Text Classification

Text Classification

- Spam detection
- Authorship identification
- Age/gender identification
- Language identification
- Sentiment analysis
- ...

Text classification: problem definition

Input

- A document d
- A fixed set of classes $C = \{c_1, c_2, \dots, c_n\}$

Output

A predicted class $c \in C$

Classification Methods: Hand-coded rules

Rules based on combinations of words or other features

Spam

black-list-address OR (“dollars” AND “have been selected”)

Pros and Cons

Accuracy can be high if rules carefully refined by expert, *but building and maintaining these rules is expensive.*

Classification Methods: Unsupervised approaches

Clustering different classes of sentences (Graph-based)

- Encode each sentence (node) as TF-IDF vector
- Compute edge weights as cosine similarity between nodes
- Define threshold (optional) for edge pruning
- Employ clustering algorithms to find clusters
- For each new sentence, find the best cluster by comparing similarity between its vector and corresponding centroids.

Classification Methods: Supervised Machine Learning

- Naïve Bayes
- Logistic regression
- Support-vector machines
- Neural Networks

Naïve Bayes Intuition

- Simple classification method based on Bayes' rule
- Relies on very simple representation of document - Bag of words

For a document d and a class c

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)}$$

Naïve Bayes Classifier

$$c_{NB} =_{c \in C} P(c|d)$$

$$=_{c \in C} P(d|c)P(c)$$

$$=_{c \in C} P(x_1, x_2, \ldots, x_n|c)P(c)$$

Naïve Bayes classification assumptions

$$P(x_1, x_2, \ldots, x_n | c)$$

Bag of words assumption

Assume that the position of a word in the document doesn't matter

Conditional Independence

Assume the feature probabilities $P(x_i | c_j)$ are independent given the class c_j.

$$P(x_1, x_2, \ldots, x_n | c) = P(x_1 | c) \cdot P(x_2 | c) \ldots P(x_n | c)$$

$$c_{NB} =_{c \in C} P(c) \prod_{x \in X} P(x | c)$$

Learning the model parameters

Maximum Likelihood Estimate

$$\hat{P}(c_j) = \frac{doc-count(C = c_j)}{N_{doc}}$$

$$\hat{P}(w_i | c_j) = \frac{count(w_i, c_j)}{\sum_{w \in V} count(w, c_j)}$$

Problem with MLE

Use smoothing to assign nonzero likelihood probabilities to unseen words in given class.

A worked example

$$\hat{P}(c) = \frac{N_c}{N}$$

$$\hat{P}(w \mid c) = \frac{count(w,c)+1}{count(c)+|V|}$$

	Doc	Words	Class
Training	1	Chinese Beijing Chinese	c
	2	Chinese Chinese Shanghai	c
	3	Chinese Macao	c
	4	Tokyo Japan Chinese	j
Test	5	Chinese Chinese Chinese Tokyo Japan	?

Priors:

$P(c) = \frac{3}{4}$

$P(j) = \frac{1}{4}$

Choosing a class:

$P(c|d5) \propto 3/4 * (3/7)^3 * 1/14 * 1/14 \approx 0.0003$

$P(j|d5) \propto 1/4 * (2/9)^3 * 2/9 * 2/9 \approx 0.0001$

Conditional Probabilities:

$P(\text{Chinese}|c) = (5+1) / (8+6) = 6/14 = 3/7$

$P(\text{Tokyo}|c) = (0+1) / (8+6) = 1/14$

$P(\text{Japan}|c) = (0+1) / (8+6) = 1/14$

$P(\text{Chinese}|j) = (1+1) / (3+6) = 2/9$

$P(\text{Tokyo}|j) = (1+1) / (3+6) = 2/9$

$P(\text{Japan}|j) = (1+1) / (3+6) = 2/9$

Machine Learning classifiers: Others

- Design feature set
- Feature set: {TF-IDF, syntactic features, dictionary-based, etc.}
- Partition dataset into train, development and test.
- Learn on training data, optimize parameters on development and evaluate on test dataset.

