Social Media Marketing Management 社會媒體行銷管理

社群ロ碑與社群網路探勘 (Social Word-of-Mouth and Web Mining on Social Media)

1002SMMM10 TLMXJ1A Tue 12,13,14 (19:20-22:10) D325

<u>Min-Yuh Day</u> <u>戴敏育</u> 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/

1

課程大綱 (Syllabus)

- 週次 日期 內容(Subject/Topics)
- 1 102/02/19 社會媒體行銷管理課程介紹 (Course Orientation of Social Media Marketing Management)
- 2 102/02/26 社群網路 (Social Media: Facebook, Youtube, Blog, Microblog)
- 3 102/03/05 社群網路行銷 (Social Media Marketing)
- 4 102/03/12 行銷管理 (Marketing Management)
- 5 102/03/19 社群網路服務與資訊系統理論 (Theories of Social Media Services and Information Systems)
- 6 102/03/26 行銷理論 (Marketing Theories)
- 7 102/04/02 教學行政觀摩日 (Off-campus study)
- 8 102/04/09 行銷管理論文研討 (Paper Reading on Marketing Management)
- 9 102/04/16 社群網路行為研究 (Behavior Research on Social Media)

課程大綱 (Syllabus)

- 週次 日期 內容(Subject/Topics)
- 10 102/04/23 期中報告 (Midterm Presentation)
- 11 102/04/30 社群網路商業模式 [Invited Speaker: Dr. Rick Cheng-Yu Lu] (Business Models and Issues of Social Media)
- 12 102/05/07 社群網路策略 (Strategy of Social Media)
- 13 102/05/14 社群口碑與社群網路探勘 (Social Word-of-Mouth and Web Mining on Social Media)
- 14 102/05/21 社群網路論文研討 (Paper Reading on Social Media)
- 15 102/05/28 探索性因素分析 (Exploratory Factor Analysis)
- 16 102/06/04 確認性因素分析 (Confirmatory Factor Analysis)
- 17 102/06/11 期末報告1 (Term Project Presentation 1)
- 18 102/06/18 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.

5

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

6

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.

9

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

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

+Positive

Opinion

Opinion

- "(1) I bought an <u>iPhone</u> a few days ago.
- (2) It was such a **nice** phone.
- (3) The touch screen was really cool.
- (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 <u>expensive</u>, and wanted me to return it to the shop. ... "-Negative

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

14

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

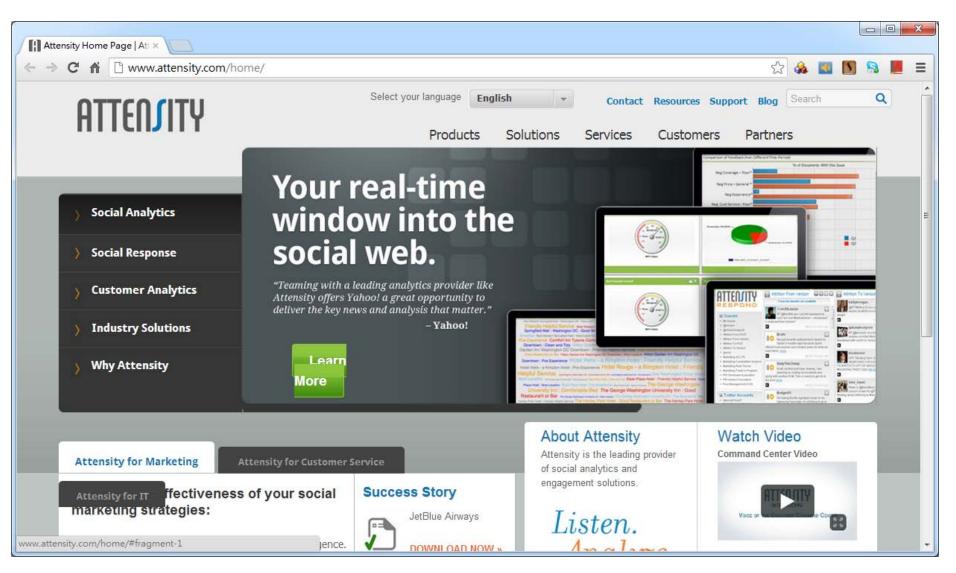
("Social Media Monitoring/Analysis")

