

Business Intelligence Trends

商業智慧趨勢

意見探勘與情感分析

(Opinion Mining and Sentiment Analysis)

1012BIT07

MIS MBA

Mon 6, 7 (13:10-15:00) Q407

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/>

2013-05-20

課程大綱 (Syllabus)

週次	日期	內容 (Subject/Topics)
1	102/02/18	商業智慧趨勢課程介紹 (Course Orientation for Business Intelligence Trends)
2	102/02/25	管理決策支援系統與商業智慧 (Management Decision Support System and Business Intelligence)
3	102/03/04	企業績效管理 (Business Performance Management)
4	102/03/11	資料倉儲 (Data Warehousing)
5	102/03/18	商業智慧的資料探勘 (Data Mining for Business Intelligence)
6	102/03/25	商業智慧的資料探勘 (Data Mining for Business Intelligence)
7	102/04/01	教學行政觀摩日 (Off-campus study)
8	102/04/08	個案分析一 (SAS EM 分群分析)： Banking Segmentation (Cluster Analysis – KMeans using SAS EM)
9	102/04/15	個案分析二 (SAS EM 關連分析)： Web Site Usage Associations (Association Analysis using SAS EM)

課程大綱 (Syllabus)

週次	日期	內容 (Subject/Topics)
10	102/04/22	期中報告 (Midterm Presentation)
11	102/04/29	個案分析三 (SAS EM 決策樹、模型評估) : Enrollment Management Case Study (Decision Tree, Model Evaluation using SAS EM)
12	102/05/06	個案分析四 (SAS EM 迴歸分析、類神經網路) : Credit Risk Case Study (Regression Analysis, Artificial Neural Network using SAS EM)
13	102/05/13	文字探勘與網路探勘 (Text and Web Mining)
14	102/05/20	意見探勘與情感分析 (Opinion Mining and Sentiment Analysis)
15	102/05/27	商業智慧導入與趨勢 (Business Intelligence Implementation and Trends)
16	102/06/03	商業智慧導入與趨勢 (Business Intelligence Implementation and Trends)
17	102/06/10	期末報告1 (Term Project Presentation 1)
18	102/06/17	期末報告2 (Term Project Presentation 2)

Outline

- Social Word-of-Mouth
- Opinion Mining and Sentiment Analysis
- Social Media Monitoring/Analysis
- Resources of Opinion Mining
- Opinion Spam Detection

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

Social Media and the Voice of the Customer

- Listen to the **Voice of the Customer (VoC)**
 - Social media can give companies a torrent of highly valuable customer feedback.
 - Such input is largely free
 - Customer feedback issued through social media is qualitative data, just like the data that market researchers derive from focus group and in-depth interviews
 - Such qualitative data is in digital form – in text or digital video on a web site.

Listen and Learn

Text Mining for VoC

- Categorization
 - Understanding what topics people are talking or writing about in the unstructured portion of their feedback.
- Sentiment Analysis
 - Determining whether people have **positive**, **negative**, or **neutral** views on those topics.

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, etc., 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

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

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

Word-of-mouth

Voice of the Customer

- 1. Attensity
 - Track social sentiment across brands and competitors
 - <http://www.attensity.com/home/>
- 2. Clarabridge
 - Sentiment and Text Analytics Software
 - <http://www.clarabridge.com/>

Attensity: Track social sentiment across brands and competitors

<http://www.attensity.com/>

The screenshot shows the Attensity website homepage. At the top, there is a navigation bar with the Attensity logo, a language selector set to 'English', and links for 'Contact', 'Resources', 'Support', and 'Blog'. A search bar is also present. Below the navigation, there are tabs for 'Products', 'Solutions', 'Services', 'Customers', and 'Partners'. The main content area features a large central banner with the headline 'Your real-time window into the social web.' and a quote from Yahoo! stating, 'Teaming with a leading analytics provider like Attensity offers Yahoo! a great opportunity to deliver the key news and analysis that matter.' To the left of the banner is a vertical menu with categories: 'Social Analytics', 'Social Response', 'Customer Analytics', 'Industry Solutions', and 'Why Attensity'. Below the banner, there are several tiles: 'Attensity for Marketing', 'Attensity for Customer Service', 'Attensity for IT', 'Success Story' (featuring JetBlue Airways), 'About Attensity', and 'Watch Video' (featuring a 'Command Center Video'). The bottom of the page shows a URL bar with 'www.attensity.com/home/#fragment-1' and a 'DOWNLOAD NOW' button.

<http://www.youtube.com/watch?v=4goxmBEg2lw#/>

Clarabridge: Sentiment and Text Analytics Software

<http://www.clarabridge.com/>

The image shows a screenshot of a web browser displaying the Clarabridge website. The browser's address bar shows the URL www.clarabridge.com. The website features a navigation menu with links for Home, About Us, News & Events, Blog, Login, and Contact Us. Below the navigation is a dark grey menu with categories: WHY TEXT ANALYTICS, PRODUCT, SERVICES, CUSTOMERS, PARTNERS, RESOURCES, and ABOUT US. The main content area has a blue background with the headline "The First Sentiment and Text Analytics Solution Built Specifically for Business." and a sub-headline "The Clarabridge sentiment and text analytics software provides enterprises with a universal view of their customers." A "Learn more about how Clarabridge works >" link is positioned at the bottom right of this section. Below the main content is a "Customers" section displaying logos for Nissan, Best Buy, Marriott, Sage, H&R Block, Choice Hotels International, Wendy's, GWLORD HOTELS, BE Aerospace, and Dell, followed by a "More >" link. The footer contains three promotional boxes: "Clarabridge Text Analytics", "Choose Your Edition" (with a sub-section for "Clarabridge for Enterprises" described as ideal for enterprise-class text analytics solutions), and "Clarabridge Webinar" (presented by Hypatia Research Group on social media).

