Fakultas Informatika
School of Computing
Telkom University

CTI-3A3
Applied Social Network Analysis

Dr. Warih Maharani

wmaharani.staff.telkomuniversity.ac.id

Lecture 13 – SNA Applications

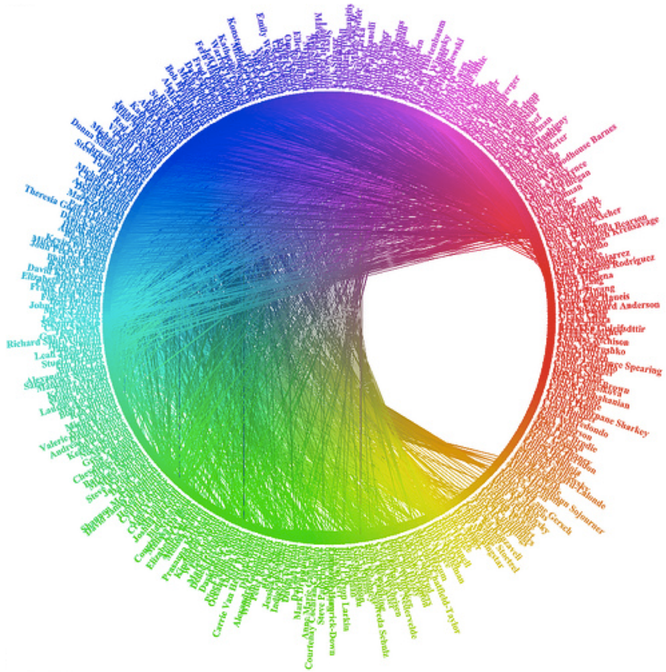
Outline

- Sentiment Analysis
  - Aspect Extraction
  - Sentiment Classification
- SNA Applications

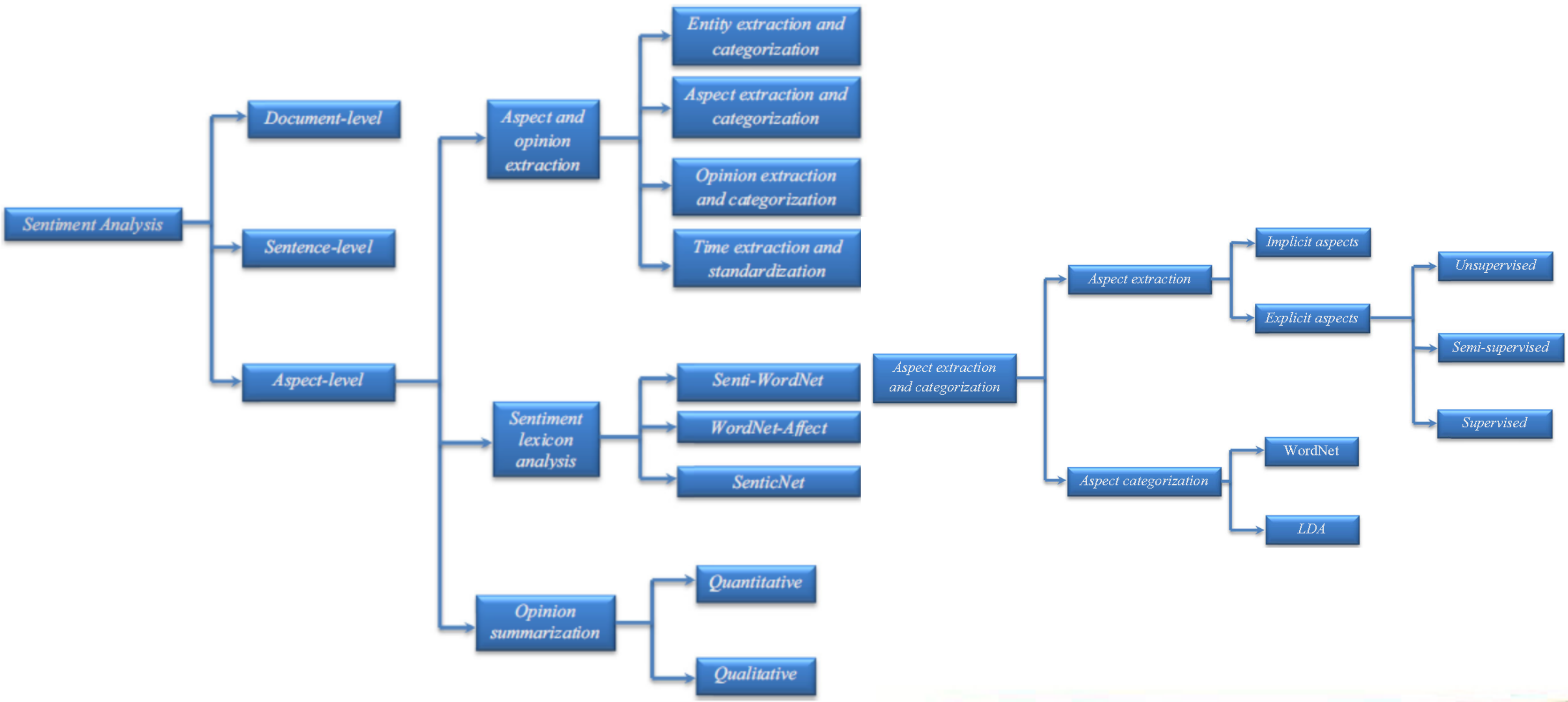
Course Learning Outcomes (CLO)

- apply the network method in Applied Social Network Analysis, as well as analyze and visualize using Social Network Analysis tools
- apply the concept of Social Network Analysis

Sentiment Analysis
Aspect Extraction



Aspect extraction in sentiment analysis: comparative analysis and survey. Toqir A. Rana, Yu Cheah. 2016.



ASPECT EXTRACTION

- 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.

Fine-grained sentiment

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, …*"

- Two products, iPhone and Blackberry
- Overall positive to iPhone, negative to Blackberry
- Postive aspect/features of iPhone: touch screen, voice quality. Negative (for the mother): expensive.

Sub Problems

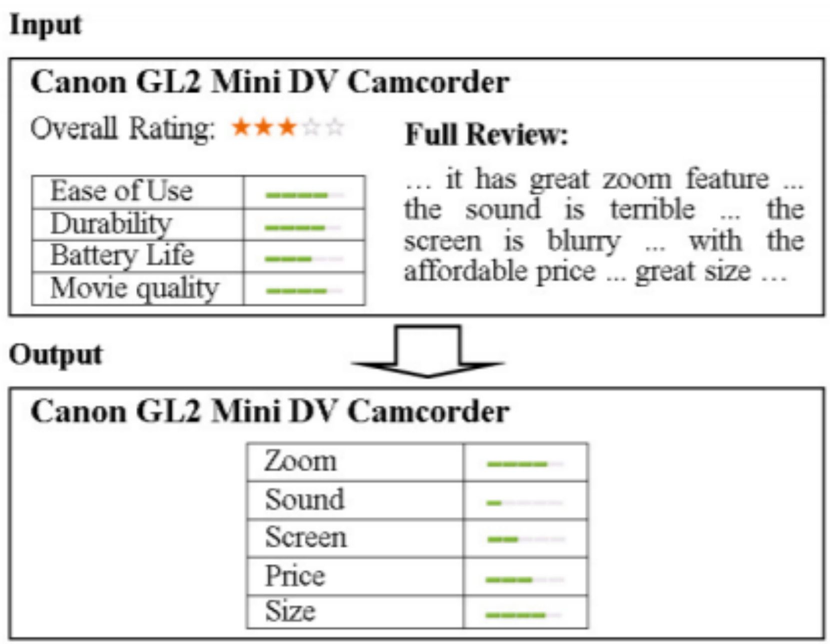
- *e* is a target entity: Named Entity Extraction
- *a* is an aspect of e: Information Extraction
- *so* is sentiment: Sentiment Identification & Classification
- *h* is an opinion holder: Information Extraction
- *t* is the time: Information Extraction

Aspect-based Opinion Mining

- Tasks:
  - Aspect Extraction
  - Rating/Polarity prediction
  - Aspect grouping
  - Coreference resolution
  - Entity, opinion holder, time extraction