- 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

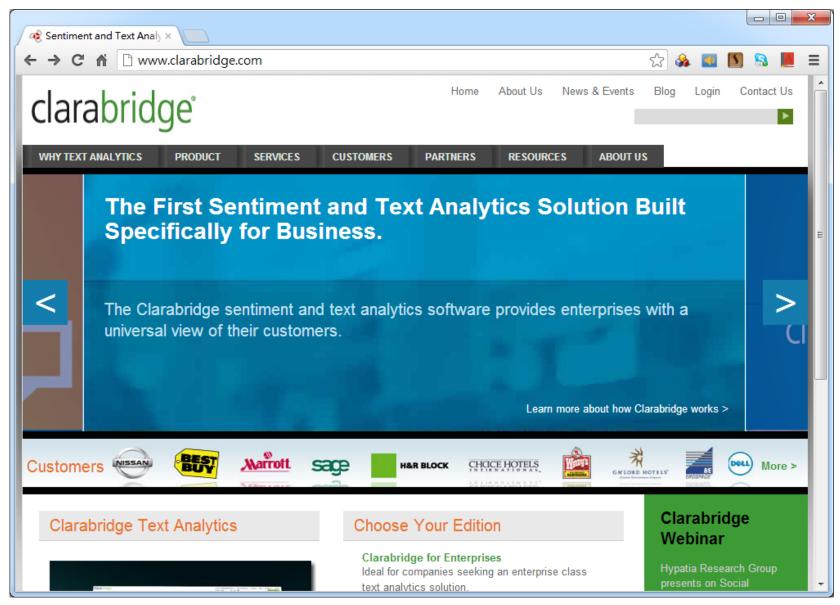
- 1. Attensity
 - Track social sentiment across brands and competitors
 - <u>http://www.attensity.com/home/</u>
- 2. Clarabridge
 - Sentiment and Text Analytics Software
 - <u>http://www.clarabridge.com/</u>

Attensity: Track social sentiment across brands and competitors http://www.attensity.com/



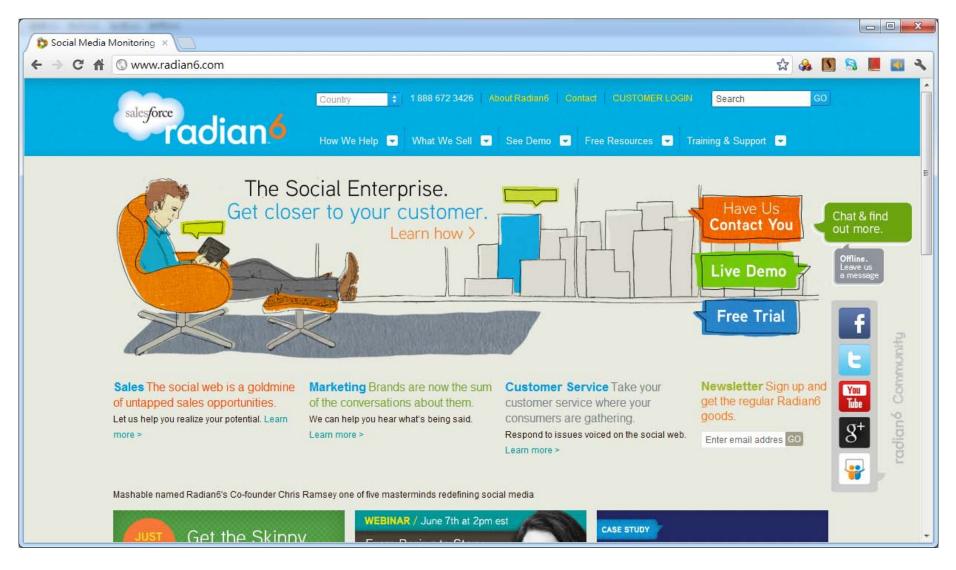
http://www.youtube.com/watch?v=4goxmBEg2Iw#!