<http://www.youtube.com/watch?v=IDHudt8M9P0>

<http://www.radian6.com/>

Social Media Monitoring x

www.radian6.com

Country 1 888 672 3426 About Radian6 Contact CUSTOMER LOGIN Search GO

salesforce **radian6**

How We Help What We Sell See Demo Free Resources Training & Support

The Social Enterprise.
Get closer to your customer.
Learn how >

Have Us Contact You

Live Demo

Free Trial

Chat & find out more.

Offline. Leave us a message

Sales The social web is a goldmine of untapped sales opportunities. Let us help you realize your potential. [Learn more >](#)

Marketing Brands are now the sum of the conversations about them. We can help you hear what's being said. [Learn more >](#)

Customer Service Take your customer service where your consumers are gathering. Respond to issues voiced on the social web. [Learn more >](#)

Newsletter Sign up and get the regular Radian6 goods. Enter email address GO

Mashable named Radian6's Co-founder Chris Ramsey one of five masterminds redefining social media

JUST Get the Skinny

WEBINAR / June 7th at 2pm est

CASE STUDY

radian6 Community

http://www.youtube.com/watch?feature=player_embedded&v=8i6Exg3Urg0

The screenshot shows a web browser window with the URL www.sas.com/software/customer-intelligence/social-media-analytics/. The page features the SAS logo and tagline "THE POWER TO KNOW." at the top left, with the text "Providing software solutions since 1976" below it. A navigation bar includes links for "Log In", "Worldwide Sites", "Contact Us", and "Follow Us". A secondary navigation bar lists "NEWS", "EVENTS", "CONSULTING", "CAREERS", and "RESOURCE CENTER". A search bar is located on the right side of the page.

The main content area is titled "PRODUCTS & SOLUTIONS / SOCIAL MEDIA ANALYTICS". On the left, there is a "Products and Solutions" sidebar with a list of categories: Industries, Small and Midsize Business, Nonprofit Organizations, Analytics, Business Analytics, Business Intelligence, Customer Intelligence (with sub-items: Strategy & Planning, Information & Analytics, Orchestration & Interaction, Customer Experience, Customer Experience Analytics, Social Media Analytics, Web Analytics), Financial Intelligence, Foundation Tools, Fraud & Financial Crimes, Governance, Risk & Compliance, High-Performance Analytics, Human Capital Intelligence, Information Management, and IT & CIO Enablement.

The main content area features a section for "SAS® Social Media Analytics" with the sub-headline "Integrate, archive, analyze and act on online conversations". Below this is a navigation bar with tabs for "Overview", "Benefits", "Features", "Demos & Screenshots", and "System Requirements". The "Overview" tab is selected.

The "Overview" section contains the following text: "SAS Social Media Analytics is an enterprise-hosted, on-demand solution that integrates, archives, analyzes and enables organizations to act on intelligence gleaned from online conversations on professional and consumer-generated media sites. It enables you to attribute online conversations to specific parts of your business, allowing accelerated responses to marketplace shifts." Below this text is a paragraph: "Based on your unique business challenges and enterprise goals, SAS can provide a tailored implementation that's hosted and managed by [SAS Solutions OnDemand](#)."

Below the overview text is a "Benefits" section with a list of bullet points: "Analyze conversation data.", "Identify advocates of, and threats to, corporate reputation and brand.", "Quantify interaction among traditional media/campaigns and social media activity.", and "Establish a platform for social CRM strategy."

To the right of the main content area is a "Questions?" section with a photo of a woman and links for "Phone" and "Contact Form". Below that is a "White Paper" section titled "Text Analytics for Social Media: Evolving Tools for an Evolving Environment" with a "Download Now" button. At the bottom right is a "SAS® Social Media Analytics" section with a link for "Overview" and a "RESOURCES" section with links for "Fact Sheet (PDF)", "Solution Brief (PDF)", and "White Papers".

In the center of the page, there is a quote: "The great thing about SAS is that it's so powerful and has such a broad offering." attributed to Jonathan Prantner, Manager of Statistics Organic. Below the quote is a "Read full story" link. Below the quote is a "Product Demo" section with a video player showing a screenshot of the SAS Social Media Analytics interface.