ASPECT EXTRACTION

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

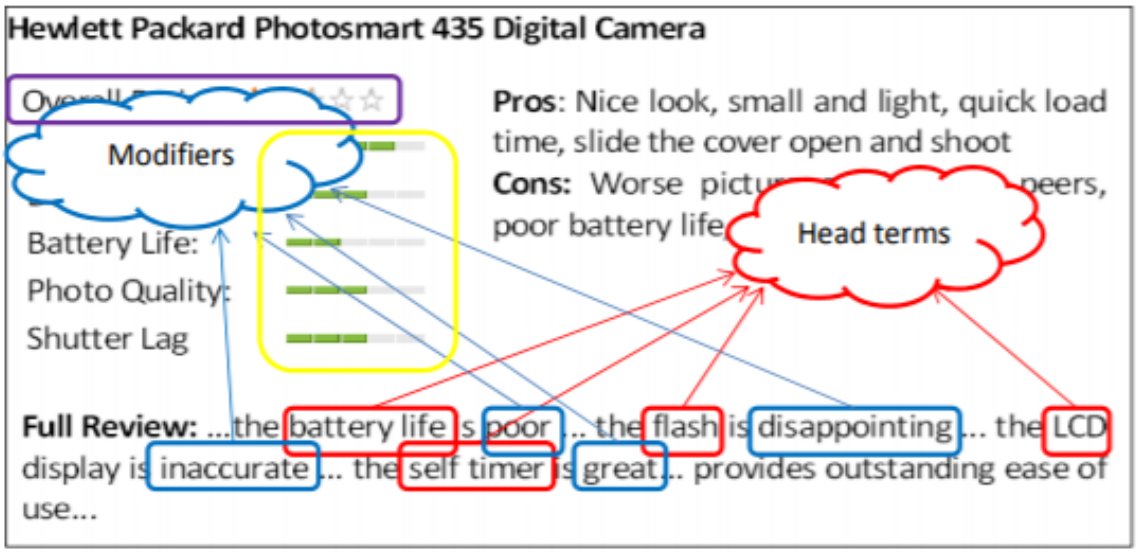


ASPECT EXTRACTION

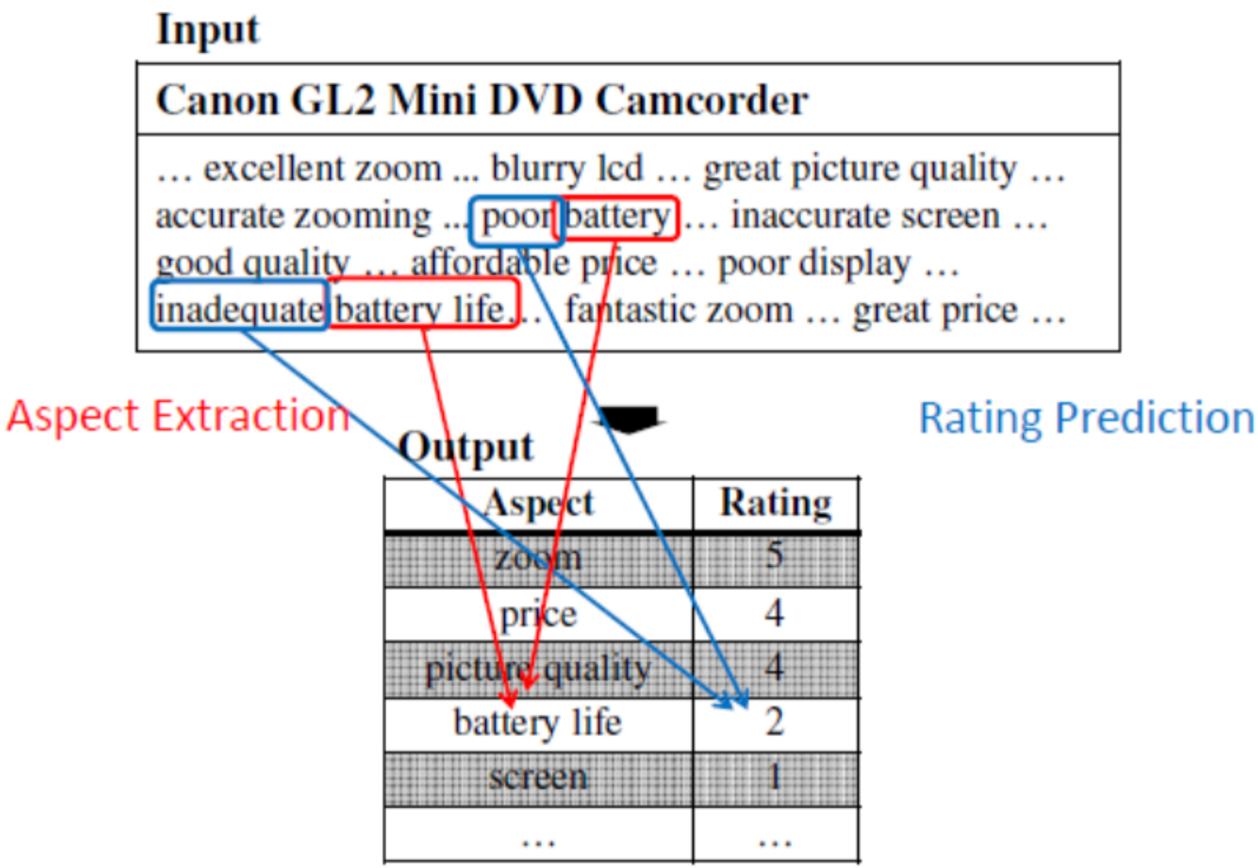
Typically, three step-approach:

1. Extract opinion phrases: <head term, modifier>

   e.g., <LCD, blurry>,<screen, inaccurate>, <display, poor>
2. Cluster head terms referring to same aspect and modifiers referring to same rating
3. Choose “names” for aspects and ratings



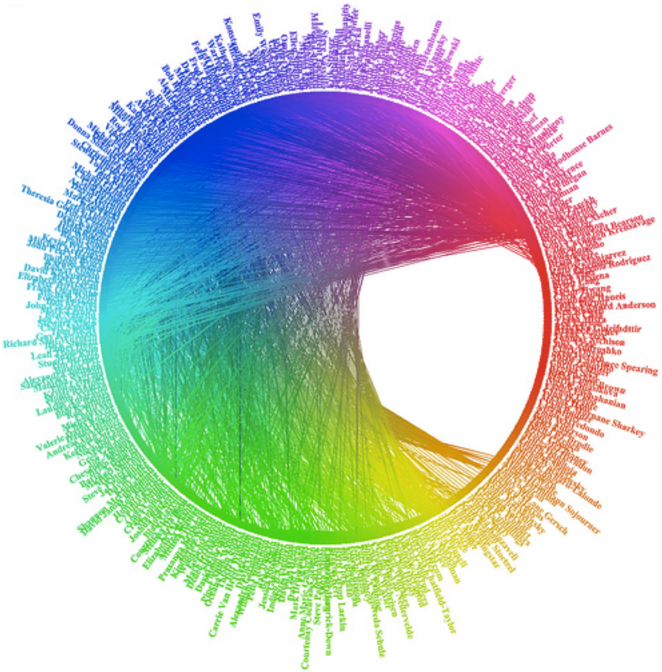
ASPECT EXTRACTION



Challenges

- Different head terms for same aspect
  - Photo quality is a little better than most of the cameras in this class
  - That gives the better image quality, especially in low light
- Different modifiers for same rating
  - For a camera of this price, the picture quality is amazing
  - I am going on a trip to France and wanted something that could take stunning pictures with, but didn't cost a small fortune
- Noise
  - I am an amateur photographer
  - I have fat hands but short fingers
- Implicit aspects or ratings
  - After a twenty-one mile bike ride a four mile backpacking river hike, the size, weight, and performance of this camera has been the answer to my needs
  - The grip and weight make it easy to handle and the mid zoom pictures have exceeded expectation

Sentiment Analysis
Sentiment Classification



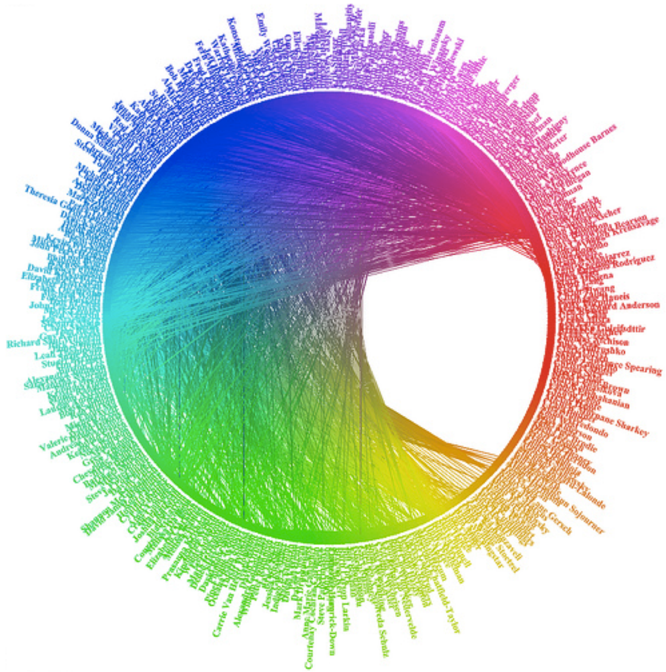
Goals

- Identify and classify opinions
- Task 1: Sentiment Identification
  - Identify whether a piece of text expresses opinions
- Task 2: Sentiment Classification
  - Determine the orientation of an opinionated text