Clarabridge: Sentiment and Text Analytics Software http://www.clarabridge.com/



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

http://www.radian6.com/

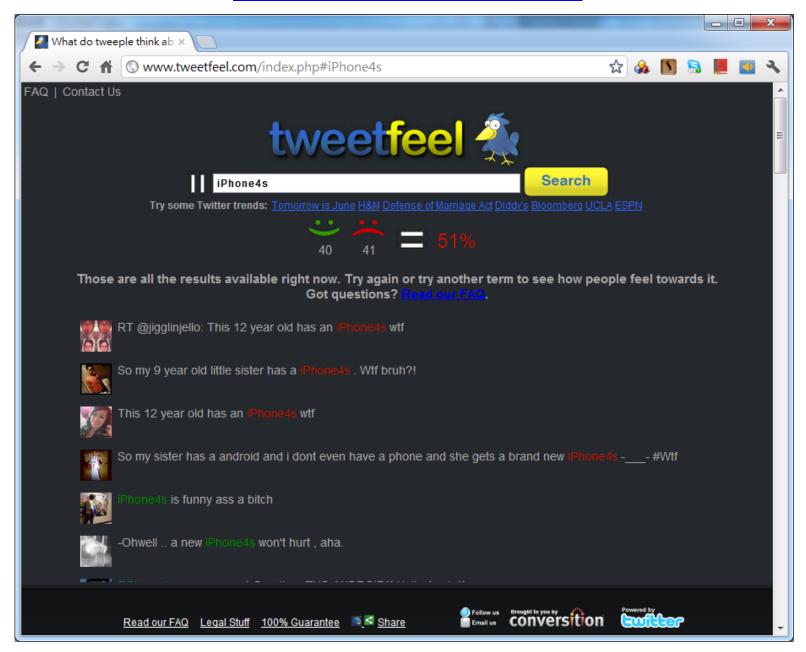


http://www.youtube.com/watch?feature=player_embedded&v=8i6Exg3Urg0

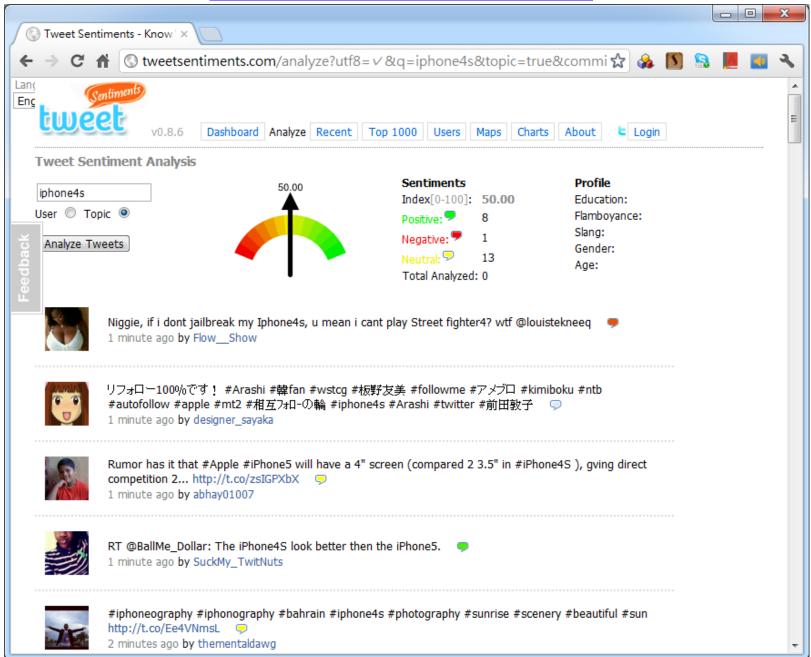
http://www.sas.com/software/customer-intelligence/social-media-analytics/