Evaluation: Constructing Confusion matrix c

For each pair of classes $< c_1, c_2 >$ how many documents from c_1 were incorrectly assigned to c_2? (when $c_2 \neq c_1$)

Docs in test set	Assigned UK	Assigned poultry	Assigned wheat	Assigned coffee	Assigned interest	Assigned trade
True UK	95	1	13	0	1	0
True poultry	0	1	0	0	0	0
True wheat	10	90	0	1	0	0
True coffee	0	0	0	34	3	7
True interest	-	1	2	13	26	5
True trade	0	0	2	14	5	10

Per class evaluation measures

Recall

Fraction of docs in class i classified correctly: $\frac{c_{ii}}{\sum_j c_{ij}}$

Precision

Fraction of docs assigned class i that are actually about class i: $\frac{c_{ii}}{\sum_i c_{ji}}$

Accuracy

Fraction of docs classified correctly: $\frac{\sum_i c_{ii}}{N}$

Micro- vs. Macro-Average

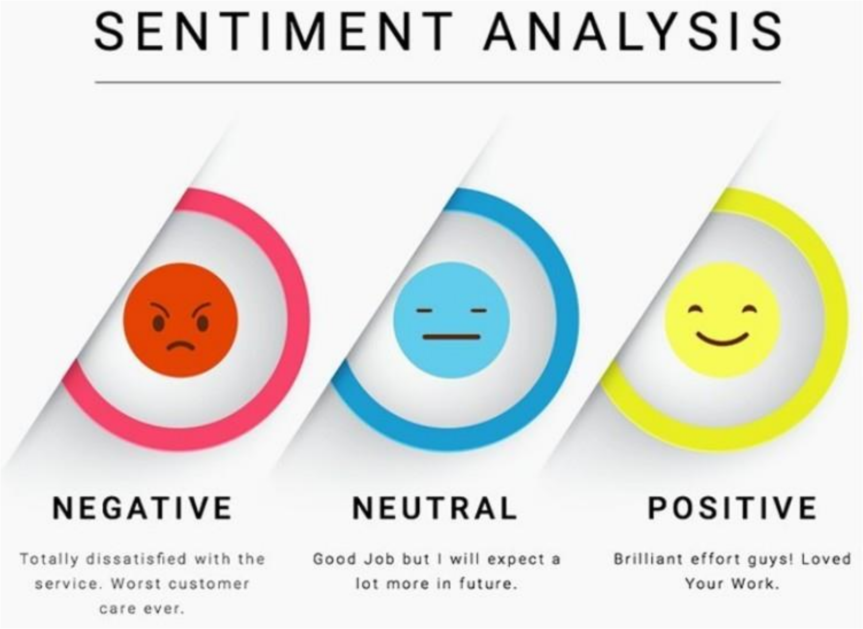
If we have more than one class, how do we combine multiple performance measures into one quantity?

Macro-averaging

Compute performance for each class, then average

Micro-averaging

Collect decisions for all the classes, compute contingency table, evaluate.


SENTIMENT ANALYSIS
NEGATIVE
Totally dissatisfied with the service. Worst customer care ever.
NEUTRAL
Good Job but I will expect a lot more in future.
POSITIVE
Brilliant effort guys! Loved Your Work.

Sentiments are everywhere!

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?
Politics	What do people think about this candidate or issue?
Prediction	Predict election outcomes or marked trends from sentiment

Affective States Typology

Emotion: angry, sad, joyful, fearful, ashamed, proud, elated

Mood: cheerful, gloomy, irritable, listless, depressed, buoyant

Interpersonal stances: friendly, flirtatious, distant, cold, warm, supportive, contemptuous

Attitudes: liking, loving, hating, valuing, desiring

Personality Traits: nervous, anxious, reckless, morose, hostile, jealous

Sentiment Analysis

Sentiment Analysis is the detection of attitudes

enduring, affectively colored beliefs, dispositions towards objects or persons

The complete task

- Holder (source) of attitude
- Target (aspect) of attitude
- Type of attitude
  - From a set of types: *like, love, hate, value, desire*
  - Or simple weighted polarity: *positive, negative, neutral, together with strength*
- Text containing the attitude

Sentiment Analysis

Simplest Task

Is the attitude of this text positive or negative?

More complex

Rank the attitudes of this text from 1 to 5

Advanced

Detect the target, source, or complex attitude types

ML Pipeline

- Tokenization
- Feature Extraction
- Classification using different classifiers
  - Naïve Bayes
  - MaxEnt
  - SVM
  - NN

Tokenization Issues

- Capitalization - preserve words in all caps (HATE, COOL, etc.)
- Word lengthening (coooooollllll, badddd, etc.)
- Handling emoticons (https://emojipedia.org/people/)

```
[<>]?                          # optional hat/brow
[:;=8]                         # eyes
[\-o\*\']?                     # optional nose
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
|                              #### reverse orientation
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
[\-o\*\']?                     # optional nose
[:;=8]                         # eyes
[<>]?                          # optional hat/brow
```

- Handling negation
  - I **didn't** like this movie
  - I really like this movie

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..

Is sentiment analysis so easy?

Is a given review on a known topic positive or negative?