Sentiment Analysis Levels

- Document level
  - Identify if document expresses opinions and whether the opinions are positive, negative, or neutral
- Sentence level
  - Identify if sentence expresses opinions and whether the opinions are positive, negative, or neutral
- Attribute level
  - Extract the object attributes that are the subject of an opinion and the opinion orientations

Sentiment Analysis
Sentiment Lexicon



The General Inquirer

Philip J. Stone, Dexter C Dunphy, Marshall S. Smith, Daniel M. Ogilvie. 1966. The General Inquirer: A Computer Approach to Content Analysis. MIT Press

- Home page: http://www.wjh.harvard.edu/~inquirer
- List of Categories: http://www.wjh.harvard.edu/~inquirer/homecat.htm
- Spreadsheet: http://www.wjh.harvard.edu/~inquirer/inquirerbasic.xls

- Categories:
  - Positive (1915 words) and Negativ (2291 words)
  - Strong vs Weak, Active vs Passive, Overstated versus Understated
  - Pleasure, Pain, Virtue, Vice, Motivation, Cognitive Orientation, etc
- Free for Research Use

LIWC (Linguistic Inquiry and Word Count)

Pennebaker, J.W., Booth, R.J., & Francis, M.E. (2007). Linguistic Inquiry and Word Count: LIWC 2007. Austin, TX

- Home page: http://www.liwc.net/
- 2300 words, >70 classes
- **Affective Processes**
  - negative emotion (*bad, weird, hate, problem, tough*)
  - positive emotion (*love, nice, sweet*)
- **Cognitive Processes**
  - Tentative (*maybe, perhaps, guess*), Inhibition (*block, constraint*)
- **Pronouns, Negation** (*no, never*), **Quantifiers** (*few, many*)
- $30 or $90 fee

MPQA Subjectivity Cues Lexicon

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proc. of HLT-EMNLP-2005.

Riloff and Wiebe (2003). Learning extraction patterns for subjective expressions. EMNLP-2003.

- Home page: http://www.cs.pitt.edu/mpqa/subj_lexicon.html
- 6885 words from 8221 lemmas
  - 2718 positive
  - 4912 negative
- Each word annotated for intensity (strong, weak)
- GNU GPL

Bing Liu Opinion Lexicon

Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. ACM SIGKDD-2004.

- Bing Liu's Page on Opinion Mining
- http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

- 6786 words
  - 2006 positive
  - 4783 negative

SentiWordNet

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010 SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC-2010

- Home page: http://sentiwordnet.isti.cnr.it/
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
- [estimable(J,3)] “may be computed or estimated”

```
Pos   0    Neg 0    Obj 1
```

- [estimable(J,1)] “deserving of respect or high regard”

```
Pos .75   Neg 0    Obj .25
```

Disagreements between polarity lexicons

Sentiment Tutorial

	Opinion Lexicon	General Inquirer	SentiWordNet	LIWC
MPQA	33/5402 (0.6%)	49/2867 (2%)	1127/4214 (27%)	12/363 (3%)
Opinion Lexicon		32/2411 (1%)	1004/3994 (25%)	9/403 (2%)
General Inquirer			520/2306 (23%)	1/204 (0.5%)
SentiWordNet				174/694 (25%)
LIWC				

WordNet

- WordNet is an online lexical database in which English nouns, verbs, adjectives and adverbs are organized into sets of synonyms.
  - Each word represents a lexicalized concept. Semantic relations link the synonym sets (synsets).

- WordNet contains more than 118,000 different word forms and more than 90,000 senses.
  - Approximately 17% of the words in WordNet are polysemous (have more than on sense); 40% have one or more synonyms (share at lease one sense in common with other words).

WordNet

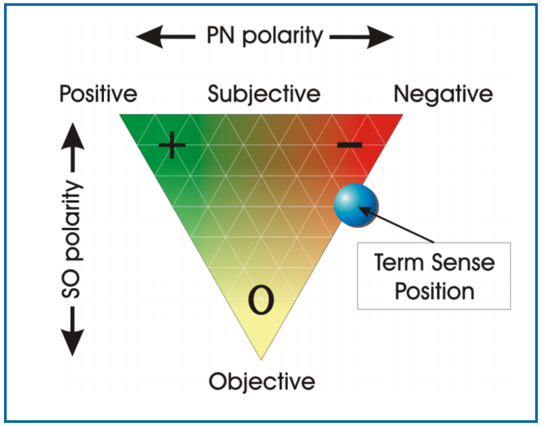
- Six semantic relations are presented in WordNet because they apply broadly throughout English and because a user need not have advanced training in linguistics to understand them. The table below shows the included semantic relations.

Semantic Relation	Syntactic Category	Examples
Synonymy (similar)	Noun, Verb, Adjective, Adverb	Pipe, tube Rise, ascent Sad, happy Rapidly, speedily
Antonymy (opposite)	Adjective, Adverb	Wet, dry Powerful, powerless Rapidly, slowly
Hyponymy (subordinate)	Noun	Maple, tree Tree, plant
Meronymy (part)	Noun	Brim, hat Ship, fleet
Troponomy (manner)	Verb	March, walk Whisper, speak
Entailment	Verb	Drive, ride Divorce, marry

- WordNet has been used for a number of different purposes in information systems, including word sense disambiguation, information retrieval, text classification, text summarization, machine translation and semantic textual similarity analysis

SentiWordNet

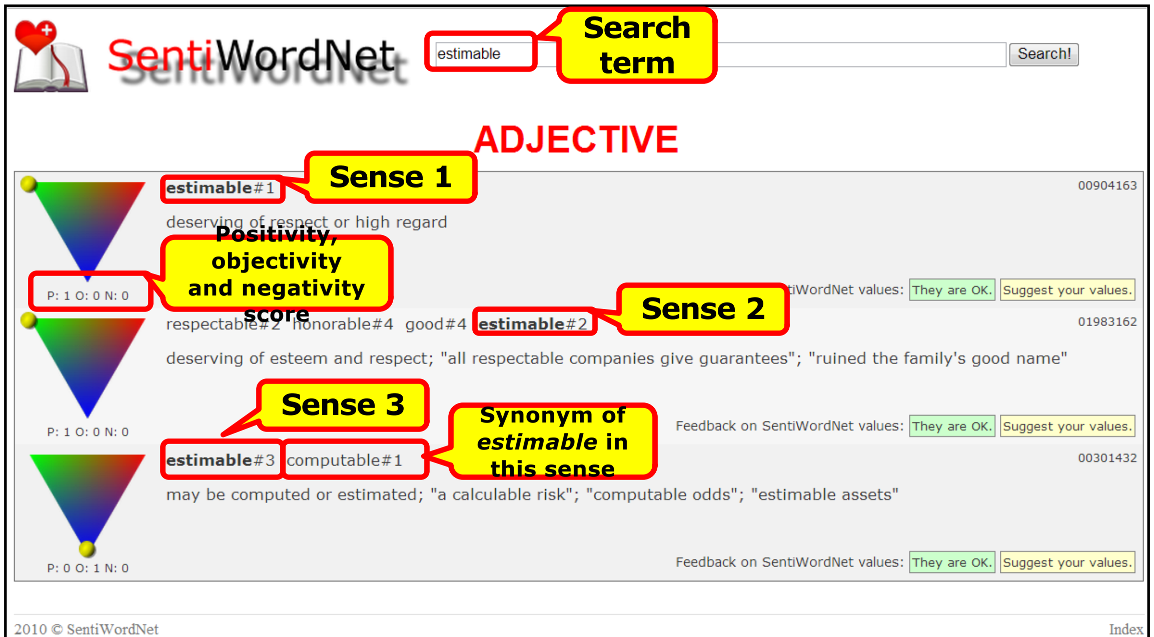
- SentiWordNet is a lexical resource explicitly devised for supporting sentiment analysis and opinion mining applications.
- SentiWordNet is the result of the automatic annotation of all the synsets of WordNet according to the notions of “positivity”, “negativity” and “objectivity”.
- Each of the “positivity”, “negativity” and “objectivity” scores ranges in the interval [0.0,1.0], and their sum is 1.0 for each synset.



The figure above shows the graphical representation adopted by SentiWordNet for representing the opinion-related properties of a term sense.

