

Introduction to Sentiment Analysis

-- *Session 1: Introduction* --

Winter Semester 2019/2020

Instructor: Michael Wiegand

Institute for Computational Linguistics

Heidelberg University, Germany

Outline

- ▶ What is Sentiment Analysis?
- ▶ Administrative Matters

Outline

- ▶ What is Sentiment Analysis?
- ▶ Administrative Matters

What is Sentiment Analysis?

- ▶ Sentiment Analysis deals with the (automatic) detection and analysis of texts that express opinions, evaluations, emotions, and uncertainty.
- ▶ *I think the decision was right.*
- ▶ *This movie is absolutely boring.*
- ▶ *I am frustrated.*
- ▶ *This idea of yours could work.*

Subtasks in Sentiment Analysis

- ▶ What is subjective language?
- ▶ How can we automatically determine the polarity of a text?
- ▶ How can we automatically extract the entities that are involved in an opinion expressed in text?
- ▶ Beyond that, take also into consideration:
 - ▶ Emotions
 - ▶ Uncertainty
 - ▶ ...

Subjectivity Detection

- ▶ Automatically identify texts that convey subjectivity:
 - ▶ *I hate this book.*
 - ▶ *The bus leaves at 9:30h.*
- ▶ Are there particular words or phrases that convey subjectivity?
 - ▶ *awful, okay, superb, expect, thought*
- ▶ Are there certain syntactic clues that convey subjectivity?
 - ▶ *Adjectives are more indicative than other parts of speech.*

Polarity Classification

- ▶ Automatically determine the polarity of texts:
 - ▶ Peter *hates* pancakes .
 - ▶ Mary *loves* pizza.
- ▶ One may also determine the degree of polarity:
 - ▶ *nice* << *great* << ***excellent***

Polarity Classification

- ▶ Need to know which words carry polarity.
 - ▶ *good, clever, fantastic, excellent*
 - ▶ *bad, idiot, horrible, shocking*
- ▶ Need to consider context:
 - ▶ Negation:
 - ▶ *Peter does **not** hate pancakes.*
 - ▶ Irony:
 - ▶ *This is the greatest operating system ever; it crashes every 5 minutes on my laptop!*

Opinion Holder/Target Extraction

- ▶ Holder/source extraction of the opinion:
 - ▶ *In a recent interview, [Noam Chomsky] praised Snowden.*
 - ▶ *This is total nonsense. [speaker]*
- ▶ Target extraction of the opinion:
 - ▶ *In a recent interview, Noam Chomsky praised [Snowden].*
 - ▶ *[This] is total nonsense.*

Other Parameters to Consider

▶ Granularity:

- ▶ document
- ▶ sentence
- ▶ phrase or word

▶ Domain:

- ▶ Train and test on the same domain (*in-domain classification*).
- ▶ Train and test in different domains (*cross-domain classification*).

▶ Language:

- ▶ Train and test on the same language (*monolingual setting*).
- ▶ Train and test on different languages (*cross-lingual setting*).

Why do we want to do Sentiment Analysis?

- ▶ General aim of Natural Language Processing is to find algorithms to compute the meaning of text.
- ▶ Sentiment analysis can be regarded as one building block towards that goal.

Why do we want to do Sentiment Analysis?



Why do we want to do Sentiment Analysis?




One of the most popular applications is automatically classifying sentiment in movie reviews.

Why do we want to do Sentiment Analysis?



Why do we want to do Sentiment Analysis?

A collage of various commercial products is arranged around a central text box. At the top left is a laptop with a blue screen. Next to it is a smartphone displaying a home screen with a clock and weather widget. To the right is a white front-loading washing machine. Below the text box, on the left, is a green soda can with a white logo. On the right is a silver sedan car.

Various companies of commercial products are keen to know the general sentiment of customers towards their products/brand.

Why do we want to do Sentiment Analysis?



Why do we want to do Sentiment Analysis?



Another application is monitoring popular opinion on political issues (e.g. prior to a general election).

Why do we want to do Sentiment Analysis?



BOOKING.COM
online hotel reservations

Why do we want to do Sentiment Analysis?



- ▶ Many hotel reviews are fakes (e.g. praising a bad hotel).
- ▶ Use Sentiment Analysis to filter this ***opinion spam***.

Aims of this Course

- ▶ What is Sentiment Analysis?
- ▶ Administrative Matters

Aims of this Course

- ▶ Learn something about Sentiment Analysis.
- ▶ ... from a linguistic point of view.
- ▶ Learn to read and understand scientific literature.
- ▶ Learn to present a paper.
- ▶ Learn to write a term paper.

Aims of this Course

- ▶ Jump in at the deep end:
 - ▶ No textbook.
 - ▶ We will read research publications.
- ▶ Difference to textbooks:
 - ▶ Condensed language.
 - ▶ Terminology may vary.
 - ▶ Basic research (*task/evaluation may be a bit abstract*).
 - ▶ There may be things you will have to look up yourself.
 - ▶ Establish connections to between different papers yourself.

Requirements for Passing the Course

- ▶ Oral presentation
- ▶ Term paper
- ▶ Reviewing a paper
- ▶ Constant active participation
- ▶ *You are allowed to be absent without excuse for only **one** session!*

Oral Presentation

- ▶ Length: up to 45 minutes (*excluding discussion*).
- ▶ In German or English.
- ▶ Based on one paper (*see course webpage*).
- ▶ You are welcome to include more related papers!
- ▶ We have a discussion after each oral presentation:
 - ▶ You will have to answer questions!
 - ▶ Make sure you really understood the problem discussed in your paper!

Oral Presentation

- ▶ I offer every student an additional (individual) meeting prior to the oral presentation:
 - ▶ Discuss open questions.
 - ▶ Feedback on presentation draft.
- ▶ Arrangement of appointment:
 - ▶ Send in draft at least **8 days prior** to your presentation!
 - ▶ (Time span may be different for the first talk!)

Term Paper

- ▶ Write an essay on the paper you orally presented.
- ▶ Length: 10-15 pages
- ▶ In German or English
- ▶ More details regarding the format will be provided in tomorrow's session.
- ▶ Use **LaTeX!**
- ▶ Deadline for term paper: **TBA**

Reviewing a Paper

- ▶ Review a paper that you do *not* orally present.
- ▶ You are given the main responsibility of the discussion session after the student presentation of that paper:
 - ▶ Ask questions about the paper.
 - ▶ Get involved into the discussion.
 - ▶ Show that you have understood the paper and thought about the problem/task presented.

How is the final mark computed?

- ▶ **Weighting:**
 - ▶ 40% oral presentation
 - ▶ 40% term paper
 - ▶ 20% reviewing
- ▶ If the final “score” is between two grades, course participation will be considered.

Timeline

- ▶ Today:
 - ▶ Recap on machine learning and evaluation
 - ▶ Subjectivity Detection
- ▶ Tomorrow:
 - ▶ How to present a paper
 - ▶ How to write a term paper
- ▶ From 14th October onwards:
 - ▶ One student presentation per week.
 - ▶ Preliminary schedule on webpage.

What to do next?

- ▶ Tell me which paper you want to present.
 - ▶ Give a list of 5 papers sorted by priority
 - ▶ deadline: **4th October (Friday)**
 - ▶ ***Need a volunteer for the first session until tomorrow (end of last session)!***

If you have further questions

- ▶ Ask now!
- ▶ Another good point in time: at the end of each session.
- ▶ ***Don't*** try to discuss complex issues via email!
- ▶ Ask for an appointment instead.