The screenshot shows a web browser window with the URL www.tweetfeel.com/index.php#iPhone4s. The page features the 'tweetfeel' logo with a blue bird icon. A search bar contains the text 'iPhone4s' and a yellow 'Search' button. Below the search bar, it displays 'Try some Twitter trends: [Tomorrow is June](#) [H&M](#) [Defense of Marriage Act](#) [Diddy's](#) [Bloomberg](#) [UCLA](#) [ESPN](#)'. A sentiment analysis graphic shows a green smiley face with '40' below it, a red frowny face with '41' below it, an equals sign, and '51%' in red. A text block reads: '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.](#)' Below this are six tweet snippets, each with a small profile picture and text mentioning 'iPhone4s' and 'wtf'. The footer contains links for 'Read our FAQ', 'Legal Stuff', '100% Guarantee', and 'Share', along with social media icons and logos for 'conversion' and 'Powered by twitter'.

Browser window: Tweet Sentiments - Know ' x
Address bar: tweetsentiments.com/analyze?utf8=✓&q=iphone4s&topic=true&commi ☆

Lang: Eng
Eng **tweet** v0.8.6 Sentiments
Dashboard Analyze Recent Top 1000 Users Maps Charts About Login


Tweet Sentiment Analysis

iphone4s
User Topic
Analyze Tweets

Feedback

Sentiments
Index[0-100]: 50.00
Positive: 8
Negative: 1
Neutral: 13
Total Analyzed: 0

Profile
Education:
Flamboyance:
Slang:
Gender:
Age:



50.00

1 minute ago by Flow__Show
Niggie, if i dont jailbreak my Iphone4s, u mean i cant play Street fighter4? wtf @louistekneeq

1 minute ago by designer_sayaka
リフォロー100%です！ #Arashi #韓fan #wstcg #板野友美 #followme #アメブロ #kimiboku #ntb #autofollow #apple #mt2 #相互フォローの輪 #iphone4s #Arashi #twitter #前田敦子

1 minute ago by abhay01007
Rumor has it that #Apple #iPhone5 will have a 4" screen (compared 2 3.5" in #iPhone4S), giving direct competition 2... <http://t.co/zsIGPXbX>

1 minute ago by SuckMy_TwitNuts
RT @BallMe_Dollar: The iPhone4S look better then the iPhone5.

2 minutes ago by thementaldawg
#iphoneography #iphonography #bahrain #iphone4s #photography #sunrise #scenery #beautiful #sun <http://t.co/Ee4VNmsL>

i-Buzz網路口碑研究中心. 網 x

www.i-buzz.com.tw/index.asp

iBuzz中國 | 關於我們

歡迎光臨。請 [加入會員](#) 或 [登錄](#)

iBuzz research 艾博司網路口碑研究中心
Internet Buzz Research Center

-議題分析-

Home 議題分析 排行榜 研究專題 數據報告 產品服務 解決方案

7-ELEVEN

連結類型	百分比
照片連結	73%
文章連結	10%
影片連結	17%
純發文	0%

議題分析

血汗醫院全民皆憤 網路熱門

研究專題

便利商店衝人氣 臉書加持少

數據報告

永慶房屋穩據房仲龍頭 資訊

f 臉書粉絲團排行榜

政治人物 名人 演藝人員

馬英九總統	蔡英文 Tsai	蘇貞昌
粉絲數 1419591	粉絲數 669928	粉絲數 317035

排行榜

Traffic 交通

	第1名
FORD (FIESTA)	
	第2名
PORSCHE (911)	
	第3名
SUBARU (legacy)	

[更多](#)

Traffic 交通

	第1名
--	-----

進口車

- 吵了這麼久, 就是不要「超速」超車, 還不懂 (224)
- 讓一下, 很難嗎。。。 (184)
- 請問大家認為這樣要賠多少?? (148)
- 當Audi 一直被鍵盤魔人攻擊時, 表示已經威 (134)
- ELANTRA系列 vs ALTIS 系列之比較 (109)

[更多](#)

高級進口車

- 吵了這麼久, 就是不要「超速」超車, 還不懂 (224)

研究專題

便利商店衝人氣 臉書加持少不得

研究專題

LUMIX與Samsung相機 口碑比拼 粉絲只顧「按讚」不「留言」

社交媒介監看

了解社群監測

OpView 服務總覽 | 意藍科技 x

www.eland.com.tw/solutions

eLAND Cloud Services 意藍科技

產品 雲端服務 壹銀標專區 我們的客戶 新聞與活動 支援 關於意藍

OpView Service

您的輿論觀測站

連上OpView · 品牌形象 · 輿論觀點監測好輕鬆

選戰情報

OpView 服務總覽

OpView 服務總覽

什麼是OpView服務？

OpView是協助您蒐集、處理、分析網路資訊的雲端服務。

OpView服務，猶如提供您無盡網路情報能量的發電廠。以雲端架構為基礎，OpView服務蒐集、處理、分析各類型網路資訊與情報，並以雲端服務平台供客戶使用。OpView服務的資料涵蓋範圍包括台灣最具代表性的新聞網站、

相關文件下載

- OpView Insight DM
- OpView品牌口碑雷達 DM
- OpView品牌口碑雷達訂購單

更多

Sentiment Analysis

- Sentiment
 - A thought, view, or attitude, especially one based mainly on emotion instead of reason
- Sentiment Analysis
 - opinion mining
 - use of natural language processing (NLP) and computational techniques to automate the extraction or classification of sentiment from typically unstructured text

Applications of Sentiment Analysis

- Consumer information
 - Product reviews
- Marketing
 - Consumer attitudes
 - Trends
- Politics
 - Politicians want to know voters' views
 - Voters want to know politicians' stances and who else supports them
- Social
 - Find like-minded individuals or communities

Sentiment detection

- How to interpret features for sentiment detection?
 - Bag of words (IR)
 - Annotated lexicons (WordNet, SentiWordNet)
 - Syntactic patterns
- Which features to use?
 - Words (unigrams)
 - Phrases/n-grams
 - Sentences

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

Sentiment Analysis

vs.