Social Media Monitoring ×			
	com /software/customer-intelligence/social-media-an	alytics/	☆ 🎄 🛐 🔒 📕 🔩 🔧
	1976 Customer Success Partners Company Support & Trainin NS / SOCIAL MEDIA ANALYTICS	Log In Worldwide Sites v NEWS EVENTS CONSULT	
Products and Solutions Industries Nonprofit Organizations Analytics Business Analytics Business Intelligence Customer Intelligence Strategy & Planning Information & Analytics Customer Experience Customer Experience Customer Experience Analytics Social Media Analytics Web Analytics	SAS® Social Media Analytics Integrate, archive, analyze and act on online conversat Overview Benefits Features Demos & Screenshots 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. Based on your unique business challenges and enterprise goals, SAS can provide a tailored implementation that's hosted and managed by <u>SAS Solutions OnDemand</u> .	System Requirements	Questions? Phone Contact Form Output White Paper Text Analytics for Social Media: Evolving Tools for an Evolving Environment Download Now
 Web Analytics Financial Intelligence Foundation Tools 	Benefits	Product Demo	SAS [®] Social Media Analytics
 Fraud & Financial Crimes Governance, Risk & Compliance 	 Analyze conversation data. Identify advocates of, and threats to, corporate reputation and brand. 		» Overview RESOURCES
 High-Performance Analytics Human Capital Intelligence Information Management 	 Quantify interaction among traditional media/campaigns and social media activity. Establish a platform for social CRM strategy. 		 » Fact Sheet (PDF) » Solution Brief (PDF) » White Papers
a IT 9 CIA Enghlomont			· · · · · · · · · · · · · · · · · · ·

http://www.tweetfeel.com



http://tweetsentiments.com/



http://www.i-buzz.com.tw/

🚫 i-Buzz網路口碑研究中心.網 × 🔲						
← → C ☆ ③ www.i-buzz.com.tw/index.asp iBuzz中國 關於我們			☆ ▲ 100 名 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100 - 100			
ジ博司網路口碑研究中心 Internet Buzz Research Center			議題分析 ▼	Q		
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 纯發文 0%		議題分析 血汗醫院全民皆慌 網路熱門	● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●	更多 三		
文章連結 17% 10% 10%	照片連結 73%	研究専題 便利商店衝人氣,臉書加持少 數據報告 永废房屋穩據房仲龍頭,資訊	馬英九總統 蔡英文 Tsai 「「「」」」」 「」」」 「「」」」 「」」」 新統數 1419591 新統數 669928	蘇貞昌		
排行榜 Traffic 交通	進口車 ・ 吵了這麼久,就是不要 (224)	「超速」超車・還不懂 🖓	研究専題			
 第1名 ・譲一下,很難嗎。。。 ・請問大家認為這樣要賠偿 		-	便利商店街人氣 臉書加持 少不得			
變第2名 PORSCHE (911) 變第3名 SUBARU (legacy)		TIS 条列之比較 (109) 更多	研究専題	社交媒介監看		
更多 Traffic 交通 第1名	高級進口車 • 吵了這麼久,就是不要 (224)	「超速」超車,還不懂 💭	LUMIX與Samsung相機 口碑比拼 粉絲只顧「按 讚」不「留言」	社交保/T监看 了解社群監測		

