Opinion Mining and Sentiment Analysis

Text and Web Mining (H6751)

School of Communication and Information

Opinion Mining and Sentiment Analysis

- The rapid growth of ***user-generated content***, called ***social media***
  - Weblogs, Discussion Boards, User and Critic Review Web sites, Twitter, etc.
- Importance of social media and online opinions
  - Online shoppers are influenced by product reviews and are willing to pay more for products highly rated by other consumers.
  - Users are more influenced by reviews of fellow consumers rather than those generated by professionals.
- ***Opinion mining*** or ***sentiment analysis*** mines user generated content.
  - A type of subjectivity analysis which analyzes sentiment in a given textual unit with the objective of understanding the ***sentiment polarities (i.e. positive, negative, or neutral)*** of the opinions toward various aspects of a subject.
  - Both terms are used interchangeably.
  - Other names: Opinion extraction, Sentiment mining, and Subjectivity analysis

Example Applications of Opinion Mining and Sentiment Analysis

- *Movie*: is this review positive or negative?
- *Products*: what do people think about the new iPhone?
- *Public sentiment*: how is consumer confidence? Is despair increasing?
  - “**Consumer confidence** is an economic indicator which measures the degree of optimism that **consumers** feel about the overall state of the economy and their personal financial situation.” (from Wikipedia)
- *Politics*: what do people think about this candidate or issue?
- *Prediction*:
  - predict election outcomes or market trends from sentiment
  - predict stock prices (up and down) with sentiment analysis of user generated content.

Scherer Typology of Affective States

- **Emotion**: brief organically synchronized … evaluation of a major event
  - *angry, sad, joyful, fearful, ashamed, proud, elated, and desperate.*
- **Mood**: diffuse non-caused low-intensity long-duration change in subjective feeling
  - *cheerful, gloomy, irritable, listless, depressed, and buoyant*
- **Interpersonal stances**: affective stance toward another person in a specific interaction
  - *friendly, distant, cold, warm, supportive, contemptuous*
- **Attitudes**: enduring, affectively colored beliefs, dispositions towards objects or persons
  - *liking, loving, hating, valuing, desiring*
  - Sentiment analysis is the detection of **attitudes**
- **Personality traits**: stable personality dispositions and typical behavior tendencies
  - *nervous, anxious, reckless, morose, hostile, jealous*

What are opinions? (Liu, 2011)

- An ***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.
  - A restricted definition in sentiment analysis
- Opinion or Sentiment expressions on some target entities
  - ***Direct opinion***
    - *"The touch screen of iPhone is really cool."* (***subjective*** sentence)
  - ***Indirect opinion***
    - *"After taking the drug, my pain has gone."* (***objective*** sentence with implicit opinion)
- Sentiment orientation of an opinion
  - Positive, negative, or neutral (no or mixed opinion)

Opinion Definition (Liu, NLP handbook, 2010)

- **An opinion is a quintuple.**

  (e, a, so, h, t), where

  - ***e*** is a target entity.
  - ***a*** is an aspect/feature of the entity ***e***.
  - ***so*** is the sentiment value of the opinion from the opinion holder ***h*** on feature ***a*** of entity ***e*** at time ***t***.
    - ***so***'s sentiment value is positive, negative, or neutral, or more granular ratings.
  - ***h*** is an opinion holder.
  - ***t*** is the time when the opinion is expressed.

E.g., (iPhone8, battery, +, reviewer#1, 26th Oct. 2017)

Opinion Definition

- An example blog in quintuples
  - User ID: Abc123 on 5-1-2018 "*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, …*"
- In quintuples
  - (iPhone, general, +, Abc123, 5-1-2018)
  - (iPhone, touch_screen, +, Abc123, 5-1-2018)
  - (iPhone, voice_quality, +, Abc123, 5-1-2018)
  - (iPhone, price, -, Abc123's mother, 5-1-2018)
  - …

If Blackberry is considered as a target,
(Blackberry, general, -, Abc123, 5-1-2018) and (Blackberry, keyboard, -, Abc123, 5-1-2018)
Machine will be confused if you don't consider the current target.

Opinion Mining and Sentiment Analysis

- Mainly uses a ***supervised learning*** approach
- There are mainly two opinion mining tasks on texts:
  - ***Sentiment Classification***
    - Treats the problem as a text document classification problem with three classes (*positive, negative,* and *neutral)* or two classes *(positive and negative).*
    - More complex task is to rank the attitude of the text from 1 to 5 (Regression Problem).
  - ***Aspect-Based Sentiment Analysis and Summarization***
    - Looks at the **sentence level** of a text to discover what aspects of an object did people like or dislike.
      - E.g., *screen size* and *battery* aspects of the *iPhone* product
    - May summarizes analyzed results using visualization tools.

Sentiment Classification

- ***Topical Text Classification***
  - Classifying documents into various subjects.
    - E.g., ***Education vs. Sports***
  - Comparing individual words (bag-of-words or unigrams) in various subject areas. E.g.,
    - *"pupil", "teach", "course"* -> ***Education***
    - *"score", "referee", "ball"* -> ***Sports***
- ***Sentiment Text Classification***
  - Classifying documents according to the overall sentiment expressed in them
    - E.g., ***Positive vs. Negative***; ***Like vs. Dislike; Recommended vs. not Recommended***
  - More difficult compared to traditional topical classification. May need more linguistic processing. E.g.,
    - *"great", "excellent", "not bad", "never regret"* -> ***Positive***
    - *"horrible", "bad", "worst", "disappointed", "not satisfactory."* -> ***Negative***
  - Is sensitive to the domain.
    - *"This vacuum cleaner really sucks."* -> ***Positive*** (***sucks*** here means taking dust)
    - *"Storyline is predictable."* -> ***Negative*** (compare wtih *"The weather is predictable."*)

Document Pre-Processing

- A sample on-line review document
  - In this example, converted into a vector of terms with term frequency weighting factor: ***"document vector"***.