"It may be a bit early to make such judgments, but Battlefield Earth may well turn out to be the worst movie of this century." (Elvis Mitchell, May 12, 2000)

Is sentiment analysis so easy?

Is a given review on a known topic positive or negative?

> "It may be a bit early to make such judgments, but Battlefield Earth may well turn out to be the worst movie of this century." (Elvis Mitchell, May 12, 2000)

don't we just need to look for "worst", "best", "love", hate", etc.?

In a small scale experiment (Pang et al., 2002)

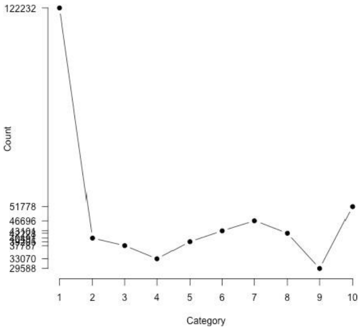
	Proposed word lists	Accuracy
Human 1	Positive: dazzling, brilliant, phenomenal, excellent, fantastic Negative: suck, terrible, awful, unwatchable, hideous	58%
Human 2	Positive: gripping, mesmerizing, riveting, spectacular, cool, awesome, thrilling, badass, excellent, moving, exciting Negative: bad, cliched, sucks, boring, stupid, slow	64%
Statistics-based	Positive: love, wonderful, best, great, superb, beautiful, **still** Negative: bad, worst, stupid, waste, boring, **?**, **!**	69%

Why can’t we just look for words like “great” and “terrible”?

- This laptop is *a great deal*.
- *A great deal* of media attention surrounded the release of the new laptop.
- 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.

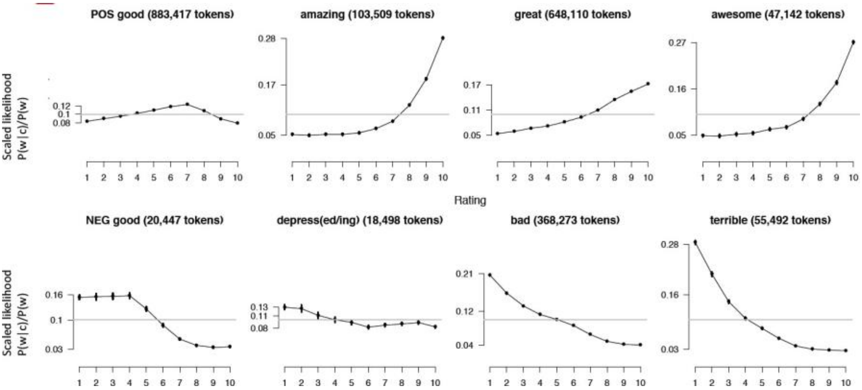
Analyzing polarity of each word in IMDB

- How likely is each word to appear in each sentiment class?
- Let’s take count(“bad”) in 1-star, 2-star, 3-star etc.



- We should use likelihood instead of counts: $P(w|c) = \frac{f(w,c)}{\sum_{w \in c} f(w,c)}$
- Make them comparable between words scaled likelihood: $\frac{P(w|c)}{P(w)}$

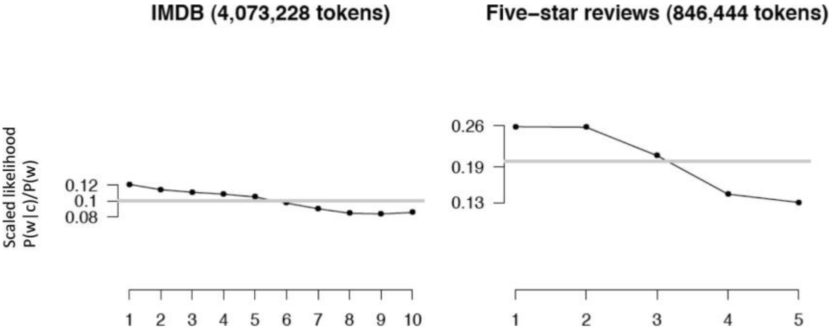
Analyzing polarity in IMDB



Logical Negation

- Is logical negation (no, not) associated with negative sentiment?
- Potts experiment:
  - Count negation (not, n’t, no, never) in online reviews
  - Regress against the review rating

More negation in negative sentiment



Using Linguistic Intuitions

Using a sentiment lexicon also works.

Some linguistic intuitions on top of that tends to give better results.