http://www.eland.com.tw/solutions

C M C WWW.eland.com.tw/solutions C M C M C M C M C M C M C M C M C M C M								
Court Strings 產品 累減服務 查拔標專區 我們的客戶 新聞與活動 文預 服於思述 	()● OpView 服務總覽 意藍科} × ●							
Services 產品 雲湖服務 查銀標專區 批問的容白 新期與活動 支援 翻於意意	← → C f (S www.eland.com.tw/	'solutions				<u>ି</u> ଛ	S 🔝 📕	بة 🔝
<section-header><section-header><complex-block><complex-block><complex-block><complex-block><complex-block></complex-block></complex-block></complex-block></complex-block></complex-block></section-header></section-header>		產品 雲端服務	臺銀標專區	我們的客戶	新聞與活動	支援	關於意藍	
<section-header><section-header><complex-block><section-header><complex-block><complex-block><complex-block></complex-block></complex-block></complex-block></section-header></complex-block></section-header></section-header>								=
Coperational (************************************			c	Dieme.		9 mg 为 ta	- (B)	
您的奧論觀測站 連上OpView · 品牌形象 · 奧論觀點監測好輕鬆 </th <th>Opview Serv</th> <th>/ice</th> <th>-</th> <th></th> <th></th> <th>-</th> <th>Calves BriAss</th> <th></th>	Opview Serv	/ice	-			-	Calves BriAss	
連上OpView · 品牌形象、輿論觀點監測好輕鬆 <li< th=""><th>您的輿論觀測站</th><th></th><th>l.</th><th></th><th></th><th>Street and a state of state of</th><th>* 171.0018</th><th></th></li<>	您的輿論觀測站		l.			Street and a state of	* 171.0018	
Image: Contract of the state of the st				9· 0· 0· 0		terr Balancer a	OFICA.	
Weight Weigh	連上OpView · 品牌形象、	輿論觀點監測好輕鬆	이는 눈 눈 듯	I MOTALS DOR - C CALL - CALL CALL - CALL		1.00	0 KER - 0 Die -	
● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●								
OpView服務總覽 搜尋 OpView 服務總覽 相關文件下載 Image: CopView Insight DM) 🛃 🍊				- 1.00	124	
OpView 服務總覽 聞 OpView Insight DM				6 6	22 :0-		And and the	
OpView 服務總覽 聞 OpView Insight DM								
相關文件下載 図 OpView Insight DM	OpView 服務總覽						搜尋	₽ P
相關文件下載 図 OpView Insight DM								
	Opview 服務總見				用	相關文件下載		
					Ŧ			
什麽是OpView服務?	什麽是OpView服務?				▶ OpView品牌□碑雷達 DM			
	OpView是協助您蒐集、處理、分析網路資訊的雲端服務。					☑ OpView品牌□碑雷達訂購 單		
		D I W JPH 5-2 HIGH J 2-5 MILAR (V)				卑		
更多更多。 OpView服務,猶如提供您每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	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				各類型網		更多	
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、	-							

http://opview-eland.blogspot.tw/2012/05/blog-post.html

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 policitians' 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

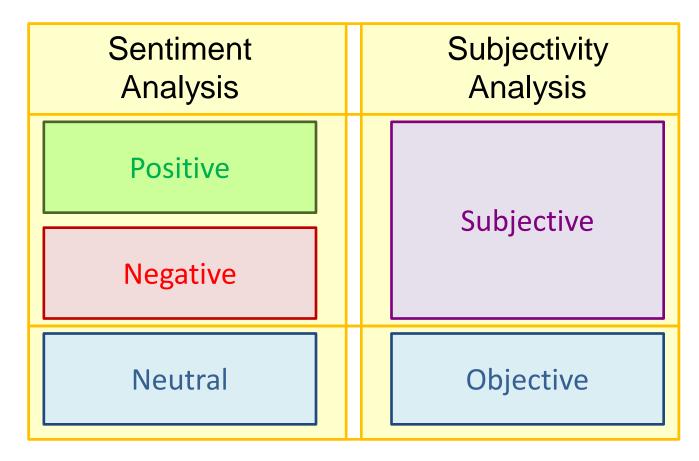
• Objective

- 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

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

35

Sentiment Analysis vs. Subjectivity Analysis



A (regular) opinion

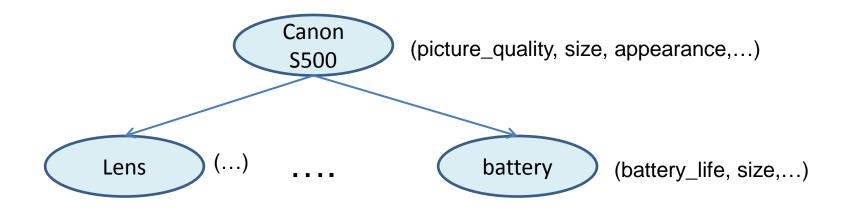
- Opinion (a restricted definition)
 - An opinion (regular opinion) is simply a positive or negative sentiment, view, attitude, emotion, or appraisal about an entity or an aspect of the entity from an opinion holder.
- Sentiment orientation of an opinion
 - Positive, negative, or neutral (no opinion)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

38

Entity and aspect



Opinion definition

- - $-e_j$ 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.
 - $-t_1$ is the time when the opinion is expressed.