A Review Document

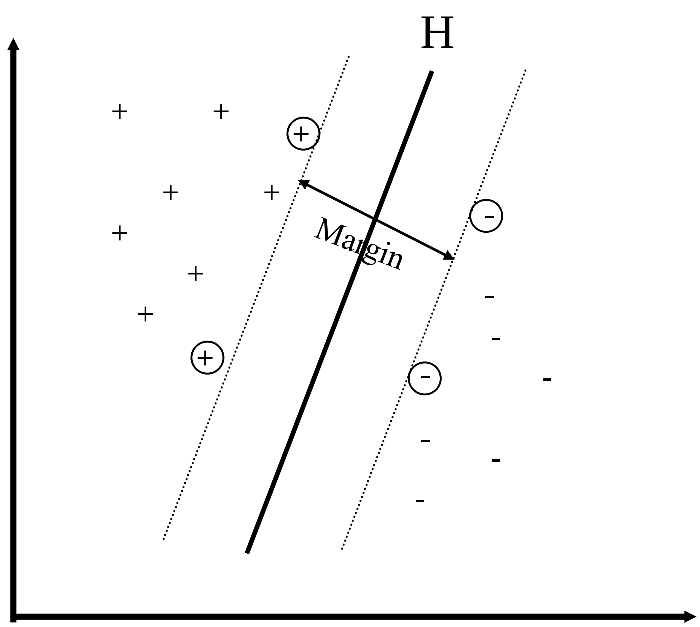
After 10 years of this phone, I have now decided to upgrade just because the digital tariffs are cheaper, and this phone does not do roaming. The phone has brilliant reception every where. This phone cannot be beaten for signal strength.

…….

TF	Terms
0	bad
1	brilliant
0	charge
1	cheap
1	digital
0	document
0	new
…	…
4	phone
0	receive
1	roam
0	sign
0	television
0	ultimate
0	unix

Machine Learning Algorithms

- SVM (Support Vector Machine)
  - Outperforms other machine-learning methods, such as ***decision tree induction*** and ***Naïve Bayes Classifier***.
  - Document vectors are input to SVM for sentiment classification.
  - Finds the hyperplane H, which separates the positive and negative examples with the maximum margin.



Sentiment Classification in Movie Reviews

- Polarity detection:
  - Is an IMDB movie review positive or negative?
- Data: *Polarity Data 2.0:* 1000 positive and 1000 negative reviews.
  - http://www.cs.cornell.edu/people/pabo/movie-review-data
- Papers
  - Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.
  - Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. ACL, 271-278

IMDB data in the Pang and Lee database

✓

in this movie , steven spielberg , one of today's finest directors , attempts to spice up the 1800s story of a long courtroom battle over the fate of prisoner cinque (djimon hounsou) - - a young angry man from sierra leone who was kidnapped into slavery - - and his fellow prisoners .

[…]

spielberg gives us a visually spicy and historically accurate real life story .

djimon hounsou and anthony hopkins turn in excellent performances .

✗

“ snake eyes ” is the most aggravating kind of movie : the kind that shows so much potential then becomes unbelievably disappointing .

it’s not just because this is a brian depalma film , and since he’s a great director and one who’s films are always greeted with at least some fanfare .

and it’s not even because this was a film starring nicolas cage and since he gives a bravura performance , this film is hardly worth his talents .

Baseline Algorithm (adapted from Pang and Lee)

- Tokenization
- Feature Extraction
  - Which words to use?
    - Only adjectives
    - All words
      - All words turns out to work better, at least on this data
- Classification using different classifiers
  - Naïve Bayes
  - SVM
  - MaxEnt (Maximum Entropy Model)
  - SVM and MaxEnt tend to do better than Naïve Bayes.

Problems:
What makes reviews hard to classify?

- Subtlety:
  - Perfume review in *Perfumes: the Guide*:
    - "If you are reading this because it is your darling fragrance, please wear it at home exclusively, and tape the windows shut."
  - Dorothy Parker on Katherine Hepburn
    - "She runs the gamut of emotions from A to B"

Thwarted Expectations and Ordering Effects

- “This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it **can’t hold up**.”
- Well as usual Keanu Reeves is nothing special, but surprisingly, the very talented Laurence Fishbourne is **not so good** either, I was surprised.

Analysis from wrongly classified documents

Possible reasons for failures in automatic sentiment classification of **mobile phone reviews**

Reason
Negation Phrase **E.g., "I'd *never* regretted purchasing it".**
Comments on parts **E.g. "The best phone I've had yet. The ONLY bad point is that…".**
Need inferencing **E.g., "if the price dropped, the company would be surprised how it would sell".**
Sarcasm **E.g., "What a great phone, it didn't work the first day".**
Comments on other products **E.g., "8210 is better. More valuable".**

Sentiment Tokenization Issues

- Deal with HTML and XML markup
- Emoticons (from Aproorv Agarwal et al, 2011: Sentiment Analysis of Twitter Data)
  - :-), :), :0), :], :3, :c) -> positive emoticons
  - :D, C: -> extremely-positive emoticons
  - :-(, :(, :c, :[-> negative emoticons
  - D8, D;, D=, DX, v.v -> Extremely-Negative emoticons
  - : | -> Neutral emoticons
- Acronym
  - Gr8, gr8t -> great
  - lol -> laughing out loud
  - Bff -> best friend forever

Sentiment Lexicon

- Sentiment Lexicons contain positive and negative terms.
  - E.g., The **General Inquirer** and **SentiWordNet**
    - Positive: good, excellent, fantastic, etc.
    - Negative: bad, evil, ridiculous, etc.
- Can be used in Sentiment Analysis
  - As a baseline approach, simply **count the numbers of positive and negative words** in a document or a sentence.
  - In a machine learning approach, used **as document features**, such as the number of positive words, and the number of negative words
    - Useful for sentiment classification of tweets
  - In a linguistic approach, **prior scores (e.g., terrible: -1.0; bad: -0.5; good:+0.5; great: +1.0) of words** in sentiment lexicons can be used to calculate a contextual sentiment score of a sentence.

Features	Acc.	F1 Measure	
		Pos	Neg
Unigram baseline	71.35	71.13	71.50
+ $f_5, f_6, f_7, f_{10}, f_{11}$	70.1	69.66	70.46
+ f_1, f_8	74.84	74.4	75.2
+ f_2, f_3, f_4, f_9	**75.39**	74.81	75.86

Table 6: Accuracy and F1-measure for 2-way classification task using Unigrams and Senti-features. All f_i refer to Table 4 and are cumulative.

Sentiment Document Features

- Document Features (from Aproorv Agarwal et al, 2011: Sentiment Analysis of Twitter Data)

$\mathbb{N}$	Polar	POS	# of (+/-) POS (JJ, RB, VB, NN)	f_1
		Other	# of negation words, positive words, negative words	f_2
			# of extremely-pos., extremely-neg., positive, negative emoticons	f_3
			# of (+/-) hashtags, capitalized words, exclamation words	f_4
	Non-Polar	POS	# of JJ, RB, VB, NN	f_5
		Other	# of slangs, latin alphabets, dictionary words, words	f_6
			# of hashtags, URLs, targets, newlines	f_7
$\mathbb{R}$	Polar	POS	For POS JJ, RB, VB, NN, $\sum$ prior pol. scores of words of that POS	f_8
		Other	$\sum$ prior polarity scores of all words	f_9
	Non-Polar	Other	percentage of capitalized text	f_{10}
$\mathbb{B}$	Non-Polar	Other	exclamation, capitalized text	f_{11}

Table 4: $\mathbb{N}$ refers to set of features whose value is a positive integer. They are primarily count features; for example, count of number of positive adverbs, negative verbs etc. $\mathbb{R}$ refers to features whose value is a real number; for example, sum of the prior polarity scores of words with part-of-speech of adjective/adverb/verb/noun, and sum of prior polarity scores of all words. $\mathbb{B}$ refers to the set of features that have a boolean value; for example, presence of exclamation marks, presence of capitalized text.