Handling negation in simple addition of scores

Example words

- Excellent +5
- good +3
- terrible -5
- bad -3

Reversing the polarity

- Not Excellent -5
- Not good -3
- Not terrible +5
- Not bad +3

Handling negation in simple addition of scores

Example words

- Excellent +5
- good +3
- terrible -5
- bad -3

Instead, a polarity shift works better

- Not Excellent (5-4) +1
- Not good (3-4) -1
- Not terrible (-5+4) -1
- Not bad (-3+4) 1

Handling Intensifiers

Intensifiers can be classified into two major categories,

- Amplifiers (e.g., very) increase the semantic intensity
- Downtoners (e.g., slightly) decrease it

Rough values for some intensifiers

Intensifier	Modifier (%)
slightly	−50
somewhat	−30
pretty	−10
really	+15
very	+25
extraordinarily	+50
(the) most	+100

Somewhat sleazy

sleazy: -3, *somewhat sleazy*: $-3 \times (100\% - 30\%) = -2.1$

Sentiment Lexicons

- The general inquirer
- MPQA Subjectivity Cues Lexicon
- SentiWordnet
- LIWC (Linguistic Inquiry and Word Count)

The General Inquirer

http://www.wjh.harvard.edu/~inquirer/homecat.htm

Categories

- Positive (1915 words) and Negative (2291) words
- Strong vs weak, active vs passive, overstated vs understated
- pleasure, pain, virtue, vice, motivation, cognitive orientation etc.

Entry	Source	Positiv	Negativ	Pstv	Affil	Ngtv	Hostile	Strong	Power	Weak	Submit	Active	Pa
A	H4Lvd												
ABANDON	H4Lvd		Negativ			Ngtv				Weak			
ABANDONMENT	H4		Negativ							Weak			
ABATE	H4Lvd		Negativ										Pa
ABATEMENT	Lvd												
ABDICATE	H4		Negativ							Weak	Submit		Pa
ABHOR	H4		Negativ				Hostile						Pa
ABIDE	H4	Positiv			Affil							Active	
ABILITY	H4Lvd	Positiv						Strong					
ABJECT	H4		Negativ							Weak	Submit		Pa
ABLE	H4Lvd	Positiv		Pstv				Strong					
ABNORMAL	H4Lvd		Negativ			Ngtv							
ABOARD	H4Lvd												
ABOLISH	H4Lvd		Negativ			Ngtv	Hostile	Strong	Power			Active	
ABOLITION	Lvd												
ABOMINABLE	H4		Negativ					Strong					
ABORTIVE	Lvd												
ABOUND	H4	Positiv											Pa
ABOUT#1	H4Lvd												
ABOUT#2	H4Lvd												
ABOUT#3	H4Lvd												
ABOUT#4	H4Lvd												
ABOUT#5	H4Lvd												
ABOUT#6	H4Lvd												

SentiWordNet

All Wordnet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness URL:
https://github.com/aesuli/sentiwordnet

Using NLTK

```
>>> from nltk.corpus import sentiwordnet as swn
>>> breakdown = swn.senti_synset('breakdown.n.03')
>>> print(breakdown)
<breakdown.n.03: PosScore=0.0 NegScore=0.25>
```

SentiWordNet

All Wordnet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness URL: https://github.com/aesuli/sentiwordnet

Using NLTK

```
>>> from nltk.corpus import sentiwordnet as swn

>>> breakdown = swn.senti_synset('breakdown.n.03')
>>> print(breakdown)
<breakdown.n.03: PosScore=0.0 NegScore=0.25>

>>> list(swn.senti_synsets('slow')) # doctest: +NORMALIZE_WHITESPACE
[SentiSynset('decelerate.v.01'), SentiSynset('slow.v.02'),
SentiSynset('slow.v.03'), SentiSynset('slow.a.01'),
SentiSynset('slow.a.02'), SentiSynset('slow.a.04'),
SentiSynset('slowly.r.01'), SentiSynset('behind.r.03')]
```

Other Lexicons

MPQA Subjectivity Cues Lexicon

- 6885 words from 8221 lemmas: 2718 positive, 4912 negative
- Each word annotated for intensity (strong, weak)
- URL: https://mpqa.cs.pitt.edu/lexicons/subj_lexicon

(5) Thousands of coup supporters celebrated (**positive**) overnight, waving flags, blowing whistles ...

(6) The criteria set by Rice are the following: the three countries in question are repressive (**negative**) and grave human rights violators (**negative**) ...

(7) Besides, politicians refer to good and evil (**both**) only for purposes of intimidation and exaggeration.

(8) Jerome says the hospital feels (**neutral**) no different than a hospital in the states.

