



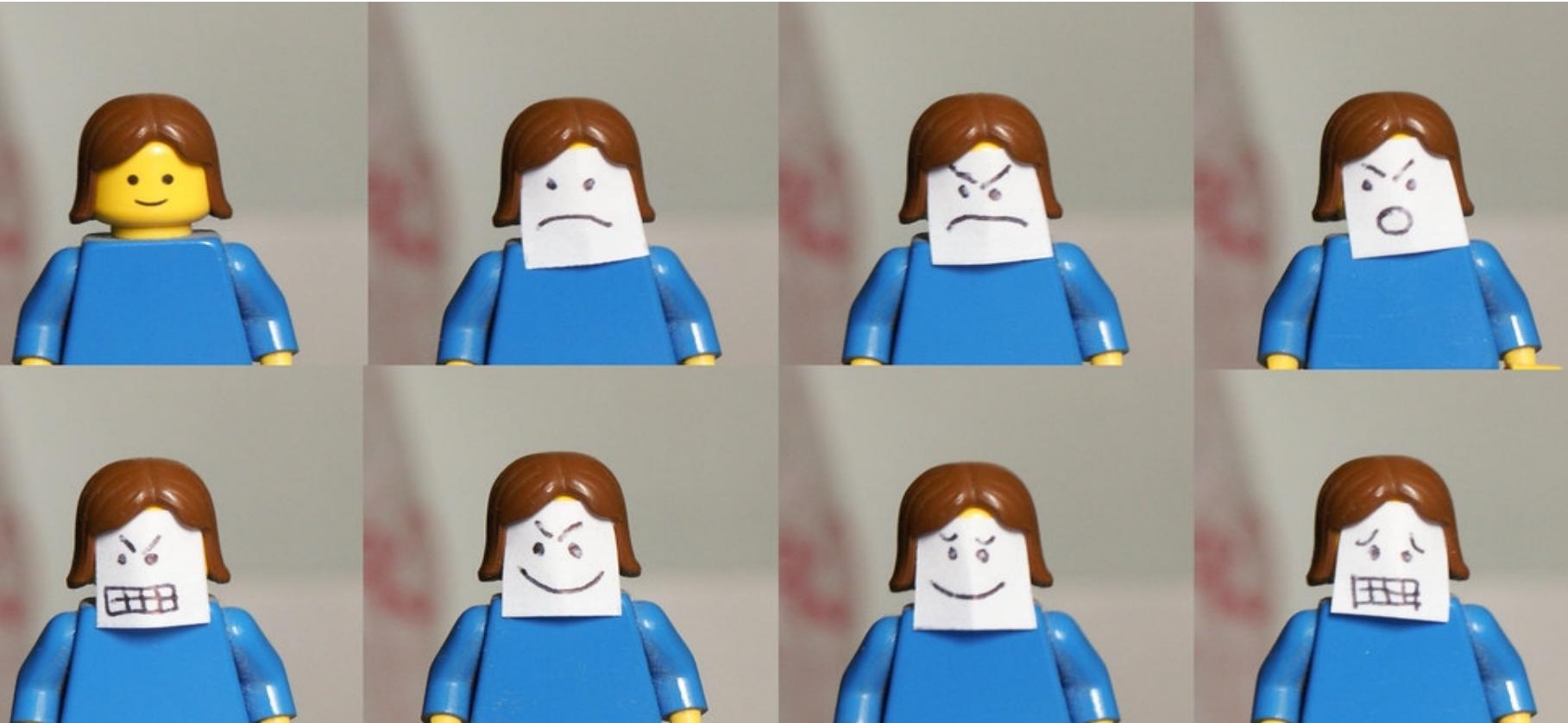
Natural Language Processing

08: Sentiment Analysis

Fabian Haak, Technische Hochschule Köln, Cologne, Germany

Version: 2021-05-31

What People Think about Things: Sentiment Analysis



The „Sentiment“ in Sentiment Analysis

- **affective dimension** of text
- Text contains sentiment as **affective connotation**
- model the lexical semantics relevant to:
 - Sentiment
 - Emotion
 - Personality
 - Mood
 - Attitudes

Why do We Care about Affective Meaning?

- Sentiments are expression of **opinions** and **stances**
- Many relevant scenarios:
 - Sentiment towards politicians, products, countries, ideas, etc.
 - Detecting hate and biased language
 - Understanding product performance
 - ...
- Where could we use this?
 - **Understanding:** Virtual assistants, recommendation systems, „shitstorm“ and hate detection, business decisions
 - **Replicating:** Chatbots, AI

How is Affective Connotation Expressed?

Choice of words:

- „Das ist ja gestern nicht so gut gelaufen.“

Other tokens:

- „Was hast du dir dabei gedacht?!?!?!?!“
- „Das war echt nett! :)“
- „Ich habe die Klausur grade so bestanden __“
→ patterns and structures that carry sentiment in text
- Context dependency:

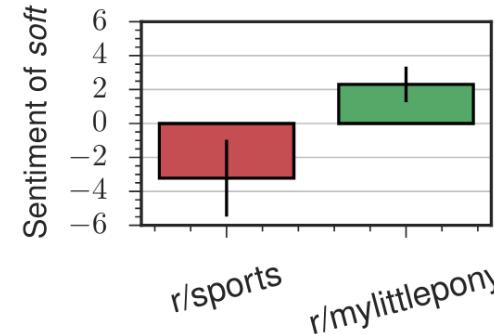


Figure: Hamilton, William & Clark, Kevin & Leskovec, Jure & Jurafsky, Dan. (2016). Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora.