Negation

Das, Sanjiv and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific Finance Association Annual Conference (APFA).

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. EMNLP-2002, 79—86.

Add **NOT_** to every word between negation and following punctuation:

```
didn't like this movie , but I
```

```
didn't NOT_like NOT_this NOT_movie but I
```

Negation Phrase

Simple linguistic processing to address the problems of negation phrase (Na et al., 2005)

- Each negation (not, no, and never) and its adjacent words are combined to generate a new composite term (i.e., negation phrase)
- For extracting negation phrases automatically from corpus, use **syntactic patterns**, such as
  - <Verb>-<Negative Particle>-<Verb>
    - E.g., Do not buy
  - <Verb>-<Negative Particle>-<Adverb>-<Adjective>
    - E.g., Be not very impressed

Samples of Negation Phrases

Samples of Negation Phrases extracted automatically (out of 1,200 training reviews)

- Instead of using n-grams, negation phrases can be used as document features (# of features is smaller).

Negation Phrases	DF	Negation Phrases	DF
Do not buy	34	Will not regret	3
Do not work	24	Be not as good as	3
Would not recommend	14	Would not buy	2
Do not want	14	Be not very impressed	2
Do not like	9	Be not happy	2
Be not worth	6	Not so bad	2
Not bad	5	Do not purchase	1
Not the good	5	Do not dislike	1
Have not regret	4	Not so good	1
Will not work	4	Not too bad	1
Do not recommend	3	Be not a good choice	1

Sentiment Lexicon

- Sentiment Lexicons contain positive and negative terms.
  - The General Inquirer
  - MPQA Subjectiviy Lexicon
  - Bing Liu’s Opinion Lexicon
  - SentiWordNet
  - LIWC (Linguistic Inquiry and Word Count)

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
- Categories:
  - Positive (1915 words) and Negative (2291 words)
    - Positive: able, achievement, adaptive, etc.
    - Negative: abandon, abnormal, absence, etc.
  - Strong vs. Weak, Active vs. Passive, Overstated vs. Understated
  - Pleasure, Pain, Virtue, Vice, Goal, etc.
- Free for Research Use

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.

- Web page: http://mpqa.cs.pitt.edu/#subj_lexicon
- 2718 positive & 4912 negative terms
- Each word annotated for **intensity (strong, weak)**
- E.g.,
  - `type=`**`strongsubj`**` len=1 word1=`**`great`**` pos1=adj stemmed1=n priorpolarity=positive`
  - `type=`**`weaksubj`**` len=1 word1=`**`good`**` pos1=anypos stemmed1=n priorpolarity=positive`
  - `type=`**`strongsubj`**` len=1 word1=`**`bad`**` pos1=adj stemmed1=n priorpolarity=negative`
  - `type=`**`weaksubj`**` len=1 word1=`**`high`**` pos1=adj stemmed1=n priorpolarity=neutral`
- GNU GPL

Bing Liu’s 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
- **Misspelled words** are included in the list.
  - Included as these misspelled words appear frequently in social media content.
- 6786 words
  - 2006 positive
    - Abundant, accurate, achievable, **achievible**, achievement, agreeably, all-around, worthy, youthful, zest, etc.
  - 4783 negative
    - Abnormal, abort, absent-minded, addicted, assault, **assult**, worst, wound, zealous, etc.

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/: freely distributed for noncommercial use
- **All WordNet synsets** automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness
  - Accuracy is a bit questionable; Has scores for each sense of a term -> needs **word-sense disambiguation**

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://liwc.wpengine.com/
- 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*)
- Not free

Disagreements between polarity lexicons

Christopher Potts, Sentiment Tutorial, 2011

The disagreements arise mainly from **sense ambiguities**.

- Cheap (inexpensive) vs. Cheap (poor quality)

	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%)**

Semi-supervised learning of lexicons

- Use a small amount of information
  - A few labeled examples
  - A few hand-built patterns
- To bootstrap a lexicon

Hatzivassiloglou and McKeown intuition for identifying word polarity

Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the Semantic Orientation of Adjectives. ACL, 174–181

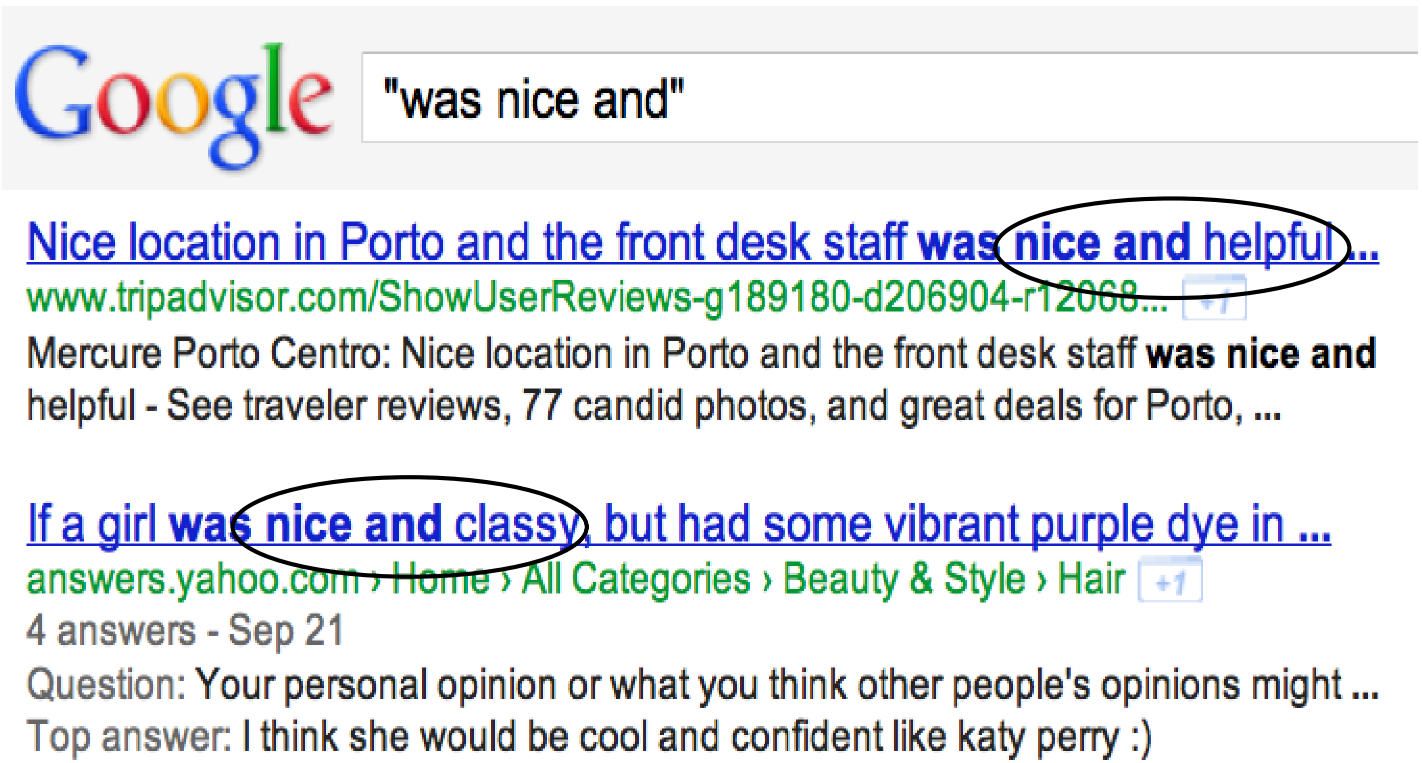
- Aimed to learn (or predict) the polarity of **adjective** terms.
- Adjectives conjoined by **"*and*"** have same polarity.
  - Fair **and** legitimate; corrupt **and** brutal
    - Unlikely cases: fair **and** brutal; corrupt **and** legitimate
- Adjectives conjoined by **"*but*"** do not
  - fair **but** brutal

