





# Sentiment Analysis

# Example: Positive or negative movie review?

-  • unbelievably disappointing
-  • Full of zany characters and richly applied satire, and some great plot twists
-  • this is the greatest screwball comedy ever filmed
-  • It was pathetic. The worst part about it was the boxing scenes.

# Where is Sentiment Analysis Used?

- **Movie** - Is this review positive or negative?
- **Products** - What do people think about the new iPhone?
- **Politics** - What do people think about this candidate or issue?
- **Prediction** - Predict election outcomes or market trends from sentiment
- **Frustration** of callers to a help line
- **Stress** in drivers or pilots
- **Depression** and other medical conditions from social media

# 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.

**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

Given a text-

## **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

# Sentiment Analysis in Movie Reviews

## Polarity detection:

- Is an IMDB movie review positive or negative?

★ 10/10

**Outstanding!!!!!!!**

jamlu-16624 9 March 2019

Badla is one of the best thriller I have seen in recent times! Edge to seat suspense till the end. Amitji was terrific and tapsee as usual doing great job!

★ 9/10

**Amazing. Worth the watch.**

lillmshadab 8 March 2019

Its an amazing movie, 1 star less just cuz its a remake. Both characters r well written n shud i talk abt acting skills of AB.. NAA... Let his records speak fr him.

21 out of 51 found this helpful. Was this review helpful? [Sign in](#) to vote.

[Permalink](#)

**Bad remake of The Invisible Guest**

dini1214 10 March 2019

This is not what i was expecting from this movie , this is just out and out copy of the masterpiece spanish movie , they should have change climax scene a bit , they just copied it ..bad background music , the only +point in this movie is performance by Taapshee and Big b ..... plz watch original spanish movie which is far better than badla in every department 🙄

21 out of 43 found this helpful. Was this review helpful? [Sign in](#) to vote.

[Permalink](#)

# Baseline Algorithm

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



# Tokenization Issues

- Capitalization - preserve for word in all caps
- Word lengthening
- Handling emoticons
- Handling negation
  - I didn't **like** this movie
  - I really **like** this movie

Add NOT\_ to every word between negation and following till punctuation

- didn't like this movie, but I ...
- didn't NOT\_like NOT\_this NOT\_movie but I..

# Naïve Bayes

$$c_{NB} = \arg \max_{c_j \in C} P(c_j) \prod_{x_i} P(x_i | c_j)$$

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

$$P(w_i | c_j) = \frac{\textit{count}(w_i, c_j) + 1}{\left( \sum_{w \in V} (\textit{count}(w, c_j)) \right) + |V|}$$

# Example:

Is a given review on a movie Badla -positive or negative?

Badla is one of the best thriller I have seen in recent times! Edge to seat suspense till the end. Amitji was **terrific** and tapsee as usual doing **great** job!

Why can't we just look for words like "great", "cool", "worst", "best", "love", "hate", etc... for doing SA.

# 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 laptop is a great deal ... and I've got a nice offer you might be interested in.
- 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.

# Sentiment Analysis - Affective Lexicons

# Sentiment Lexicons

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

# The General Inquirer

## 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	Ngvtv	Hostile	Strong	Power	Weak	Submit	Active	Passive
A	H4Lvd												
ABANDON	H4Lvd		Negativ			Ngvtv				Weak			
ABANDONMENT	H4		Negativ							Weak			
ABATE	H4Lvd		Negativ										Passive
ABATEMENT	Lvd												
ABDICATE	H4		Negativ							Weak	Submit		Passive
ABHOR	H4		Negativ				Hostile						Passive
ABIDE	H4	Positiv			Affil							Active	
ABILITY	H4Lvd	Positiv						Strong					
ABJECT	H4		Negativ							Weak	Submit		Passive
ABLE	H4Lvd	Positiv		Pstv				Strong					
ABNORMAL	H4Lvd		Negativ			Ngvtv							
ABOARD	H4Lvd												
ABOLISH	H4Lvd		Negativ			Ngvtv	Hostile	Strong	Power			Active	
ABOLITION	Lvd												
ABOMINABLE	H4		Negativ					Strong					
ABORTIVE	Lvd												
ABOUND	H4	Positiv											Passive
ABOUT#1	H4Lvd												
ABOUT#2	H4Lvd												
ABOUT#3	H4Lvd												
ABOUT#4	H4Lvd												
ABOUT#5	H4Lvd												
ABOUT#6	H4Lvd												
ABOUT#7	H4Lvd												
ABOVE#1	H4Lvd												
ABOVE#2	H4Lvd												
ABOVE#3	H4Lvd												
ABOVE#4	H4Lvd												
ABRASIVE	H4		Negativ				Hostile	Strong					
ABROAD	H4Lvd												
ABRUPT	H4Lvd		Negativ			Ngvtv							
ABSCOND	H4		Negativ				Hostile			Weak	Submit	Active	
ABSENCE	H4Lvd		Negativ							Weak			
ABSENT#1	H4Lvd		Negativ							Weak			Passive
ABSENT#2	H4Lvd												Passive
ABSENT-MINDED	H4		Negativ							Weak			Passive

# SentiWordNet

- Home page: <http://sentiwordnet.isti.cnr.it/>
- All Wordnet synsets automatically annotated for degrees of positivity, negativity, and neutrality/objectiveness

## Example:

**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



# MPQA Subjectivity Cues Lexicon

- Home page:  
[http://www.cs.pitt.edu/mpqa/subj\\_lexicon.html](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)

# LIWC (Linguistic Inquiry and Word Count)

- 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)

# Constructing own lexicon

# Using WordNet to learn polarity

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