Subjectivity Analysis

Sentiment Analysis	Subjectivity Analysis
Positive	Subjective
Negative	
Neutral	Objective

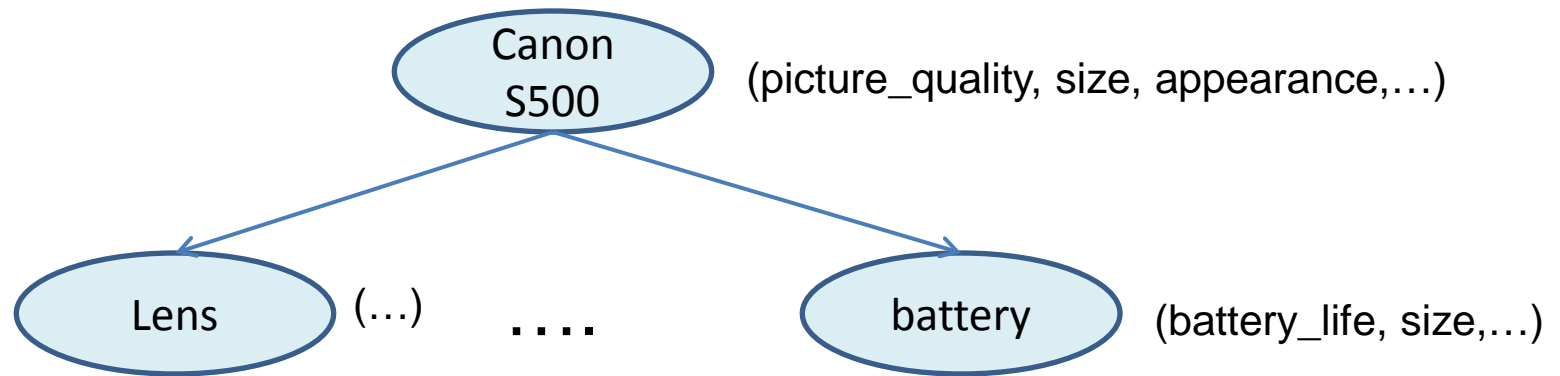
A (regular) opinion

- **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



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 \leftrightarrow 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>