Hatzivassiloglou & McKeown Step 1

- Label **seed set** of 1336 adjectives from a Wall Street Journal corpus (21 million words)
  - 657 positive
    - Adequate, central, clever, famous, intelligent, remarkable, reputed, sensitive, slender, thriving, …
  - 679 negative
    - Contagious, drunken, ignorant, lanky, listless, primitive, strident, troublesome, unresolved, unsuspecting, …

Hatzivassiloglou & McKeown Step 2

- Expand seed set to conjoined adjectives

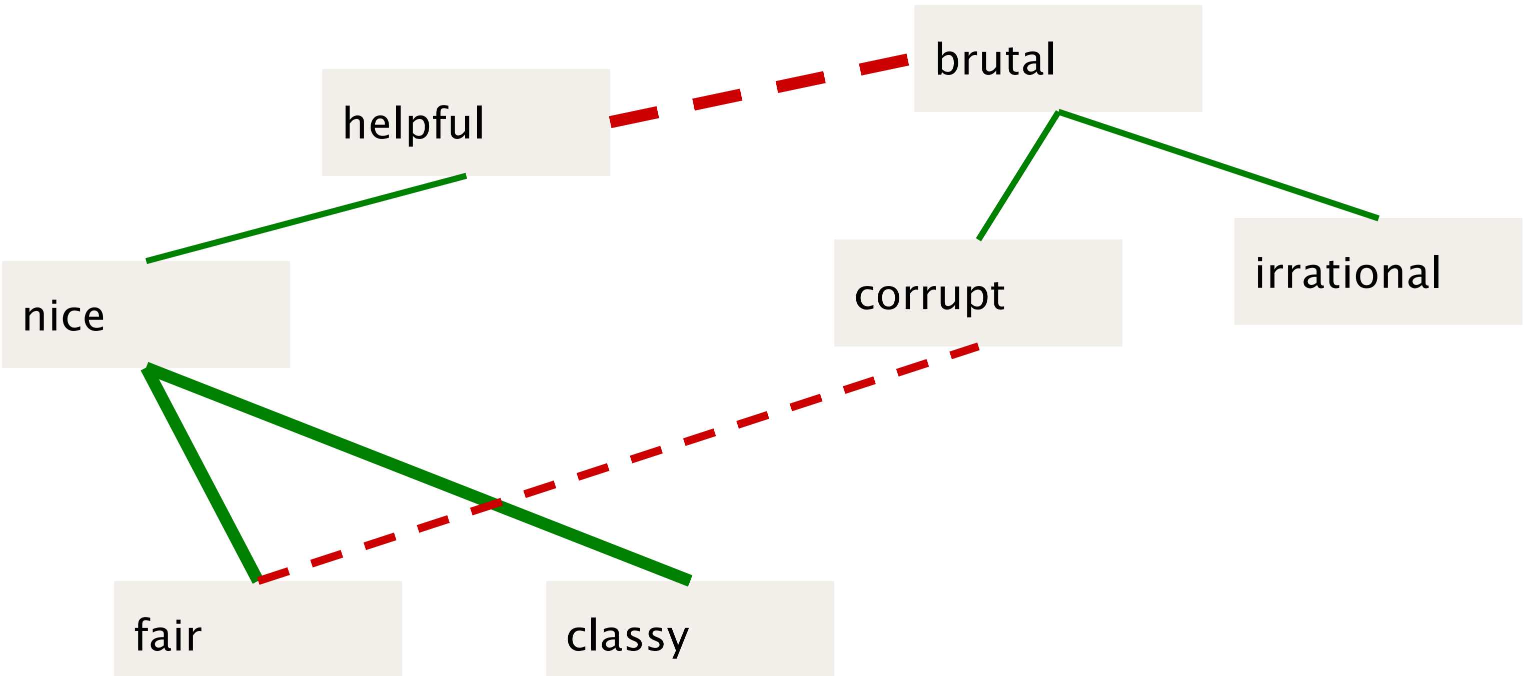


nice, helpful

nice, classy

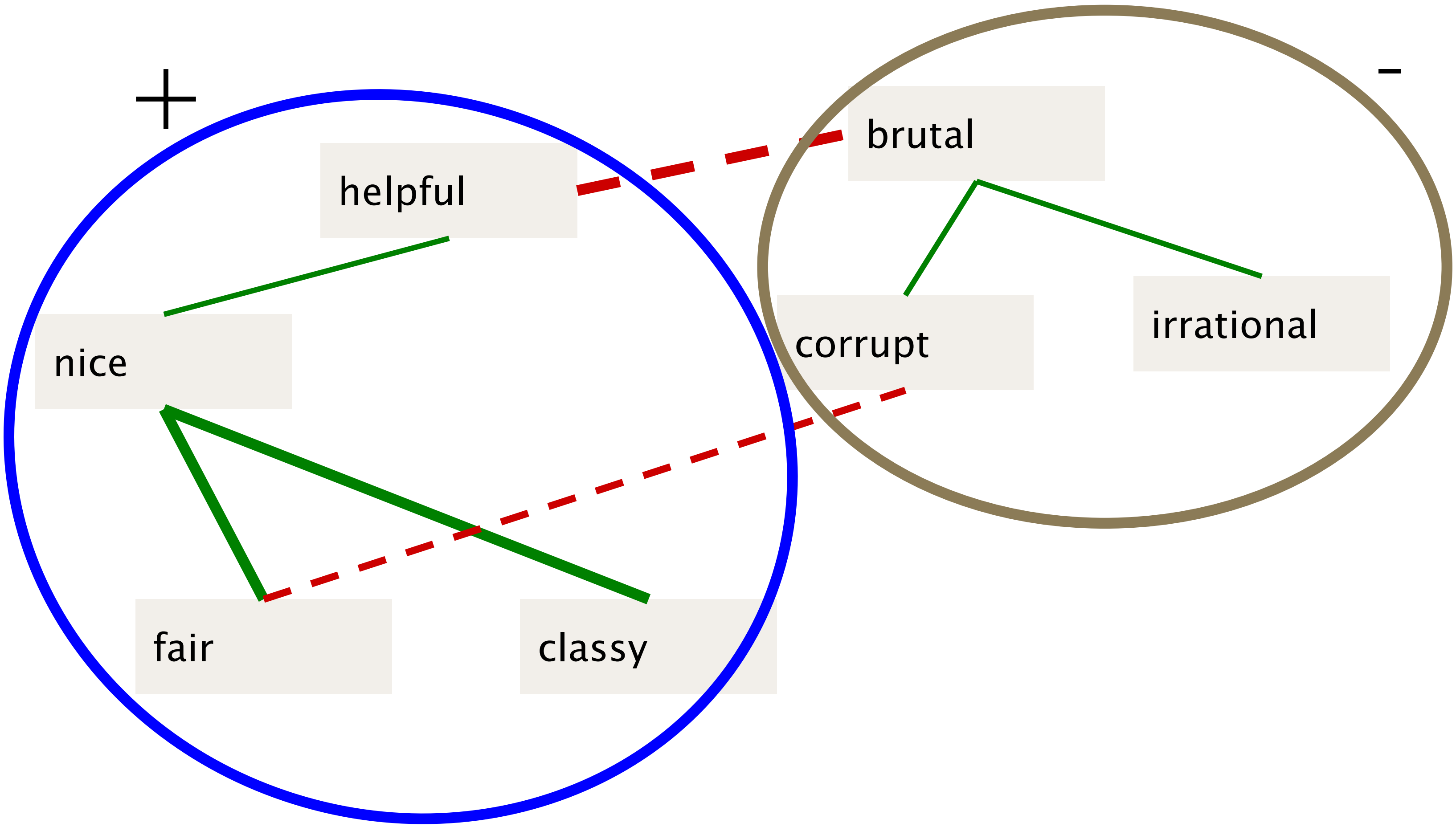
Hatzivassiloglou & McKeown Step 3

- Supervised classifier assigns "polarity similarity" to each word pair, resulting in graph:
  - Features: the observed counts from the corpus in the various conjunction categories, such as *and*, *or*, *but*, *either-or*, *neither-nor*, etc., for the particular adjective pair.



Hatzivassiloglou & McKeown Step 4

- Clustering for partitioning the graph into two.



Output polarity lexicon

- Positive
  - bold decisive **disturbing** generous good honest important large mature patient peaceful positive proud sound stimulating straightforward **strange** talented vigorous witty…
- Negative
  - ambiguous cautious cynical evasive harmful hypocritical inefficient insecure irrational irresponsible minor **outspoken** **pleasant** reckless risky selfish tedious unsupported vulnerable wasteful…

Turney Algorithm

Turney (2002): Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews

1. Extract a ***phrasal lexicon*** from reviews
2. Learn polarity of each phrase
   - E.g., very handy (+1.4); other problems (-2.8)
3. Rate a review by the average polarity of its phrases

Extract two-word phrases with adjectives

First Word	Second Word	Third Word (not extracted)
JJ	NN or NNS E.g., low fees	Anything
RB, RBR, RBS	JJ E.g., very handy	Not NN nor NNS
JJ	JJ E.g., handy little	Not NN or NNS
NN or NNS	JJ E.g., phones handy	Nor NN nor NNS
RB, RBR, or RBS	VB, VBD, VBN, VBG E.g., nicely work	anything

Alphabetical list of part-of-speech tags used in the Penn Treebank Project:
https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