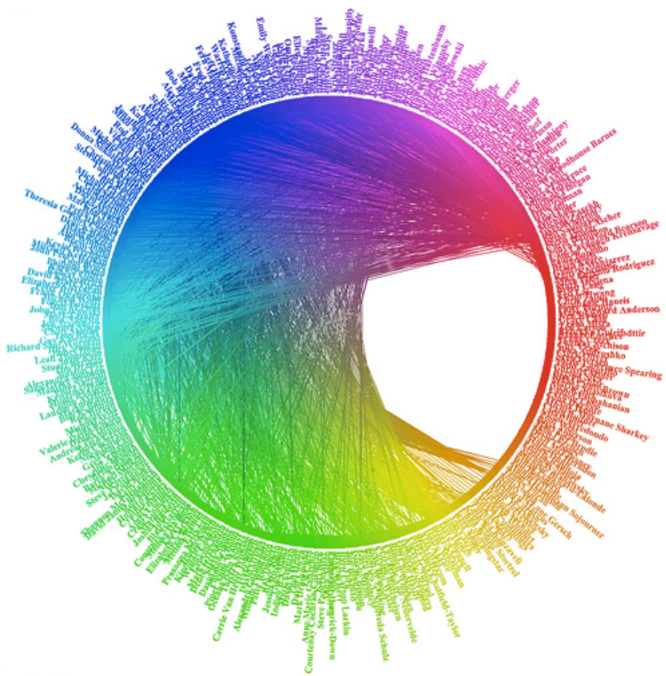
SentiWordNet

- In SentiWordNet, different senses of the same term may have different opinion-related properties.



The figure above shows the visualization of opinion related properties of the term *estimable* in SentiWordNet (http://sentiwordnet.isti.cnr.it/search.php?q=estimable).

SNA Applications



Predicting Popularity in social media

- **Popularity** of brand posts on brand fan pages: An investigation of the effects of **social media** marketing
- Does content determine information **popularity** in **social media**?: A case study of youtube videos' content and their **popularity**
- What messages to post? Evaluating the **popularity** of **social media** communications in business versus consumer markets
- Unfolding temporal dynamics: Predicting **social media popularity** using multi-scale temporal decomposition

Sociology and Human Interaction

- With the huge number of people who are involved nowadays with social networks, it is very interesting to note how they are influenced by each other in many different ways.
  - e.g., identity in the age of social media

CNN Health » Why some 13-year-olds check social media 100 times a day International Edition +

Why some 13-year-olds check social media 100 times a day

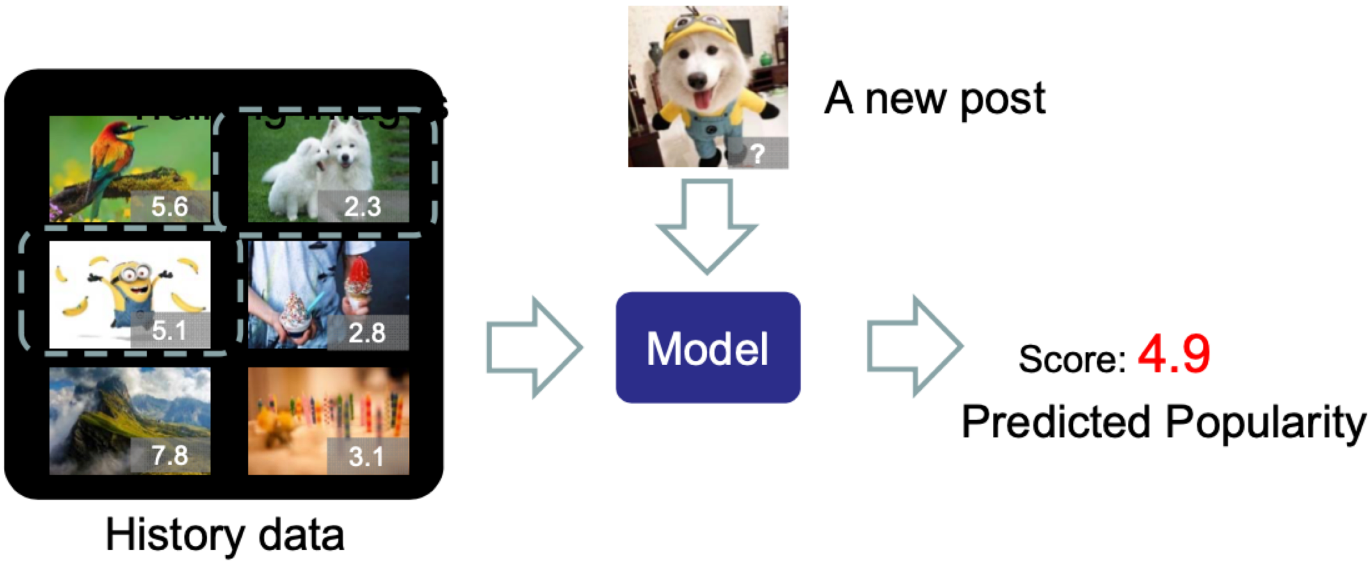
By Chuck Hadad, CNN
October 13, 2015

Why are teens so anxious about what's happening online? #Being13 found that it's largely due to a need to monitor their own popularity status, and defend themselves against those who challenge it.

- 61% of teens said they wanted to see if their online posts are getting likes and comments.
- 36% of teens said they wanted to see if their friends are doing things without them.
- 21% of teens said they wanted to make sure no one was saying mean things about them.

Social Popularity Prediction

- General Popularity Prediction: Predicting the popularity score of a new social media post by combining post content (photo, text or video) and user cues



Why is it important?

- wide applications and high business value
  - e.g., predicting the “Stars of Tomorrow” (top popular models) within the fashion Industry using social media

Fashion Model Directory (FMD) profile page

Can you tell who will be the “top”?

[Ref] “Style in the Age of Instagram: Predicting Success within the Fashion Industry using Social Media,” CSCW 2016

People are desired for knowing the future…

Inside the Washington Post's popularity prediction experiment

A peek into the clickstream analysis and production pipeline for processing tens of millions of daily clicks, for thousands of articles.

By Shuguang Wang and Eui-Hong (Sam) Han, January 25, 2017

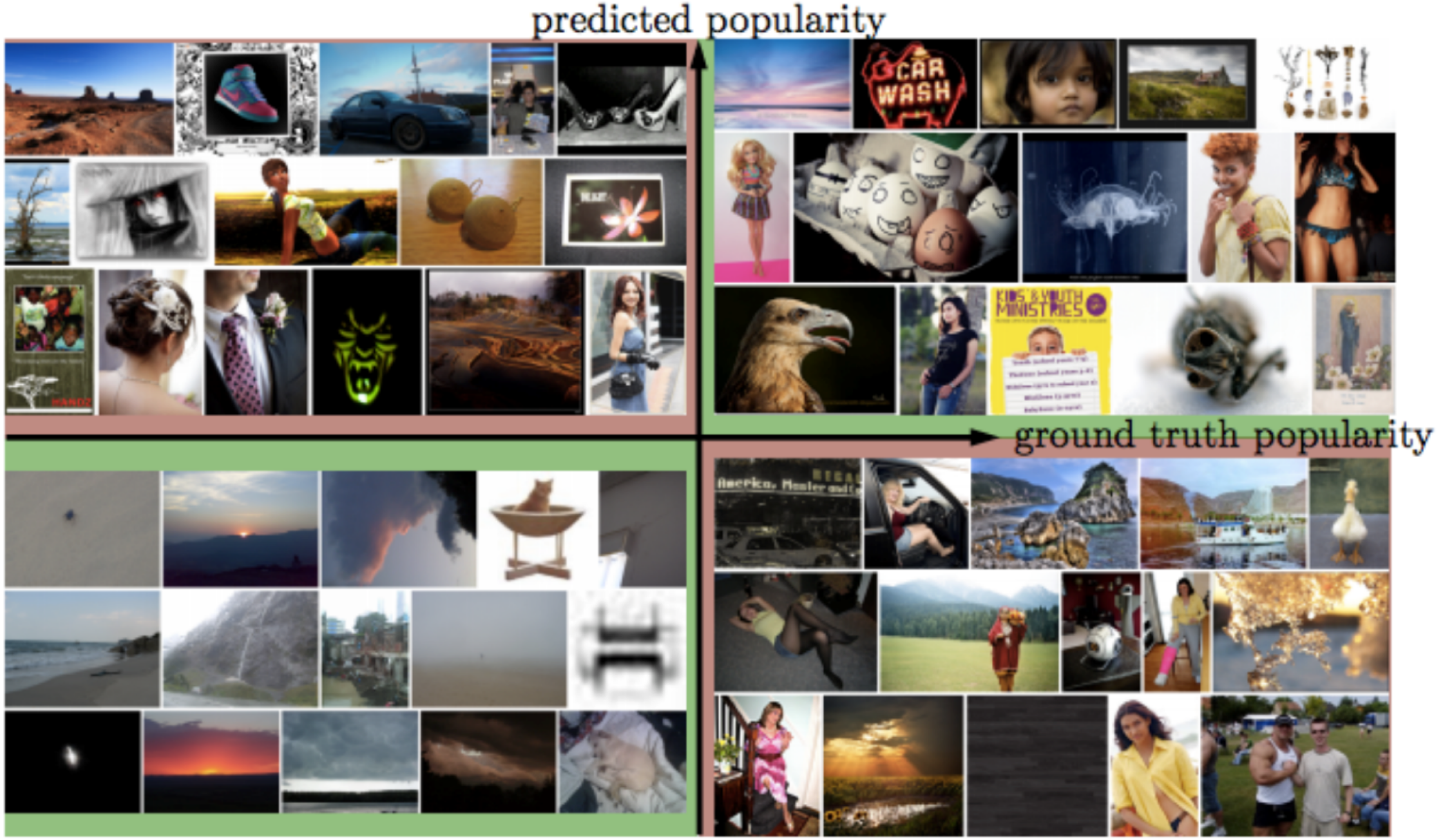
Learn more about working with stream processing in the tutorial Building real-time data pipelines with Apache Kafka at Strata + Hadoop World San Jose, March 13-16, 2017.