Opinion definition

- - $-e_j$ 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.
 - $-t_1$ is the time when the opinion is expressed.
- (*e_j*, *a_{jk}*) is also called opinion target

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

41

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 7 Negative: Aspect: Voice quality Positive: 120 Negative: 8 Aspect: Battery Positive: 80 Negative: 12

. . .

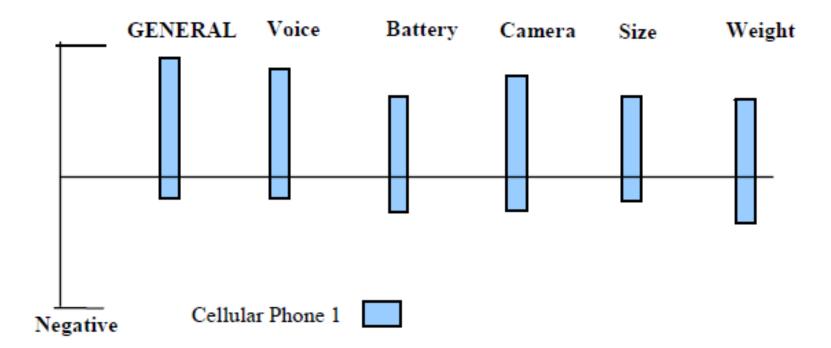
<individual review sentences> <individual review sentences>

<individual review sentences> <individual review sentences>

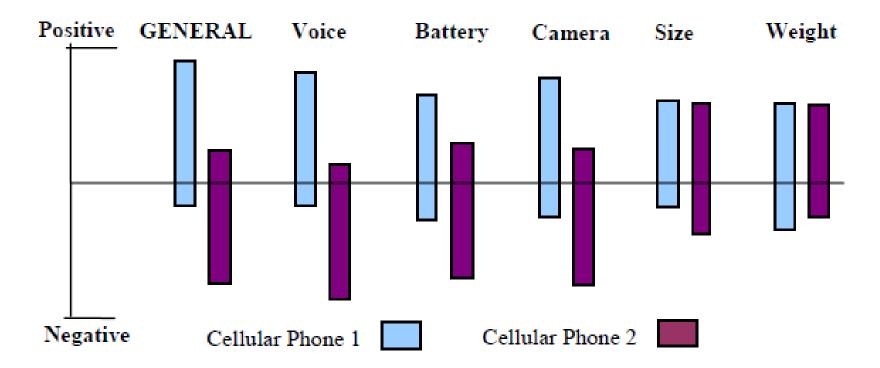
<individual review sentences> <individual review sentences>

46

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

50

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

Source: Bing Liu (2011), "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data," Springer, 2nd Edition, 51

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	Sentiment Identification		Sentiment Aggregation		Nature of
	Task	Method	Level	Method	Level	Measure
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)

Source: Zhang, Z., Li, X., and Chen, Y. (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.,