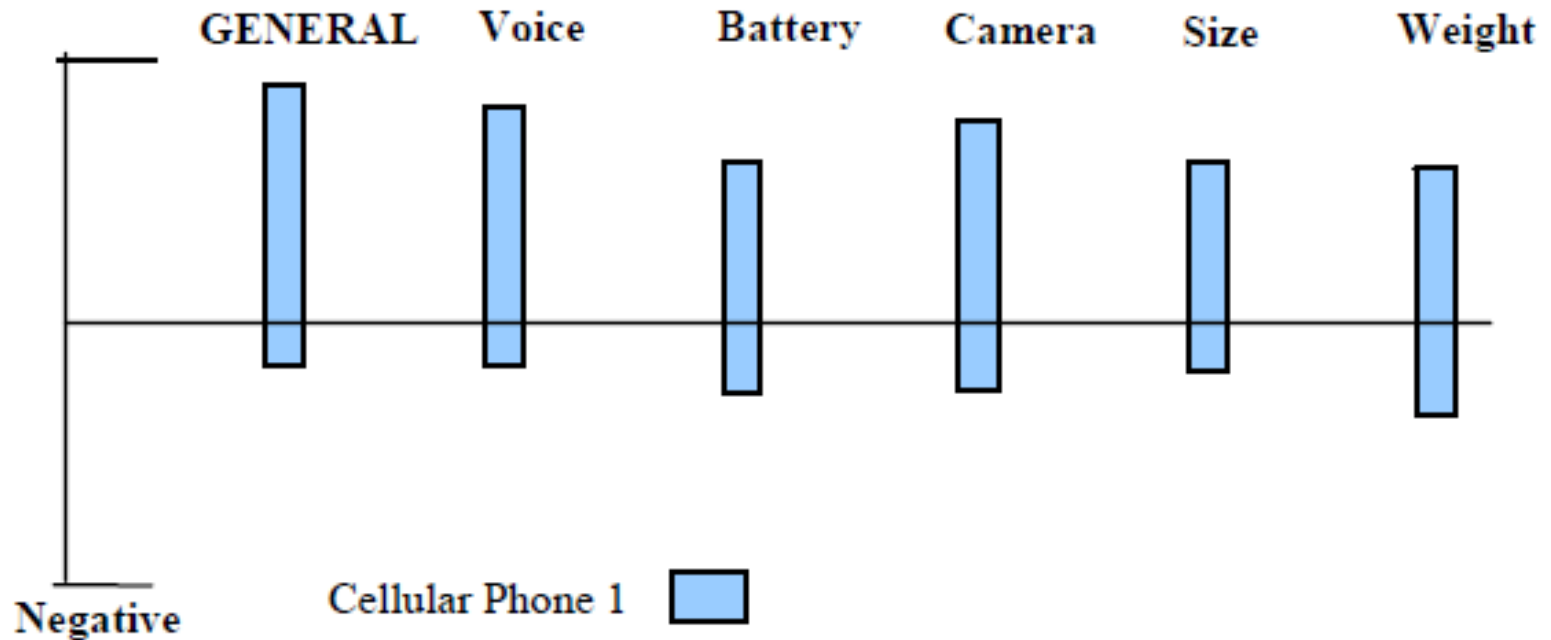
Aspect: **Battery**

Positive: 80 <individual review sentences>

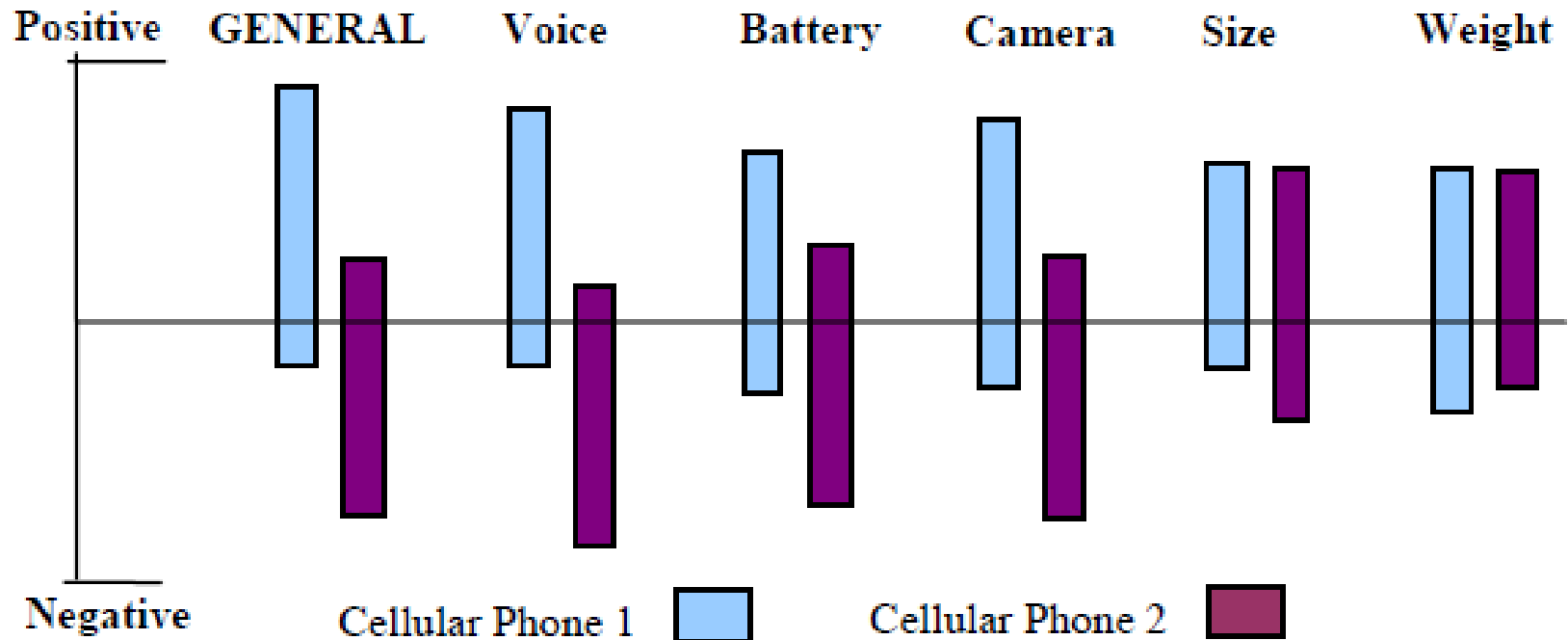
Negative: 12 <individual review sentences>

...