In the distributed age, news organizations are likely to see their stories shared more widely, potentially reaching thousands of readers in a short

amount of time. At the *Washington Post*, we asked ourselves if it was possible to predict which stories will become popular. For the *Post*

more, in order to more deeply engage the new and occasional readers clicking through to a popular story.

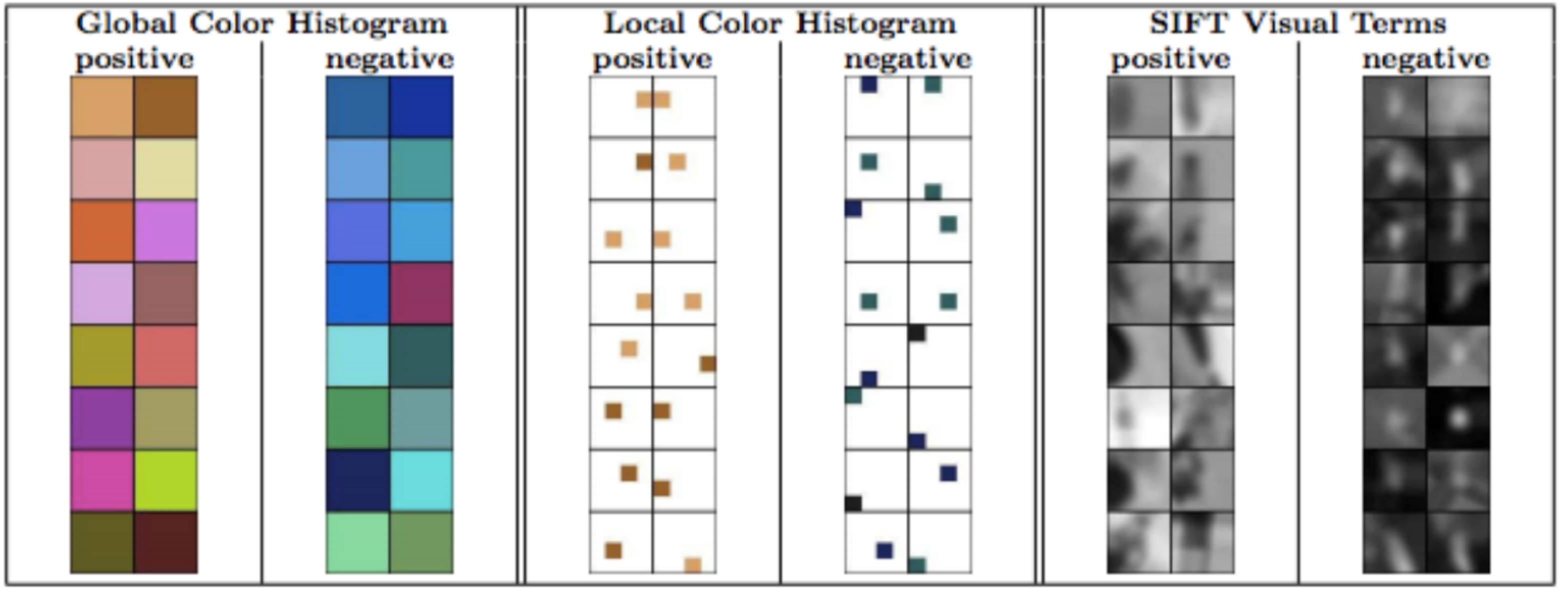
What Makes A Post Popular?



Ref] "What Makes an Image Popular?" WWW, 2014.

What Makes A Post Popular?

- Features for prediction
  - Post content
    - e.g., visual sentiment features (color and texture)



[Ref] "Analyzing and predicting sentiment of images on the social web," ACM Multimedia 2010.

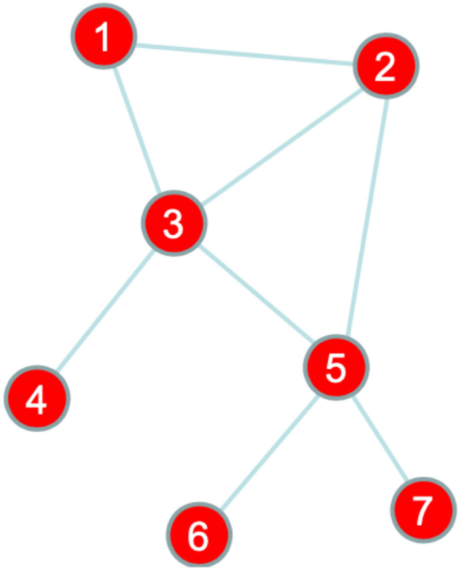
What Makes A Post Popular?

- Features for prediction
  - User cues
    - e.g., followers (a user's follower count), friends (how many users a user follows), statuses (a user's current total post count), user time (a user's account creation time), etc.

A friend graph:

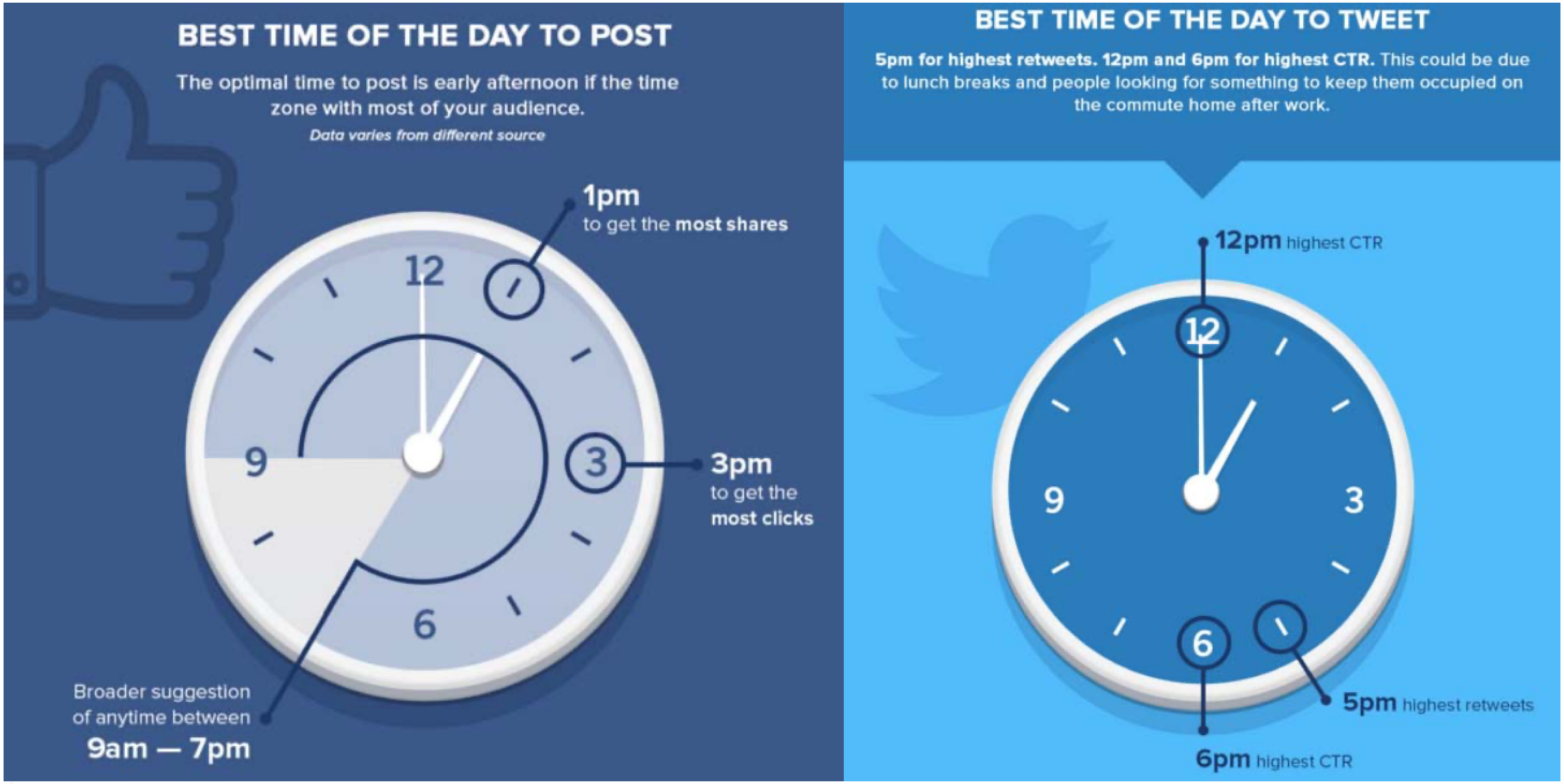
What Makes A Post Popular?