How to measure polarity of a phrase?

- **Positive phrases** co-occur more with ***"excellent"***.
- **Negative phrases** co-occur more with ***"poor"***.
- But how to measure co-occurrence?
- **Pointwise mutual information (PMI)** between two words:
  - How much more do two words co-occur than if they were independent?

$$\mathrm{PMI}(word_1, word_2) = \log_2 \frac{P(word_1, word_2)}{P(word_1)P(word_2)}$$

How to Estimate Pointwise Mutual Information

- Query search engine
  - P(word) estimated by `hits(word)/N`
  - $P(word_1, word_2)$ by `hits(word1 NEAR word2)/N`^{2}

$$\mathrm{PMI}(word_1, word_2) = \log_2 \frac{hits(word_1\ \mathrm{NEAR}\ word_2)}{hits(word_1)hits(word_2)}$$

$$\mathrm{PMI}(word_1, word_2) = \log_2 \frac{P(word_1, word_2)}{P(word_1)P(word_2)}$$

*NEAR is for a **proximity search**

Does phrase appear more with "poor" or "excellent"?

$log_2(^m/_n) = log_2(m) - log_2(n)$

$$\text{Polarity}(\textit{phrase}) = \text{PMI}(\textit{phrase}, \text{"excellent"}) - \text{PMI}(\textit{phrase}, \text{"poor"})$$

$$= \log_2 \frac{\text{hits}(\textit{phrase}\ \text{NEAR "excellent"})}{\text{hits}(\textit{phrase})\text{hits}(\text{"excellent"})} - \log_2 \frac{\text{hits}(\textit{phrase}\ \text{NEAR "poor"})}{\text{hits}(\textit{phrase})\text{hits}(\text{"poor"})}$$

$$= \log_2 \frac{\text{hits}(\textit{phrase}\ \text{NEAR "excellent"})}{\text{hits}(\textit{phrase})\text{hits}(\text{"excellent"})} \frac{\text{hits}(\textit{phrase})\text{hits}(\text{"poor"})}{\text{hits}(\textit{phrase}\ \text{NEAR "poor"})}$$

$$= \log_2 \left(\frac{\text{hits}(\textit{phrase}\ \text{NEAR "excellent"})\text{hits}(\text{"poor"})}{\text{hits}(\textit{phrase}\ \text{NEAR "poor"})\text{hits}(\text{"excellent"})} \right)$$

Alternative approach:

$$\text{score}(w) = \log_2((f(w, \text{positive}) * f(\text{negative}))/(f(w, \text{negative}) * f(\text{positive})))$$

where *f(w, positive)* is the frequency of *w* in positive documents/sentences, and *f(positive)* is the total number of terms in positive documents/sentences. Similar definitions apply to *f(w, negative) and f(negative)*.