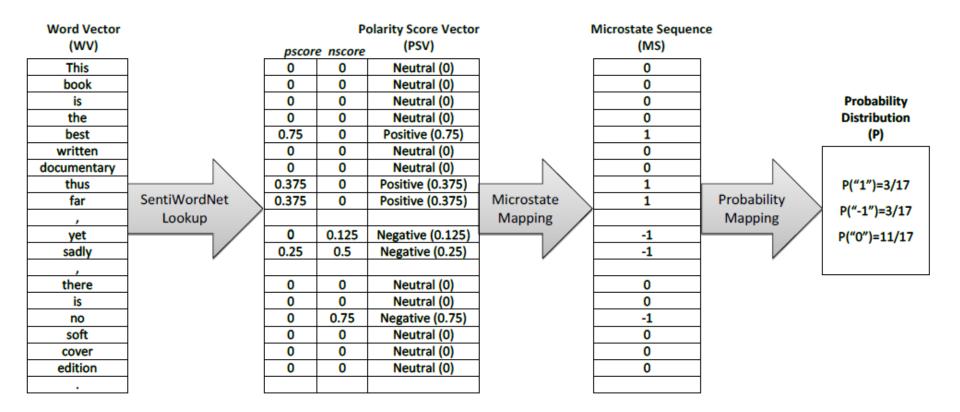
Word-of-Mouth (WOM)

 "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	VBN
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



Source: Zhang, Z., Li, X., and Chen, Y. (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.,

Datasets of Opinion Mining

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

Lexical Resources of Opinion Mining

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

Example of SentiWordNet

- POSIDPosScoreNegScoreSynsetTermsGlossa002177280.750beautiful#1delighting the senses orexciting 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(betak)

- "中英文情感分析用詞語集"
 一包含詞語約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

- •"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...

- •"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...

- "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

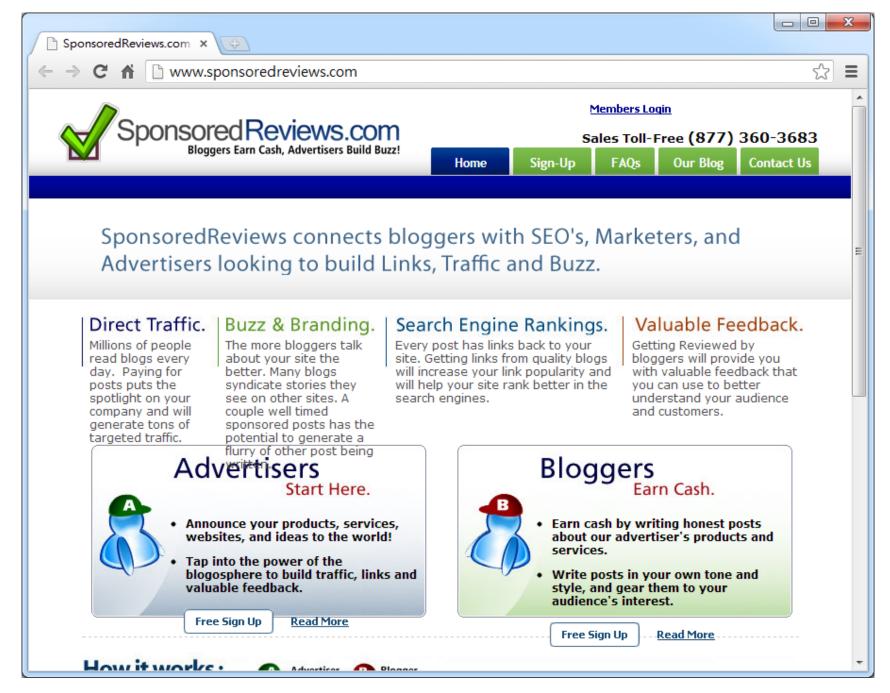
- fake reviews (also called bogus reviews)
- fake comments
- fake blogs
- fake social network postings
- deceptions
- deceptive messages