- Features for prediction
  - User cues (topological features)
    - e.g., closeness centrality, the average length of the shortest path between the node and all other nodes in the graph



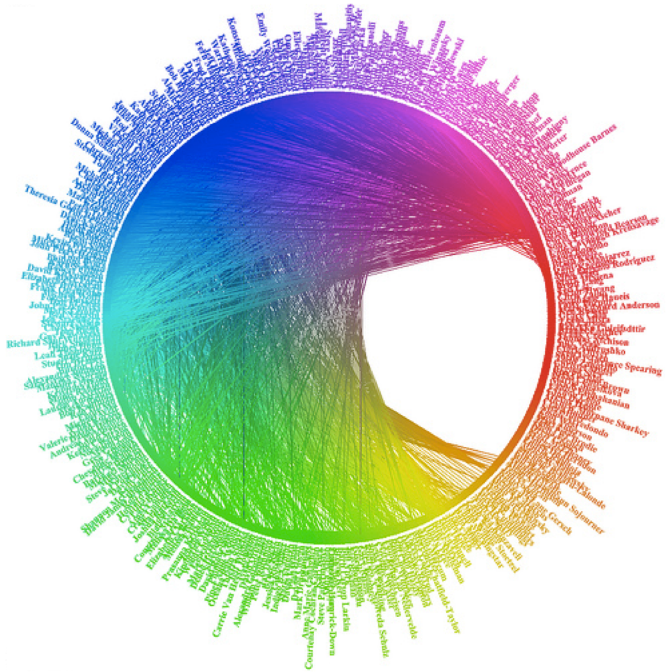
$$C(x) = \frac{1}{\sum_y d(y, x)}$$

where $d(y, x)$ is the distance between vertices x and y



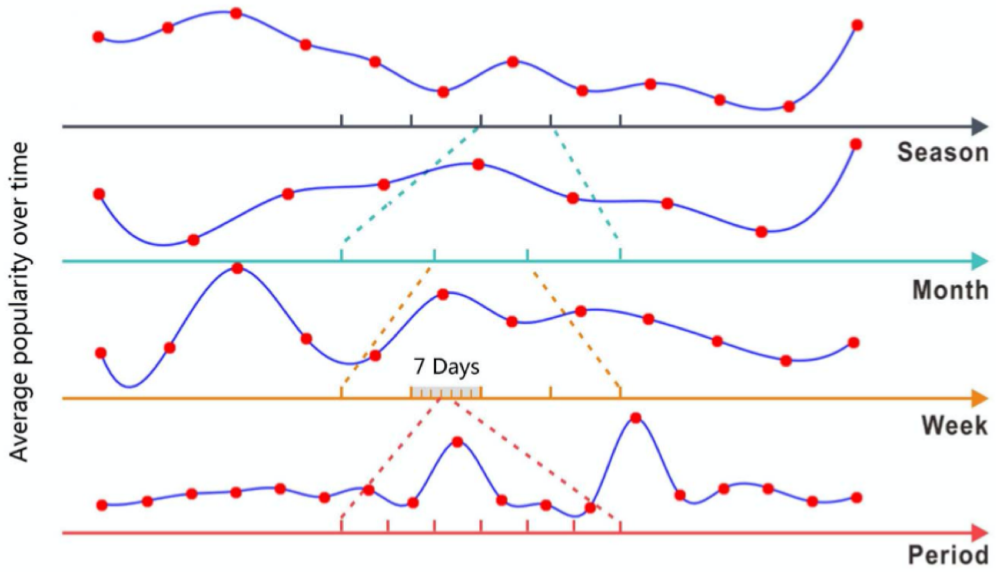
[Ref] http://www.adweek.com/socialtimes/best-time-to-post-social-media/504222

Challenges



Challenge 1: Temporal Evolving

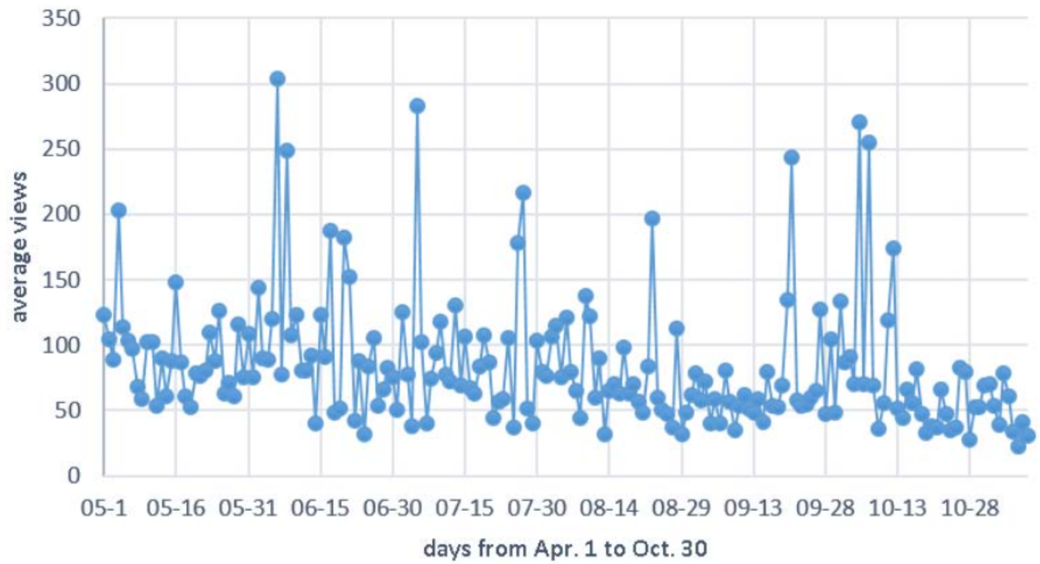
- The popularity evolving at multi-granularities with different patterns



Multi-granularities Characteristics of Popularity Dynamics

Challenge 2: Data Noise

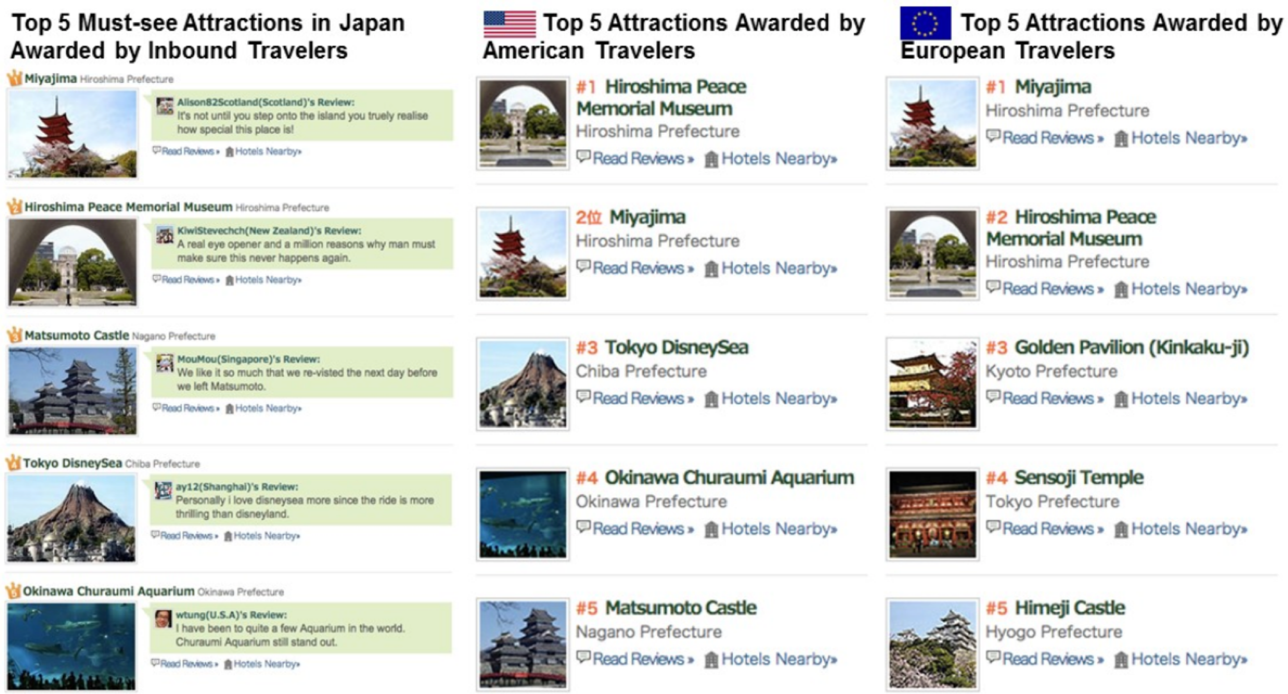
- Popularity patterns are covered in very noisy behavior data or information