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>

A Brief Summary of **Sentiment Analysis** Methods

Study	Analysis Task	Sentiment Identification		Sentiment Aggregation		Nature of Measure
		Method	Level	Method	Level	
Hu and Li, 2011	Polarity	ML (Probabilistic model)	Snippet			Valence
Li and Wu, 2010	Polarity	Lexicon/Rule	Phrase	Sum	Snippet	Valence
Thelwall et al., 2010	Polarity	Lexicon/Rule	Sentence	Max & Min	Snippet	Range
Boiy and Moens, 2009	Both	ML (Cascade ensemble)	Sentence			Valence
Chung 2009	Polarity	Lexicon	Phrase	Average	Sentence	Valence
Wilson, Wiebe, and Hoffmann, 2009	Both	ML (SVM, AdaBoost, Rule, etc.)	Phrase			Valence
Zhang et al., 2009	Polarity	Lexicon/Rule	Sentence	Weighted average	Snippet	Valence
Abbasi, Chen, and Salem, 2008	Polarity	ML (GA + feature selection)	Snippet			Valence
Subrahmanian and Reforgiato, 2008	Polarity	Lexicon/Rule	Phrase	Rule	Snippet	Valence
Tan and Zhang 2008	Polarity	ML (SVM, Winnow, NB, etc.)	Snippet			Valence
Airoldi, Bai, and Padman, 2007	Polarity	ML (Markov Blanket)	Snippet			Valence
Das and Chen, 2007	Polarity	ML (Bayesian, Discriminate, etc.)	Snippet	Average	Daily	Valence
Liu et al., 2007	Polarity	ML (PLSA)	Snippet			Valence
Kennedy and Inkpen, 2006	Polarity	Lexicon/Rule, ML (SVM)	Phrase	Count	Snippet	Valence
Mishne 2006	Polarity	Lexicon	Phrase	Average	Snippet	Valence
Liu et al., 2005	Polarity	Lexicon/Rule	Phrase	Distribution	Object	Range
Mishne 2005	Polarity	ML (SVM)	Snippet			Valence
Popescu and Etzioni 2005	Polarity	Lexicon/Rule	Phrase			Valence
Efron 2004	Polarity	ML (SVN, NB)	Snippet			Valence
Wilson, Wiebe, and Hwa, 2004	Both	ML (SVM, AdaBoost, Rule, etc.)	Sentence			Valence
Nigam and Hurst 2004	Polarity	Lexicon/Rule	Chunk	Rule	Sentence	Valence
Dave, Lawrence, and Pennock, 2003	Polarity	ML (SVM, Rainbow, etc.)	Snippet			Valence
Nasukawa and Yi 2003	Polarity	Lexicon/Rule	Phrase	Rule	Sentence	Valence
Yi et al., 2003	Polarity	Lexicon/Rule	Phrase	Rule	Sentence	Valence
Yu and Hatzivassiloglou 2003	Both	ML (NB) + Lexicon/Rule	Phrase	Average	Sentence	Valence
Pang, Lee, and Vaithyanathan 2002	Polarity	ML (SVM, MaxEnt, NB)	Snippet			Valence
Subasic and Huettner 2001	Polarity	Lexicon/Fuzzy logic	Phrase	Average	Snippet	Valence
Turney 2001	Polarity	Lexicon/Rule	Phrase	Average	Snippet	Valence

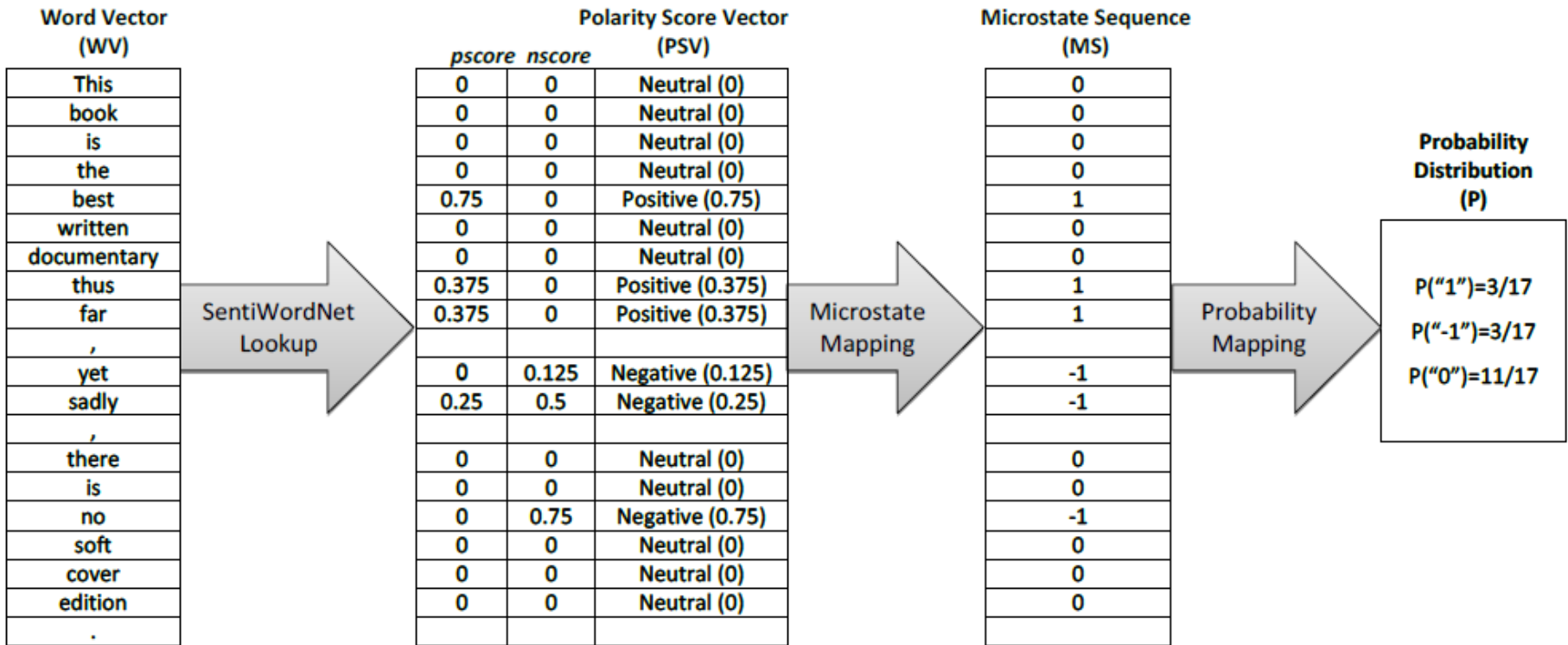
(Both = Subjectivity and Polarity; ML= Machine Learning; Lexicon/Rule= Lexicon enhanced by linguistic rules)

Word-of-Mouth (WOM)

- “This book is the best written documentary thus far, yet sadly, there is no soft cover edition.”
- “This book is the **best** written documentary **thus far**, **yet** **sadly**, there is **no** soft cover edition.”