Fake Review Detection

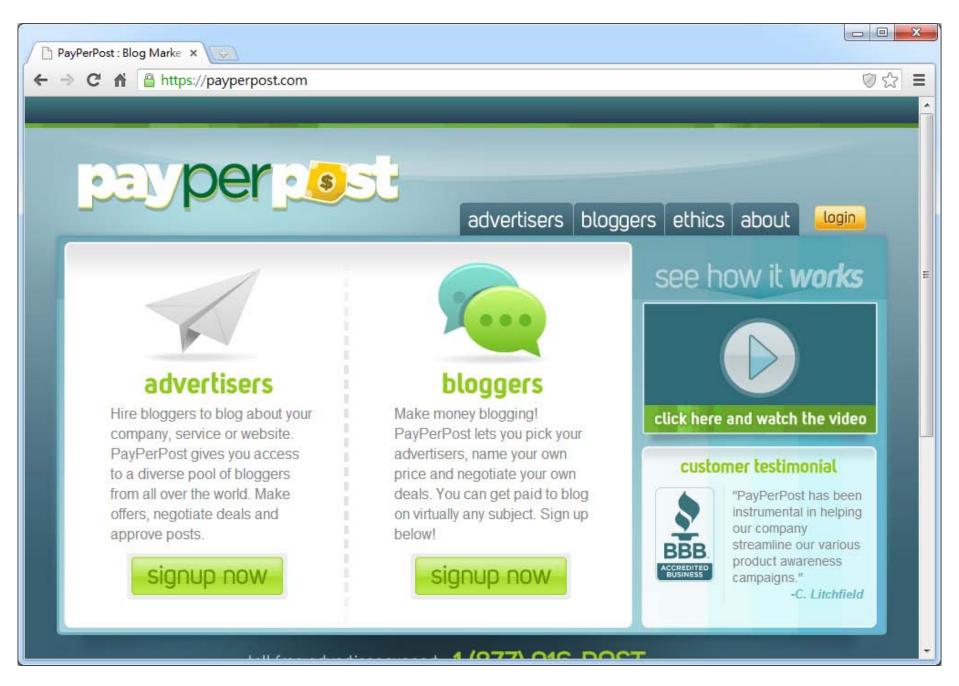
- Methods
 - supervised learning
 - pattern discovery
 - graph-based methods
 - relational modeling
- Signals
 - Review content
 - Reviewer abnormal behaviors
 - Product related features
 - Relationships

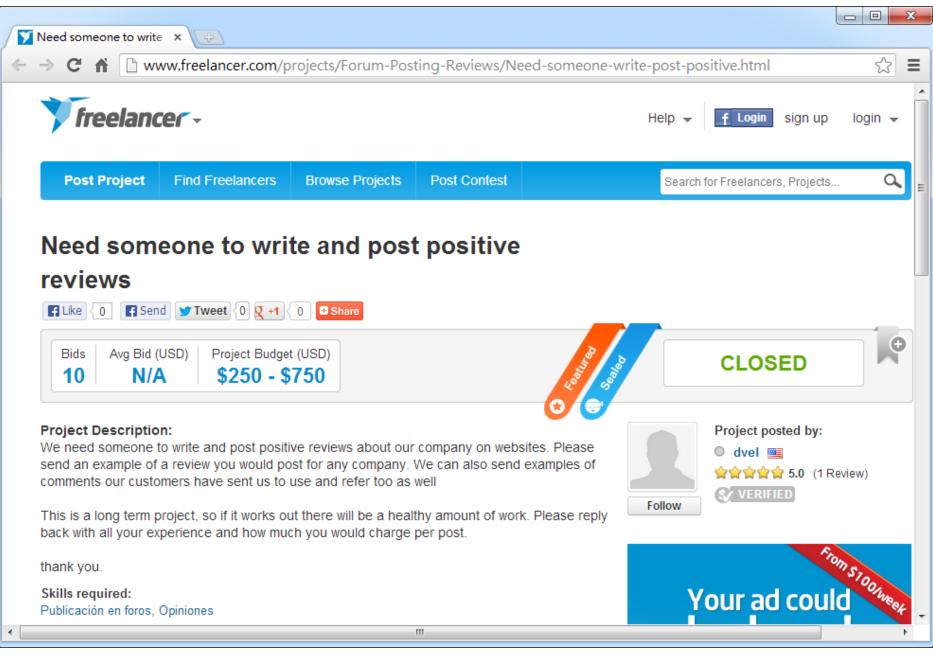
Professional Fake Review Writing Services (some Reputation Management companies)

- 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)



Source: http://www.sponsoredreviews.com/





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.
- 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, <u>http://www.cs.uic.edu/~liub/FBS/fake-reviews.html</u>
- 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.