Phrases from a thumbs-up review

Phrase	POS tags	Polarity
online service	JJ NN	2.8
online experience	JJ NN	2.3
direct deposit	JJ NN	1.3
local branch	JJ NN	0.42
…		
low fees	JJ NNS	0.33
true service	JJ NN	-0.73
other bank	JJ NN	-0.85
inconveniently located	RB JJ	-1.5
Average		0.32

Phrases from a thumbs-down review

Phrase	POS tags	Polarity
direct deposits	JJ NNS	5.8
online web	JJ NN	1.9
very handy	RB JJ	1.4
…		
virtual monopoly	JJ NN	-2.0
lesser evil	RBR JJ	-2.3
other problems	JJ NNS	-2.8
low funds	JJ NNS	-6.8
unethical practices	JJ NNS	-8.5
Average		-1.2

Results of Turney algorithm

- 410 reviews from Epinions
  - 170 (41%) negative
  - 240 (59%) positive
- Majority class baseline: 59%
- Turney algorithm: 74%
- Phrases rather than words
- Learns domain-specific information
  - Can be used to learn a domain specific lexicon.

Using WordNet to learn polarity

S.M. Kim and E. Hovy. 2004. Determining the sentiment of opinions. COLING 2004
M. Hu and B. Liu. Mining and summarizing customer reviews. In Proceedings of KDD, 2004

- **WordNet**: online thesaurus
- Create positive ("good") and negative ("terrible") seed-words
- Find **Synonyms** and **Antonyms**
  - Positive Set: Add synonyms of positive words ("well") and antonyms of negative words ("comfortable")
  - Negative Set: Add synonyms of negative words ("awful") and antonyms of positive words ("evil")
- Repeat, following chains of synonyms
- Filter

Summary on Learning Lexicons

- Advantages:
  - Can be **domain-specific** compared to general sentiment lexicons, such as the General Inquirer.
  - Can be more robust (more words)
- Intuition
  - Start with a seed set of words ('good', 'poor')
  - Find other words that have similar polarity:
    - Using "and" and "but"
    - Using words that occur nearby in the same document (i.e. phrases with excellent vs. poor)
    - Using WordNet synonyms and antonyms
  - Use seeds and semi-supervised learning to induce lexicons

Finding sentiment of a sentence

- Important for finding **aspects** or attributes
  - Target of sentiment

- `The food was great but the service was awful.`

Finding aspect/attribute/target of sentiment

M. Hu and B. Liu. 2004. Mining and summarizing customer reviews. In Proceedings of KDD.
S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop.

- Frequent phrases + rules for aspect extraction
  - Find all highly frequent phrases across reviews (“`fish tacos`”)
  - Filter by rules like “occurs right after sentiment word”
    - “…**`great fish tacos`**” means `fish tacos` a likely aspect.

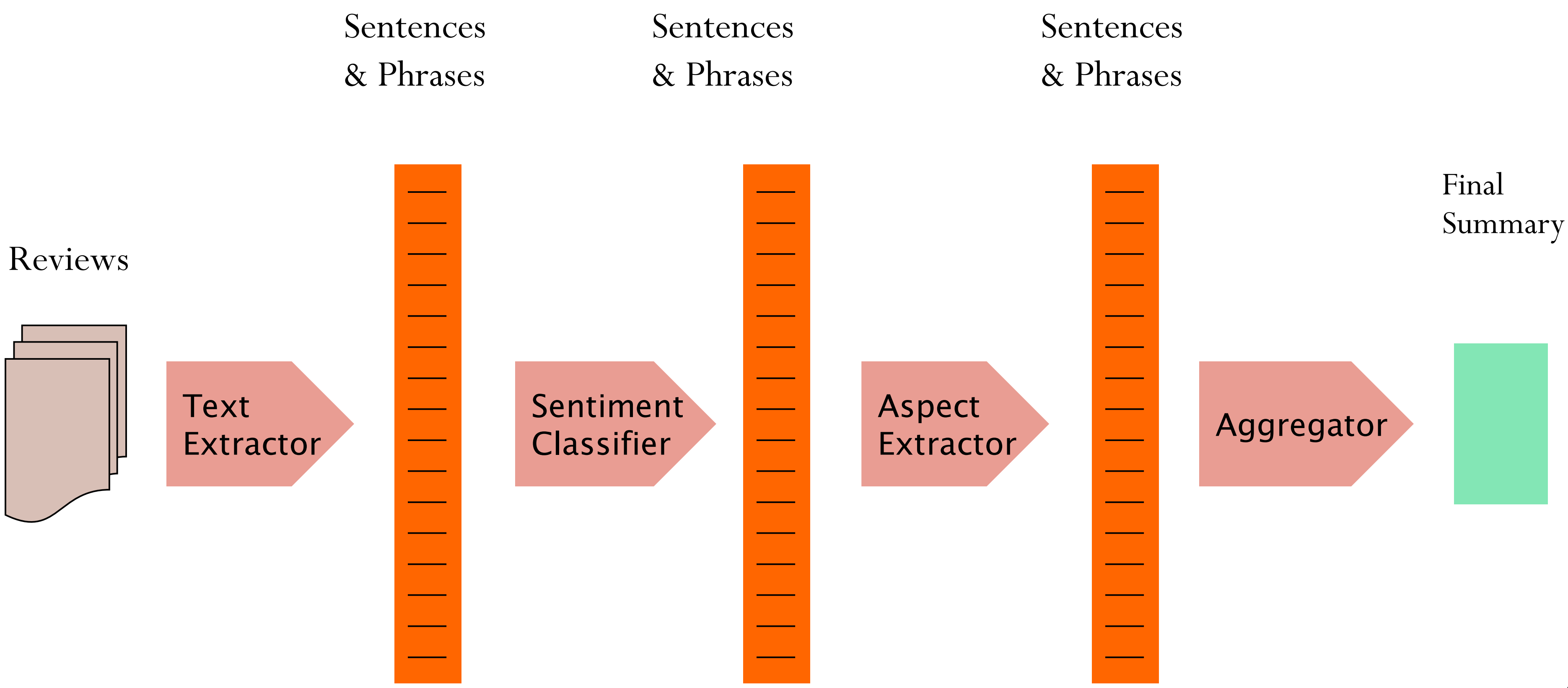
Casino	casino, buffet, pool, resort, beds
Children’s Barber	haircut, job, experience, kids
Greek Restaurant	food, wine, service, appetizer, lamb
Department Store	selection, department, sales, shop, clothing

Finding aspect/attribute/target of sentiment

- The aspect name may not be in the sentence.
  - **Value/Cost** aspect: "it is expensive".
- For restaurants/hotels, aspects are well-understood.
- Supervised classification for aspect detection
  - Hand-label a small corpus of restaurant review sentences with aspect
    - **food, décor, service, value, NONE**
  - Train a classifier to assign an aspect to a sentence
    - "Given this sentence, is the aspect *food, décor, service, value,* or *NONE* ?"