	Word	POS
This	This	DT
book	book	NN
is	is	VBZ
the	the	DT
best	best	JJS
written	written	VCN
documentary	documentary	NN
thus	thus	RB
far	far	RB
,	,	,
yet	yet	RB
sadly	sadly	RB
,	,	,
there	there	EX
is	is	VBZ
no	no	DT
soft	soft	JJ
cover	cover	NN
edition	edition	NN
.	.	.

Conversion of text representation



Datasets of Opinion Mining

- Blog06
 - 25GB TREC test collection
 - [http://ir.dcs.gla.ac.uk/test collections/access to data.html](http://ir.dcs.gla.ac.uk/test%20collections/access%20to%20data.html)
- Cornell movie-review datasets
 - <http://www.cs.cornell.edu/people/pabo/movie-review-data/>
- Customer review datasets
 - <http://www.cs.uic.edu/~liub/FBS/CustomerReviewData.zip>
- Multiple-aspect restaurant reviews
 - <http://people.csail.mit.edu/bsnyder/naacl07>
- NTCIR multilingual corpus
 - NTCIR Multilingual Opinion-Analysis Task (MOAT)

Lexical Resources of Opinion Mining

- SentiWordnet
 - <http://sentiwordnet.isti.cnr.it/>
- General Inquirer
 - <http://www.wjh.harvard.edu/~inquirer/>
- OpinionFinder's Subjectivity Lexicon
 - <http://www.cs.pitt.edu/mpqa/>
- NTU Sentiment Dictionary (NTUSD)
 - <http://nlg18.csie.ntu.edu.tw:8080/opinion/>
- HowNet Sentiment
 - http://www.keenage.com/html/c_bulletin_2007.htm

Example of SentiWordNet

POS	ID	PosScore	NegScore	SynsetTerms	Gloss
a	00217728	0.75	0	beautiful#1	delighting the senses or exciting intellectual or emotional admiration; "a beautiful child"; "beautiful country"; "a beautiful painting"; "a beautiful theory"; "a beautiful party"
a	00227507	0.75	0	best#1	(superlative of `good') having the most positive qualities; "the best film of the year"; "the best solution"; "the best time for planting"; "wore his best suit"
r	00042614	0	0.625	unhappily#2	sadly#1 in an unfortunate way; "sadly he died before he could see his grandchild"
r	00093270	0	0.875	woefully#1	sadly#3 lamentably#1 deplorably#1 in an unfortunate or deplorable manner; "he was sadly neglected"; "it was woefully inadequate"
r	00404501	0	0.25	sadly#2	with sadness; in a sad manner; "'She died last night,' he said sadly"

《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

Opinion Spam Detection

- Opinion Spam Detection: Detecting Fake Reviews and Reviewers
 - Spam Review
 - Fake Review
 - Bogus Review
 - Deceptive review
 - Opinion Spammer
 - Review Spammer
 - Fake Reviewer
 - Shill (Stooge or Plant)

Opinion Spamming

- Opinion Spamming
 - "illegal" activities
 - e.g., writing fake reviews, also called shilling
 - try to mislead readers or automated opinion mining and sentiment analysis systems by giving undeserving positive opinions to some target entities in order to promote the entities and/or by giving false negative opinions to some other entities in order to damage their reputations.

Forms of Opinion spam

- fake reviews (also called bogus reviews)
- fake comments
- fake blogs
- fake social network postings
- deceptions
- deceptive messages

Fake Review Detection

- Methods
 - supervised learning
 - pattern discovery
 - graph-based methods
 - relational modeling
- Signals
 - Review content
 - Reviewer abnormal behaviors
 - Product related features
 - Relationships

Professional Fake Review Writing Services (some Reputation Management companies)

- Post positive reviews
- Sponsored reviews
- Pay per post
- Need someone to write positive reviews about our company (budget: \$250-\$750 USD)
- Fake review writer
- Product review writer for hire
- Hire a content writer
- Fake Amazon book reviews (hiring book reviewers)
- People are just having fun (not serious)

SponsoredReviews.com x www.sponsoredreviews.com

 **SponsoredReviews.com**
Bloggers Earn Cash, Advertisers Build Buzz!

[Members Login](#)
Sales Toll-Free (877) 360-3683

[Home](#) [Sign-Up](#) [FAQs](#) [Our Blog](#) [Contact Us](#)

SponsoredReviews connects bloggers with SEO's, Marketers, and Advertisers looking to build Links, Traffic and Buzz.

Direct Traffic.
Millions of people read blogs every day. Paying for posts puts the spotlight on your company and will generate tons of targeted traffic.

