

- Opinion (a restricted definition)
  - An opinion (regular opinion) is simply a positive or negative sentiment, view, attitude, emotion, or appraisal about an entity or an aspect of the entity from an opinion holder.
- Sentiment orientation of an opinion
  - Positive, negative, or neutral (no opinion)
  - Also called:
    - Opinion orientation
    - Semantic orientation
    - Sentiment polarity

Entity and aspect

- Definition of Entity:

  – An *entity e* is a product, person, event, organization, or topic.

  – e is represented as

    - A hierarchy of components, sub-components.
    - Each node represents a components and is associated with a set of attributes of the components

- An opinion can be expressed on any node or attribute of the node

- Aspects(features)

  – represent both components and attribute

Entity and aspect
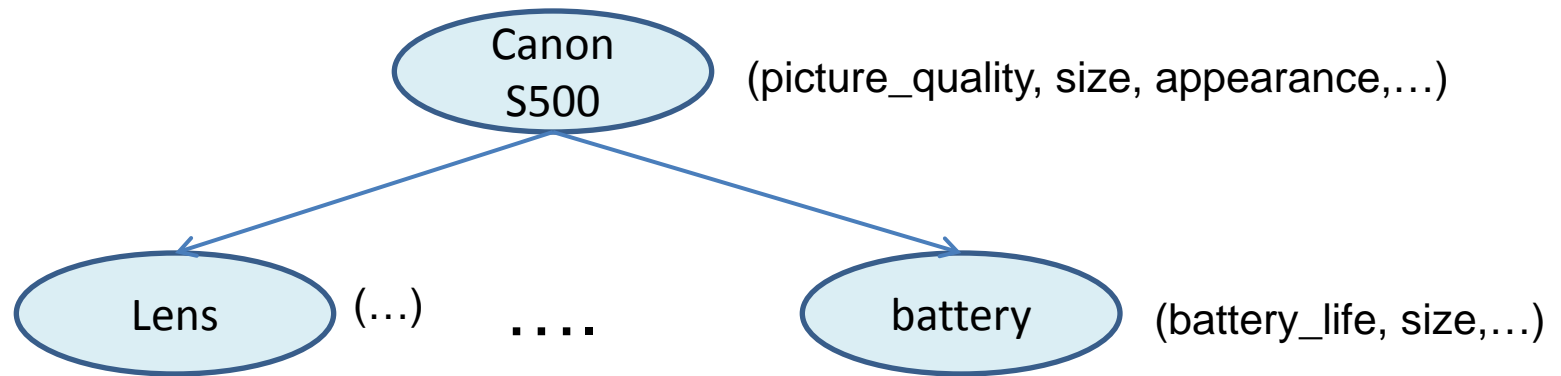
Canon S500 — (picture_quality, size, appearance,…)

Lens (…) …. battery (battery_life, size,…)

Opinion definition

- 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 from the opinion holder on feature of entity at time.
    so_{ijkl} is +ve, -ve, or neu, or more granular ratings

  - h_i is an opinion holder.

  - t_l is the time when the opinion is expressed.

Opinion definition

- 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 from the opinion holder on feature of entity at time.
    so_{ijkl} is +ve, -ve, or neu, or more granular ratings
  - h_i is an opinion holder.
  - t_l is the time when the opinion is expressed.

- $(e_j,\ a_{jk})$ is also called opinion target

Terminologies

- Entity: object
- Aspect: feature, attribute, facet
- Opinion holder: opinion source

- Topic: entity, aspect

- Product features, political issues

Subjectivity and Emotion

- Sentence subjectivity
  - An objective sentence presents some factual information, while a subjective sentence expresses some personal feelings, views, emotions, or beliefs.

- Emotion
  - Emotions are people's subjective feelings and thoughts.

Emotion

- Six main emotions
  - Love
  - Joy
  - Surprise
  - Anger
  - Sadness
  - Fear

Abstraction (2): opinion summary

- **With a lot of opinions, a summary is necessary.**
  - A multi-document summarization task
- For factual texts, summarization is to select the most important facts and present them in a sensible order while avoiding repetition
  - 1 fact = any number of the same fact
- But for opinion documents, it is different because opinions have a quantitative side & have targets
  - 1 opinion <> a number of opinions
  - Aspect-based summary is more suitable
  - Quintuples form the basis for opinion summarization

An aspect-based opinion summary

Cellular phone 1:

    Aspect: **GENERAL**

        Positive: 125   <individual review sentences>

        Negative: 7   <individual review sentences>

    Aspect: **Voice quality**

        Positive: 120   <individual review sentences>

        Negative: 8   <individual review sentences>

    Aspect: **Battery**

        Positive: 80   <individual review sentences>

        Negative: 12   <individual review sentences>
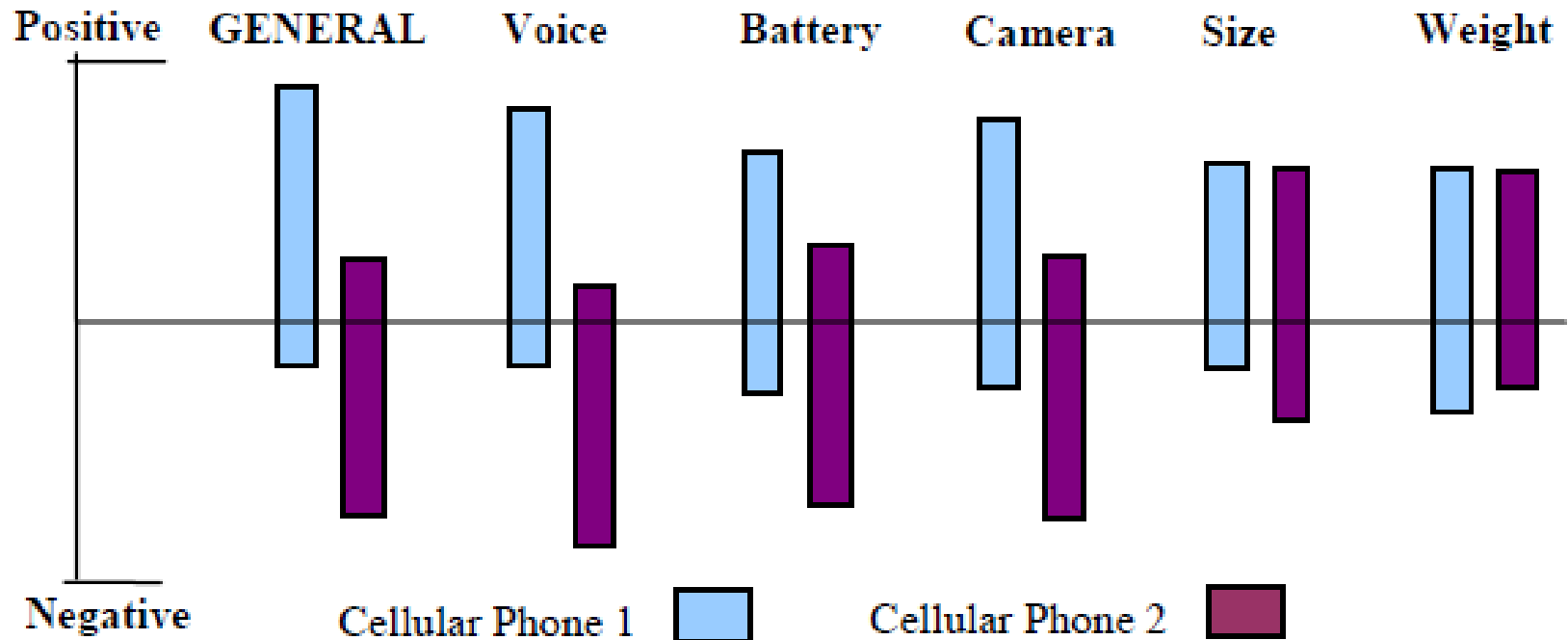
…

Visualization of aspect-based summaries of opinions
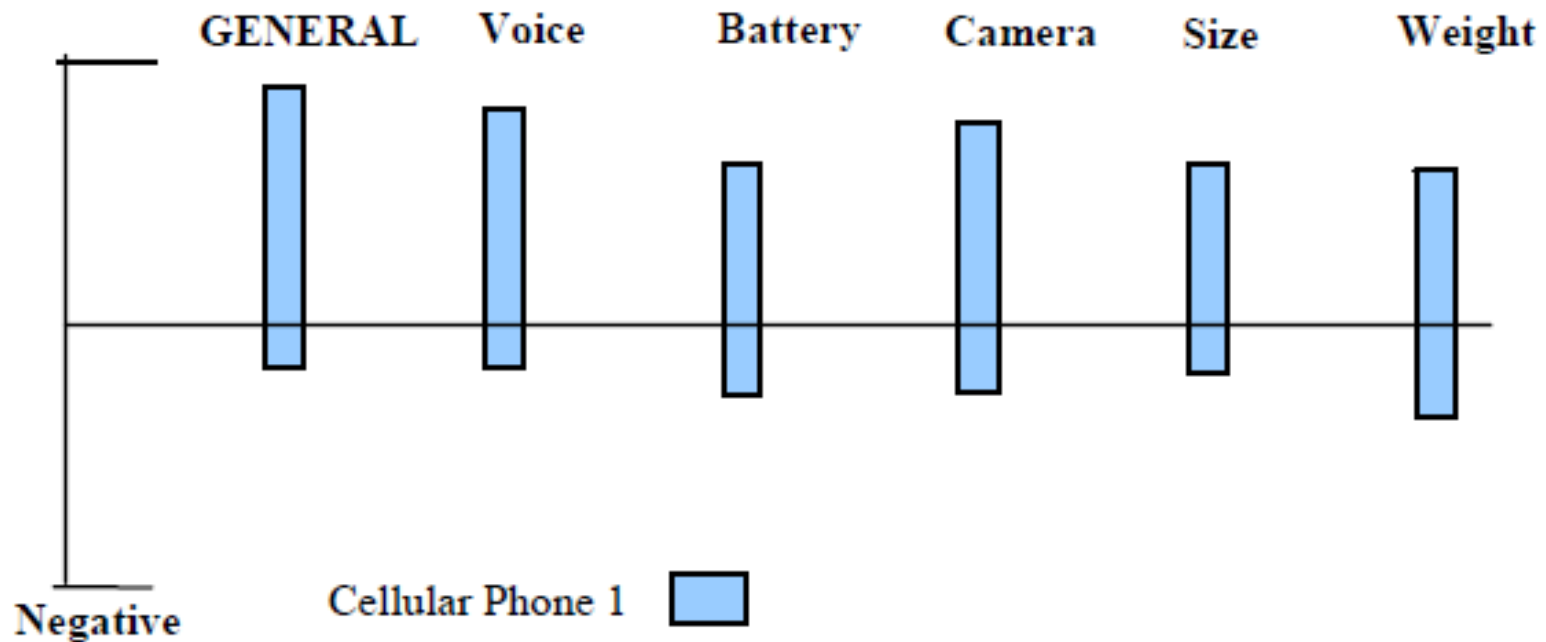
Visualization of aspect-based summaries of opinions