Popularity distribution on time series

More Influential Factors: Cultures

- A voting survey of the 2014 TripAdvisor's Top 10 Attractions in Japan by visitors from different countries shows how much the favorites for attractions can vary among people from different regions, i.e., different cultures.

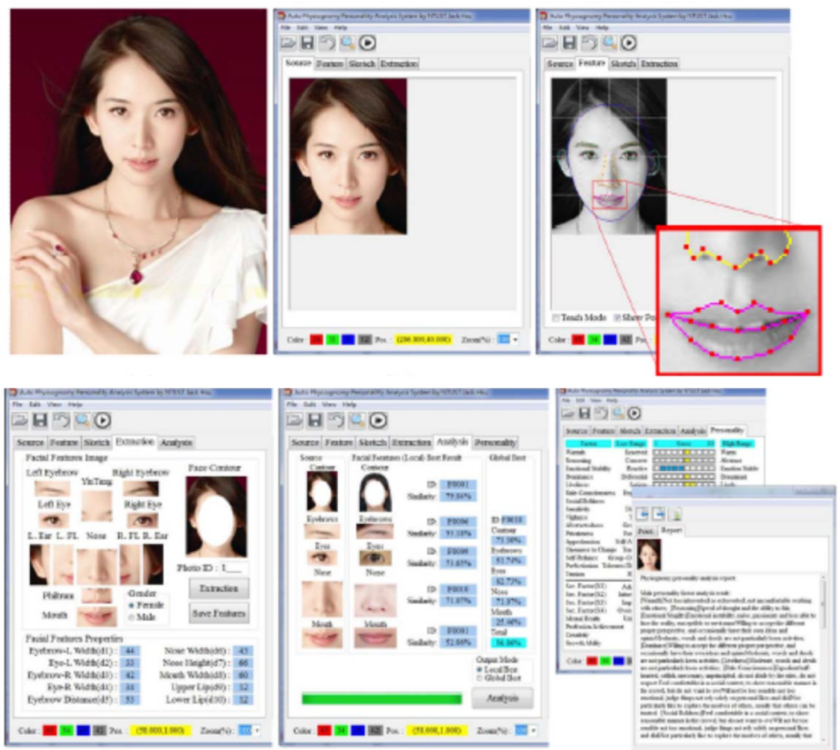


[Ref] 2014 TripAdvisor's Top 20 Attractions in Japan: http://www.tripadvisor.com/pages/- HotSpotsJapan.html.

More Influential Factors: Personalization

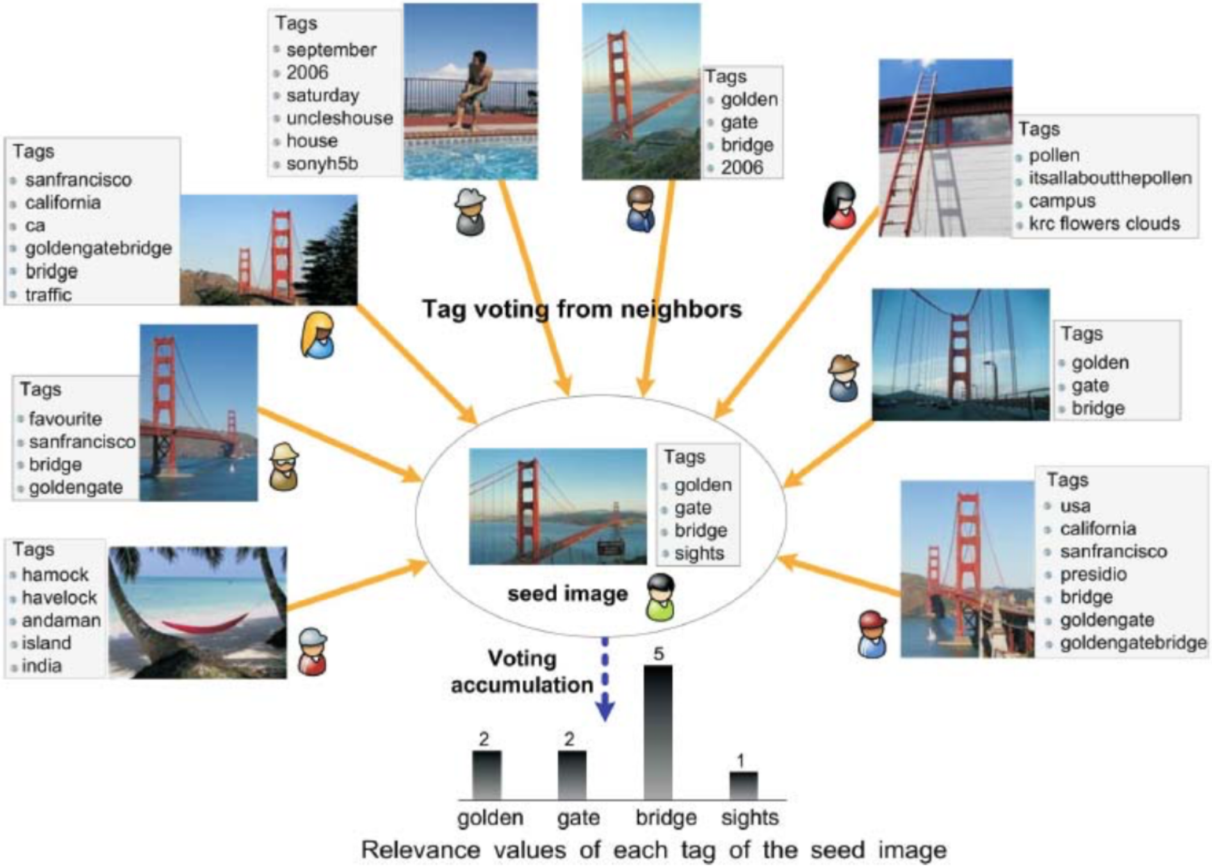
- What Your Facial Features Say About Your Personality (MM13)