Putting it all together: Finding sentiment for aspects

S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a Sentiment Summarizer for Local Service Reviews. WWW Workshop



Results of Blair-Goldensohn et al. method

Hotel/Casino (46 Reviews)

Rooms (3/5 stars, 41 comments)

- (+) The room was clean and everything worked fine – even the water pressure ...
- (+) We went because of the free room and was pleasantly pleased ...
- (-) ...the worst hotel I had ever stayed at ...

Service (3/5 stars, 31 comments)

- (+) Upon checking out another couple was checking early due to a problem ...
- (+) Every single hotel staff member treated us great and answered every ...
- (-) The food is cold and the service gives new meaning to SLOW.

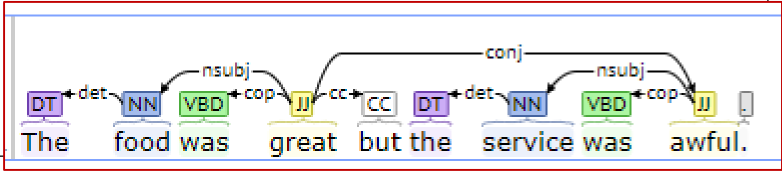
Dining (3/5 stars, 18 comments)

- (+) our favorite place to stay in biloxi. the food is great also the service ...
- (+) Offer of free buffet for joining the Play
- (-) The buffet at the casino was terrible.

Finding sentiment of Aspect Term in a sentence

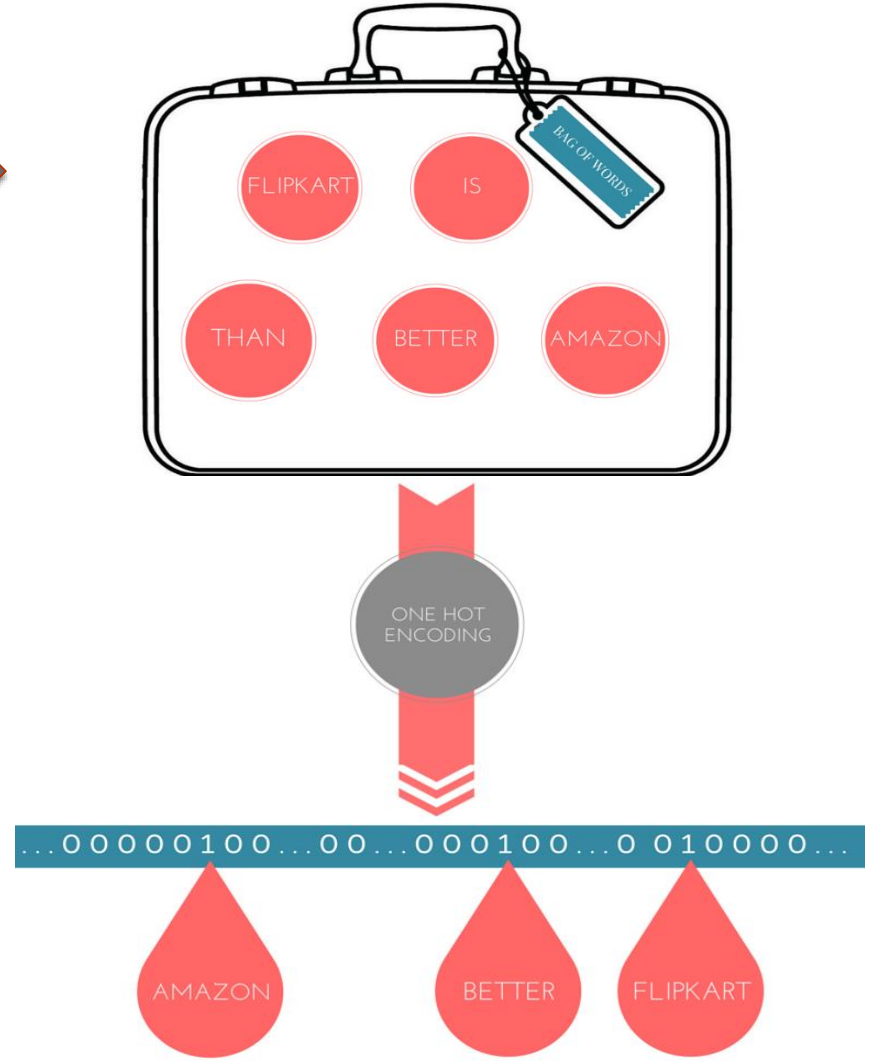
NRC-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews (2014)

- Detect **sentiment** expressed towards **a given aspect term**.
- E.g., "The food was great but the <u>service</u> was awful."
- Uses a supervised classifier in which the features are extracted from
  - The target term itself
  - Its surface context, i.e., a window of n words surrounding the term
  - The parse context, i.e., the nodes in the parse tree that are connected to the target term by at most three edges.
    - E.g., context-target bigrams, i.e., bigrams composed of a word from the parse context and a word from the term.
  - Lexicon features, i.e. the number of positive/negative tokens.



PROBLEMS WITH STATISTICAL MODELS

BAG OF WORDS



"FLIPKART IS BETTER THAN AMAZON"

PROBLEMS WITH STATISTICAL MODELS

- Word ordering information lost
- Data sparsity
- Words as atomic symbols
- Features other than BOW: Feature engineering is necessary
- Deep Neural Networks: Identify the features automatically

WORD EMBEDDINGS

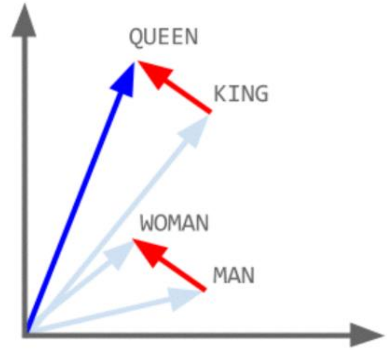
- **Word2Vec** and **GloVe** are unsupervised learning algorithms for obtaining **vector representations for words**.
- Training is performed on word-word co-occurrence statistics from a corpus.