Bing Liu Opinion Lexicon

6786 word: 2006 positive, 4780 negative

https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html

Learning Sentiment Lexicons

Basic Intuition

- Adjectives conjoined by “and” have same polarity
    - *Fair and legitimate*, *corrupt and brutal*
- Adjectives conjoined by “but” do not
    - *fair but brutal*

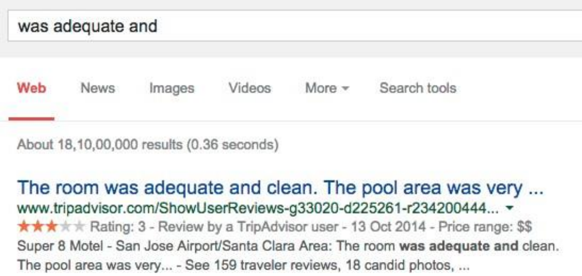
Learning Sentiment Lexicons

Step 1: Label **seed set** of adjectives

- **Positive cases:** adequate, central, clever, famous, intelligent, remarkable, reputed, sensitive, slender, thriving ...
- **Negative cases:** contagious, drunken, ignorant, lanky, listless, primitive, strident, troublesome, unresolved, unsuspecting ...

Learning Sentiment Lexicons

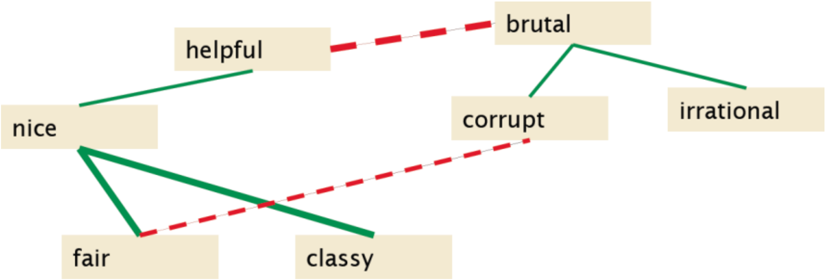
Step 2: Expand seed set to conjoined adjectives



Learning Sentiment Lexicons

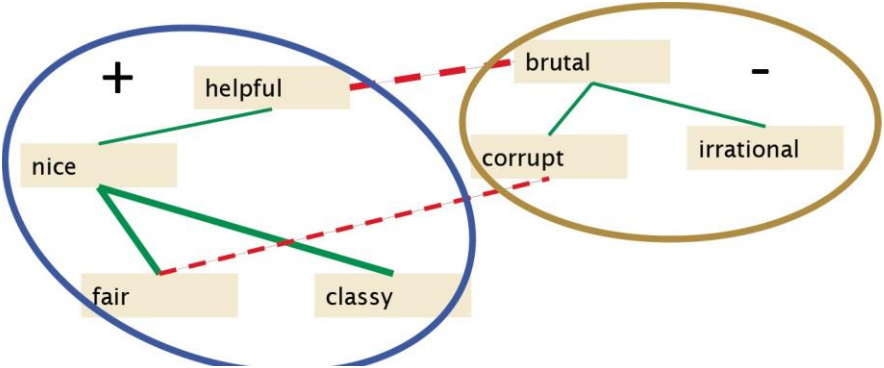
Step 3: Construct a graph

Polarity similarity is assigned to each word pair:



Learning Sentiment Lexicons

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 ...

Finding aspects or attributes

The food was great but the service was awful.

Aspects Involved

Food, service

Frequent phrases + rules

- 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

If the aspects are well understood, use supervised classification.

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

Sentiment Tutorial by Christopher Potts

URL:http://sentiment.christopherpotts.net/

Text scoring

*Shows how a variety of **sentiment lexicons** score novel texts. Such values could be used in many ways (as raw values, to derive percentages or ratios, as **classifier features**, ...).*

Enter your own text (max 140 characters), or

It sounds like a great plot but can't hold up.

☐ analyze a random tweet instead.

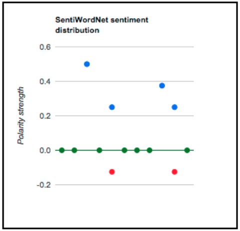
Submit

Scores

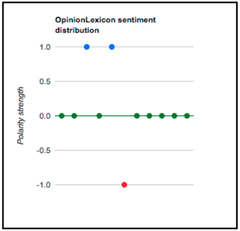
```
it sounds like a great plot but can't hold_NEG up_NEG .
```

Sentiment Tutorial by Christopher Potts

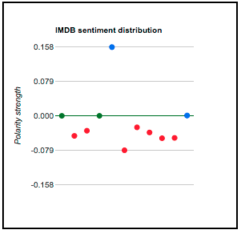
SentiWordNet



Opinion Lexicon



IMDB



IMDB 2d