facial image



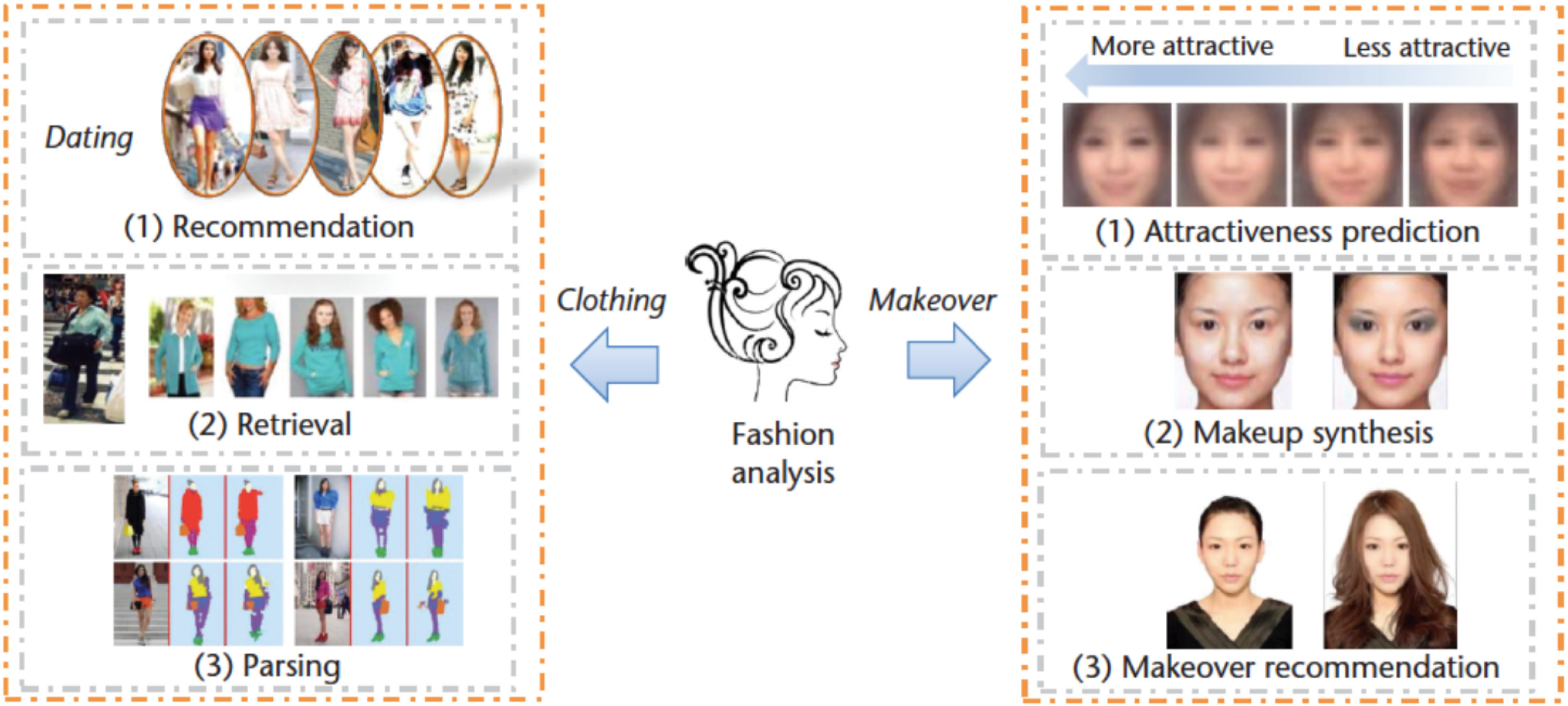
personality report

Learning Relevance by Neighbor Voting



[Ref] X. Li, C.G.M. Snoek, M. Worring, "Learning tag relevance by neighbor voting for social image retrieval," Proc. ACM Intl. Conf. Multimedia Information Retrieval (MIR), 2008.

More Influential Factors: Personal Fashion Flavor



[Ref] "Fashion Analysis: Current Techniques and Future Directions," IEEE Multimedia, 2014.

Clothing Fashion Analysis

- "i-Stylist: Finding the Right Dress Through Your Social Networks," MMM 2017.
- "A Framework of Enlarging Face Datasets Used for Makeup Face Analysis," BigMM 2016.
- "What are the Fashion Trends in New York?" MM 2014. (Grand Challenge Prize)
- "Clothing Genre Classification by Exploiting the Style Elements," MM 2012.

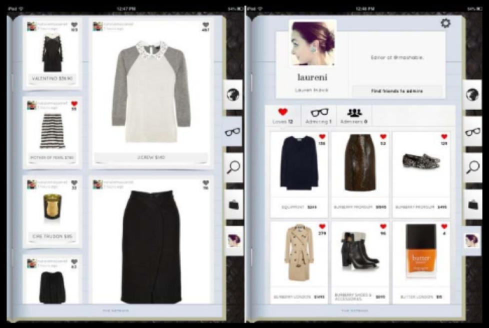
Clothing fashion is a reflection of the society of a period

- The global fashion apparel market today has surpassed 1 trillion US dollars since 2013, and accounts for nearly 2 percent of the world's Gross Domestic Product (GDP)

Applications: "Fashion is becoming mobile first with apps that help track down must-have clothes, accessories and shoes" – theguardian.com

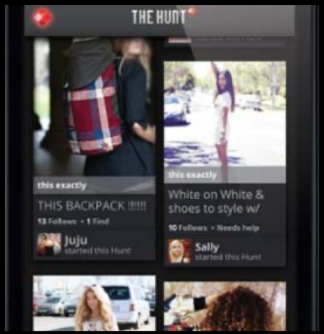
LIKEtoKNOW.it

The Netbook



Snap Fashion

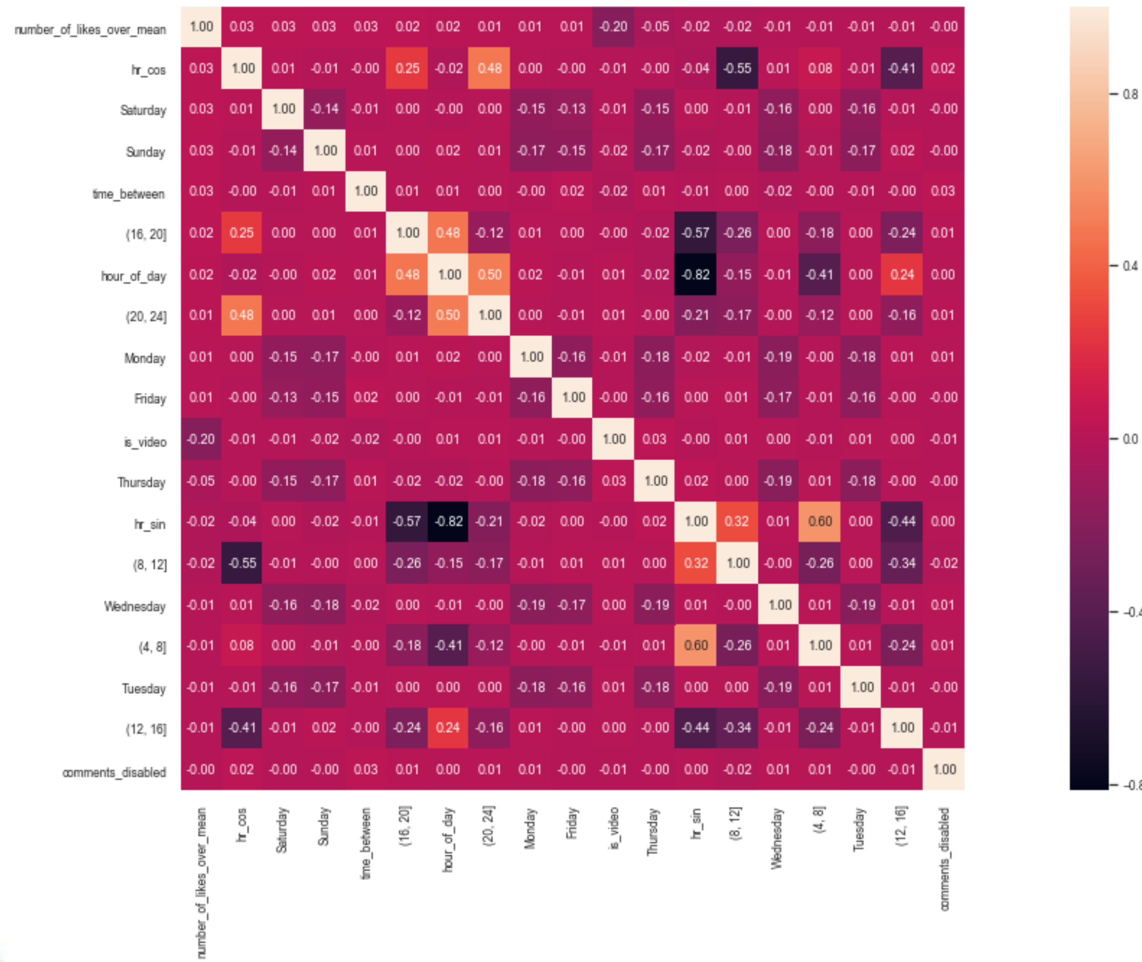
The Hunt



Predicting the Popularity of Instagram Posts

- Features:

	Attribute Pair	Correlation
1	(hour_of_day, hr_sin)	-0.815784
6	(Creators & Celebrities, is_business_account)	0.748046
5	(mean_comments, mean_likes)	0.649453
4	((4, 8], hr_sin)	0.600908
3	((16, 20], hr_sin)	-0.568493
7	(edge_followed_by, mean_likes)	0.551244
0	((8, 12], hr_cos)	-0.551196
2	((0, 4], hour_of_day)	-0.518755
8	((20, 24], hour_of_day)	0.502076



Summary

- Sociological understanding of humans and human interactions is fun but still a long way to go!

- People use their social networks to find interesting content
  - E.g., see stories friends post
  - This affects how popular stories become and how successful users are in having their stories promoted to the front page
    - Popular submitter advantage

- Simple phenomenological model explains dynamics of social voting
  - Story visibility (on front page, upcoming stories page, social stream): all parameters measured from data
  - Story interestingness: only adjustable parameter

  Model explains and predicts story popularity

THANK YOU