WORD EMBEDDINGS

FRANCE	JESUS	XBOX	REDDISH	SCRATCHED	MEGABITS
AUSTRIA	GOD	AMIGA	GREENISH	NAILED	OCTETS
BELGIUM	SATI	PLAYSTATION	BLUISH	SMASHED	MB/S
GERMANY	CHRIST	MSX	PINKISH	PUNCHED	BIT/S
ITALY	SATAN	IPOD	PURPLISH	POPPED	BAUD
GREECE	KALI	SEGA	BROWNISH	CRIMPED	CARATS
SWEDEN	INDRA	PSNUMBER	GREYISH	SCRAPED	KBIT/S
NORWAY	VISHNU	HD	GRAYISH	SCREWED	MEGAHERTZ
EUROPE	ANANDA	DREAMCAST	WHITISH	SECTIONED	MEGAPIXELS
HUNGARY	PARVATI	GEFORCE	SILVERY	SLASHED	GBIT/S
SWITZERLAND	GRACE	CAPCOM	YELLOWISH	RIPPED	AMPERES

What words have embeddings closest to a given word? From Collobert *et al.* (2011)

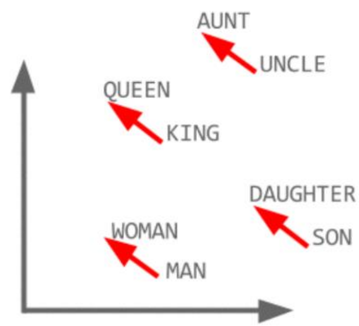
CAPTURE RELATIONSHIPS



$$W(\text{``woman''}) - W(\text{``man''}) \simeq W(\text{``aunt''}) - W(\text{``uncle''})$$

$$W(\text{``woman''}) - W(\text{``man''}) \simeq W(\text{``queen''}) - W(\text{``king''})$$

E.g., **vec(“King”) - vec(“Man”) + vec(“Woman”)** is closer to **vec(“Queen”)** than to any other word vector.



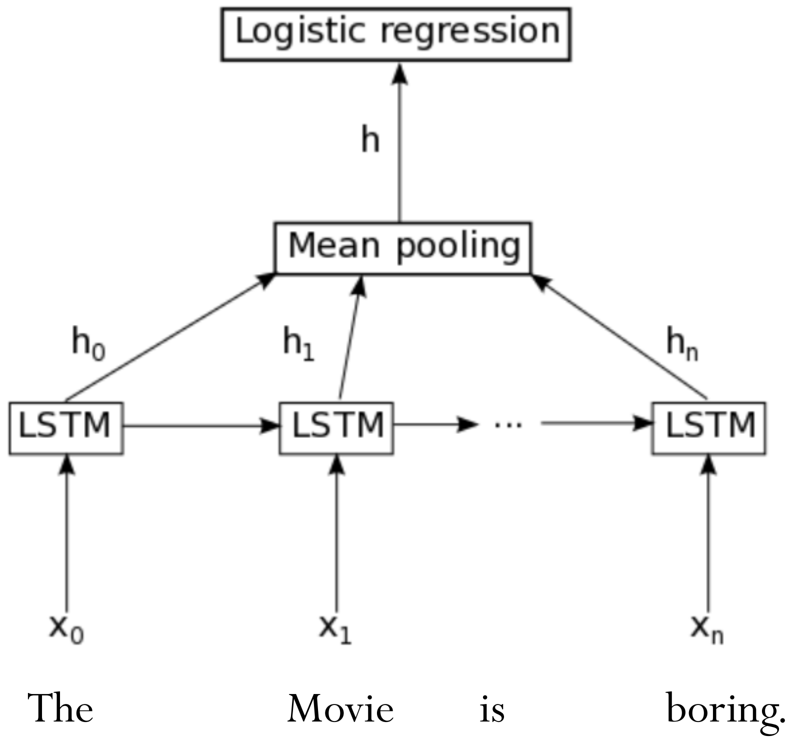
E.g., **vec(“Paris”) - vec(“France”) + vec(“Italy”)** is closer to **vec(“Rome”)** than to any other word vector.

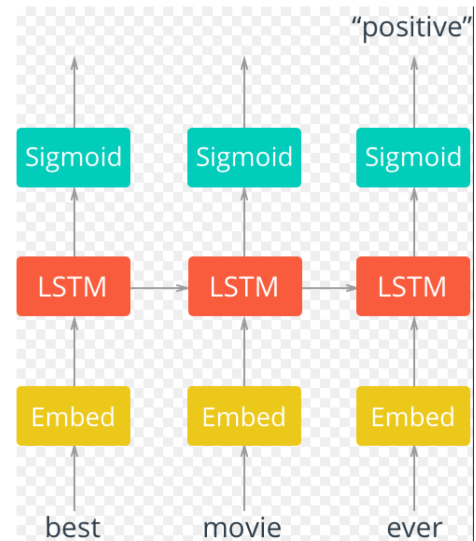
Relationship	Example 1	Example 2	Example 3
France - Paris	Italy: Rome	Japan: Tokyo	Florida: Tallahassee
big - bigger	small: larger	cold: colder	quick: quicker
Miami - Florida	Baltimore: Maryland	Dallas: Texas	Kona: Hawaii
Einstein - scientist	Messi: midfielder	Mozart: violinist	Picasso: painter
Sarkozy - France	Berlusconi: Italy	Merkel: Germany	Koizumi: Japan
copper - Cu	zinc: Zn	gold: Au	uranium: plutonium
Berlusconi - Silvio	Sarkozy: Nicolas	Putin: Medvedev	Obama: Barack
Microsoft - Windows	Google: Android	IBM: Linux	Apple: iPhone
Microsoft - Ballmer	Google: Yahoo	IBM: McNealy	Apple: Jobs
Japan - sushi	Germany: bratwurst	France: tapas	USA: pizza

Relationship pairs in a word embedding. From Mikolov *et al.* (2013b).

LSTM Networks for Sentiment Analysis

http://deeplearning.net/tutorial/lstm.html





- It is composed of a single Long Short Term Memory (LSTM) layer followed by mean pooling over time and logistic regression.
- This model is used to perform sentiment analysis on movie reviews from the Large Movie Review Dataset, known as the IMDB dataset.

Referenced Materials

- Sentiment Analysis and Opinion Mining, Bing Liu, Morgan & Claypool Publishers.
- Sentiment Analysis, Dan Jurafsky and Christopher Manning, http://www.standford.edu/~jurafsky/NLPCourseraSlides.html
- Fundamentals of Predictive Text Mining, Sholom M. Weiss, Nitin Indurkhya, and Tong Zhang, Springer.
  - Chapter 7
- Deep Learning For Natural Language Processing, Devashish Shankar and Prerana Singhal.