Buzz & Branding.
The more bloggers talk about your site the better. Many blogs syndicate stories they see on other sites. A couple well timed sponsored posts has the potential to generate a flurry of other post being written.

Search Engine Rankings.
Every post has links back to your site. Getting links from quality blogs will increase your link popularity and will help your site rank better in the search engines.

Valuable Feedback.
Getting Reviewed by bloggers will provide you with valuable feedback that you can use to better understand your audience and customers.

Advertisers
Start Here.



- Announce your products, services, websites, and ideas to the world!
- Tap into the power of the blogosphere to build traffic, links and valuable feedback.

[Free Sign Up](#) [Read More](#)

Bloggers
Earn Cash.



- Earn cash by writing honest posts about our advertiser's products and services.
- Write posts in your own tone and style, and gear them to your audience's interest.

[Free Sign Up](#) [Read More](#)

How it works:  Advertiser  Blogger

PayPerPost : Blog Marke x

https://payperpost.com

payperpost

advertisers bloggers ethics about login

advertisers

Hire bloggers to blog about your company, service or website. PayPerPost gives you access to a diverse pool of bloggers from all over the world. Make offers, negotiate deals and approve posts.

[signup now](#)

bloggers

Make money blogging! PayPerPost lets you pick your advertisers, name your own price and negotiate your own deals. You can get paid to blog on virtually any subject. Sign up below!

[signup now](#)

see how it *works*

[click here and watch the video](#)

customer testimonial

"PayPerPost has been instrumental in helping our company streamline our various product awareness campaigns."

-C. Litchfield

1 (877) 916 POST

Need someone to write x

www.freelancer.com/projects/Forum-Posting-Reviews/Need-someone-write-post-positive.html

freelancer

Help Login sign up login

Post Project Find Freelancers Browse Projects Post Contest Search for Freelancers, Projects...

Need someone to write and post positive reviews

Like 0 Send Tweet 0 +1 0 Share

Bids	Avg Bid (USD)	Project Budget (USD)	CLOSED
10	N/A	\$250 - \$750	

Project Description:
 We need someone to write and post positive reviews about our company on websites. Please send an example of a review you would post for any company. We can also send examples of comments our customers have sent us to use and refer too as well

This is a long term project, so if it works out there will be a healthy amount of work. Please reply back with all your experience and how much you would charge per post.

thank you.

Skills required:
 Publicación en foros, Opiniones

Project posted by:
 dvel
 ★★★★★ 5.0 (1 Review)
 VERIFIED

Follow

Your ad could *From \$100/week*

Papers on Opinion Spam Detection

1. Arjun Mukherjee, Bing Liu, and Natalie Glance. Spotting Fake Reviewer Groups in Consumer Reviews. International World Wide Web Conference (WWW-2012), Lyon, France, April 16-20, 2012.
2. Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu. Identify Online Store Review Spammers via Social Review Graph. ACM Transactions on Intelligent Systems and Technology, accepted for publication, 2011.
3. Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu. Review Graph based Online Store Review Spammer Detection. ICDM-2011, 2011.
4. Arjun Mukherjee, Bing Liu, Junhui Wang, Natalie Glance, Nitin Jindal. Detecting Group Review Spam. WWW-2011 poster paper, 2011.
5. Nitin Jindal, Bing Liu and Ee-Peng Lim. "Finding Unusual Review Patterns Using Unexpected Rules" Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM-2010, short paper), Toronto, Canada, Oct 26 - 30, 2010.
6. Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu and Hady Lauw. "Detecting Product Review Spammers using Rating Behaviors." Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM-2010, full paper), Toronto, Canada, Oct 26 - 30, 2010.
7. Nitin Jindal and Bing Liu. "Opinion Spam and Analysis." Proceedings of First ACM International Conference on Web Search and Data Mining (WSDM-2008), Feb 11-12, 2008, Stanford University, Stanford, California, USA.
8. Nitin Jindal and Bing Liu. "Review Spam Detection." Proceedings of WWW-2007 (poster paper), May 8-12, Banff, Canada.

Summary

- Social Word-of-Mouth
- Opinion Mining and Sentiment Analysis
- Social Media Monitoring/Analysis
- Resources of Opinion Mining
- Opinion Spam Detection

References

- Bing Liu (2011) , “Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data,” 2nd Edition, Springer.
<http://www.cs.uic.edu/~liub/WebMiningBook.html>
- Bing Liu (2013), Opinion Spam Detection: Detecting Fake Reviews and Reviewers, <http://www.cs.uic.edu/~liub/FBS/fake-reviews.html>
- Bo Pang and Lillian Lee (2008), "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval 2(1-2), pp. 1–135, 2008.
- Wiltrud Kessler (2012), Introduction to Sentiment Analysis, http://www.ims.uni-stuttgart.de/~kesslewd/lehre/sentimentanalysis12s/introduction_sentimentanalysis.pdf
- Z. Zhang, X. Li, and Y. Chen (2012), "Deciphering word-of-mouth in social media: Text-based metrics of consumer reviews," ACM Trans. Manage. Inf. Syst. (3:1) 2012, pp 1-23.
- Efraim Turban, Ramesh Sharda, Dursun Delen (2011), “Decision Support and Business Intelligence Systems,” Pearson , Ninth Edition, 2011.