Classification Based on Supervised Learning

- Sentiment classification
  - Supervised learning Problem
  - Three classes
    - *Positive*
    - *Negative*
    - *Neutral*

Opinion words in Sentiment classification

- topic-based classification
  - topic-related words are important
    - e.g., *politics, sciences, sports*
- Sentiment classification
  - topic-related words are unimportant
  - **opinion words** (also called **sentiment words)**
    - **that indicate positive or negative opinions** are important,
    e.g., *great, excellent, amazing, horrible, bad, worst*

Features in Opinion Mining

- *Terms and their frequency*
  - *TF-IDF*
- *Part of speech (POS)*
  - Adjectives
- *Opinion words and phrases*
  - *beautiful, wonderful, good, and amazing are positive opinion words*
  - *bad, poor, and terrible are negative opinion words.*
  - opinion phrases and idioms, e.g., *cost someone an arm and a leg*
- *Rules of opinions*
- *Negations*
- *Syntactic dependency*

Rules of opinions

Syntactic template

<subj> passive-verb

<subj> active-verb

active-verb <dobj>

noun aux <dobj>

passive-verb prep <np>

Example pattern

<subj> was satisfied

<subj> complained

endorsed <dobj>

fact is <dobj>

was worried about <np>

《知網》情感分析用詞語集（beta版）

- "中英文情感分析用詞語集"
  - 包含詞語約 17887
- "中文情感分析用詞語集"
  - 包含詞語約 9193
- "英文情感分析用詞語集"
  - 包含詞語 8945

中文情感分析用詞語集

中文正面情感詞語	836
中文負面情感詞語	1254
中文正面評價詞語	3730
中文負面評價詞語	3116
中文程度級別詞語	219
中文主張詞語	38
Total	9193

中文情感分析用詞語集

- "正面情感"詞語
  - 如：
    愛，讚賞，快樂，感同身受，好奇，
    喝彩，魂牽夢縈，嘉許 ...
- "負面情感"詞語
  - 如：
    哀傷，半信半疑，鄙視，不滿意，不是滋味兒
    ，後悔，大失所望 ...

中文情感分析用詞語集

- "正面評價"詞語
  - 如：
    不可或缺，部優，才高八斗，沉魚落雁，
    催人奮進，動聽，對勁兒 ...

- "負面評價"詞語
  - 如：
    醜，苦，超標，華而不實，荒涼，混濁，
    畸輕畸重，價高，空洞無物 ...

中文情感分析用詞語集

- "程度級別"詞語
  - 1. "極其|extreme / 最|most"
    - 非常，極，極度，無以倫比，最為
  - 2. "很|very"
    - 多麼，分外，格外，著實
  - ...
- "主張"詞語
  - 1. {perception|感知}
    - 感覺，覺得，預感
  - 2. {regard|認為}
    - 認為，以為，主張

Web Data Mining

Exploring Hyperlinks, Contents, and Usage Data

1. Introduction
2. Association Rules and Sequential Patterns
3. Supervised Learning
4. Unsupervised Learning
5. Partially Supervised Learning
6. Information Retrieval and Web Search
7. Social Network Analysis
8. Web Crawling
9. Structured Data Extraction: Wrapper Generation
10. Information Integration
11. Opinion Mining and Sentiment Analysis
12. Web Usage Mining

Source: http://www.cs.uic.edu/~liub/WebMiningBook.html

Summary

- Opinion Mining
- Sentiment Analysis

References

- Bing Liu (2011) , "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data," Springer, 2nd Edition, 2011, http://www.cs.uic.edu/~liub/WebMiningBook.html
- Bo Pang and Lillian Lee (2008), Opinion mining and sentiment analysis, Foundations and Trends in Information Retrieval, 2:1-135, January 2008
- Wiltrud Kessler (2012), Introduction to Sentiment Analysis, http://www.ims.uni-stuttgart.de/~kesslewd/lehre/sentimentanalysis12s/introduction_sentimentanalysis.pdf