A Definition of Emotion

Emotion (Scherer 2000):

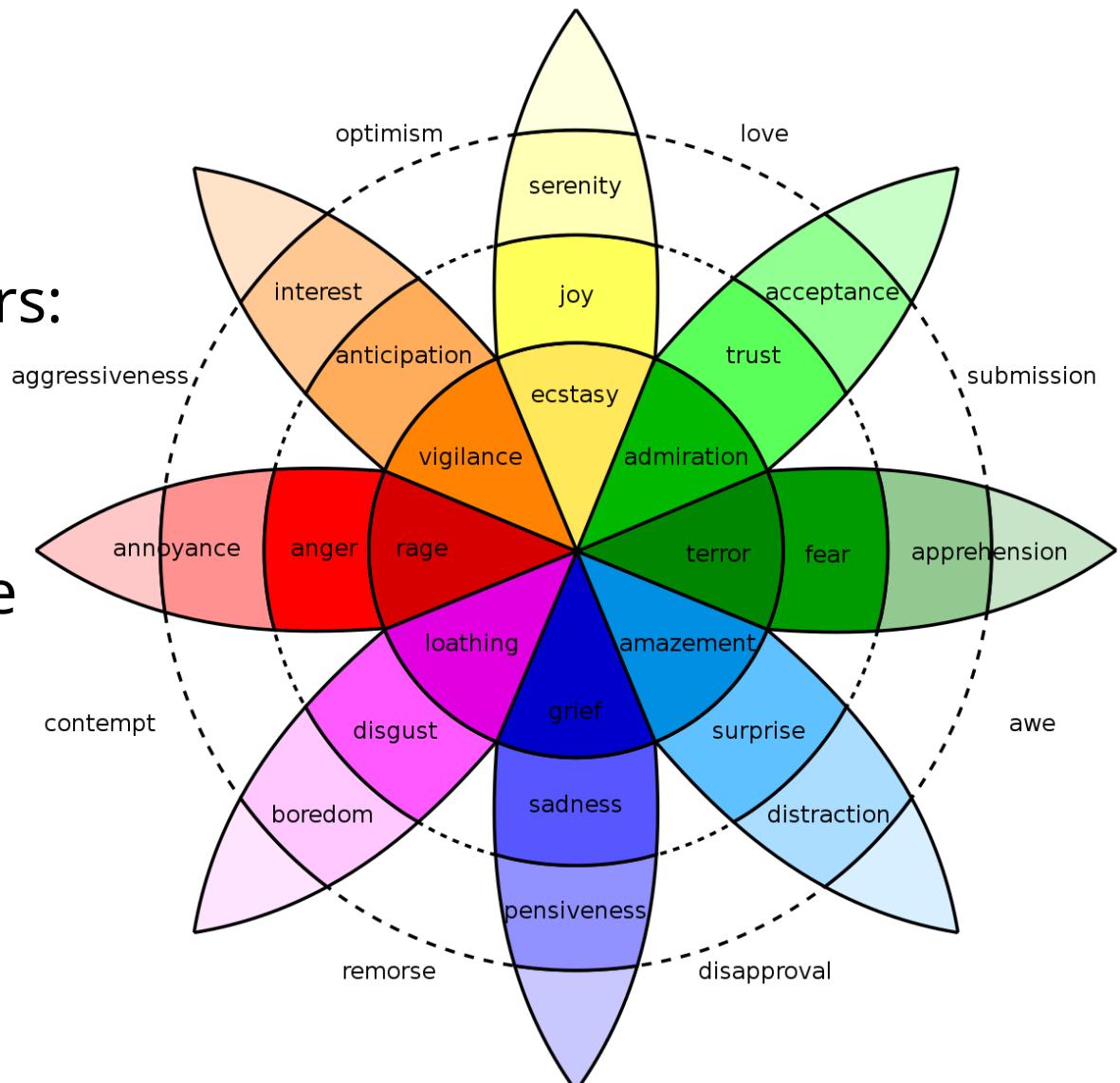
- „relatively brief episode of response to the evaluation of an external or internal event as being of major significance“

Two main concepts of emotion:

- 1) Basic Emotions Theory** (Tomkins '62, Plutchik '62)
- 2) Spatial models** with 2-3 dimensions (f.e. Russel '80)

Basic Emotions Theory: Plutchik's Wheel of Emotions

8 basic emotions
in four opposing pairs:
joy-sadness
anger-fear
trust-disgust
anticipation-surprise



Spatial Models of Emotion

1-dimensional models

- **Polarity**
 - Discrete levels: Positive, Neutral, Negative
 - Good → positive, superb → positive

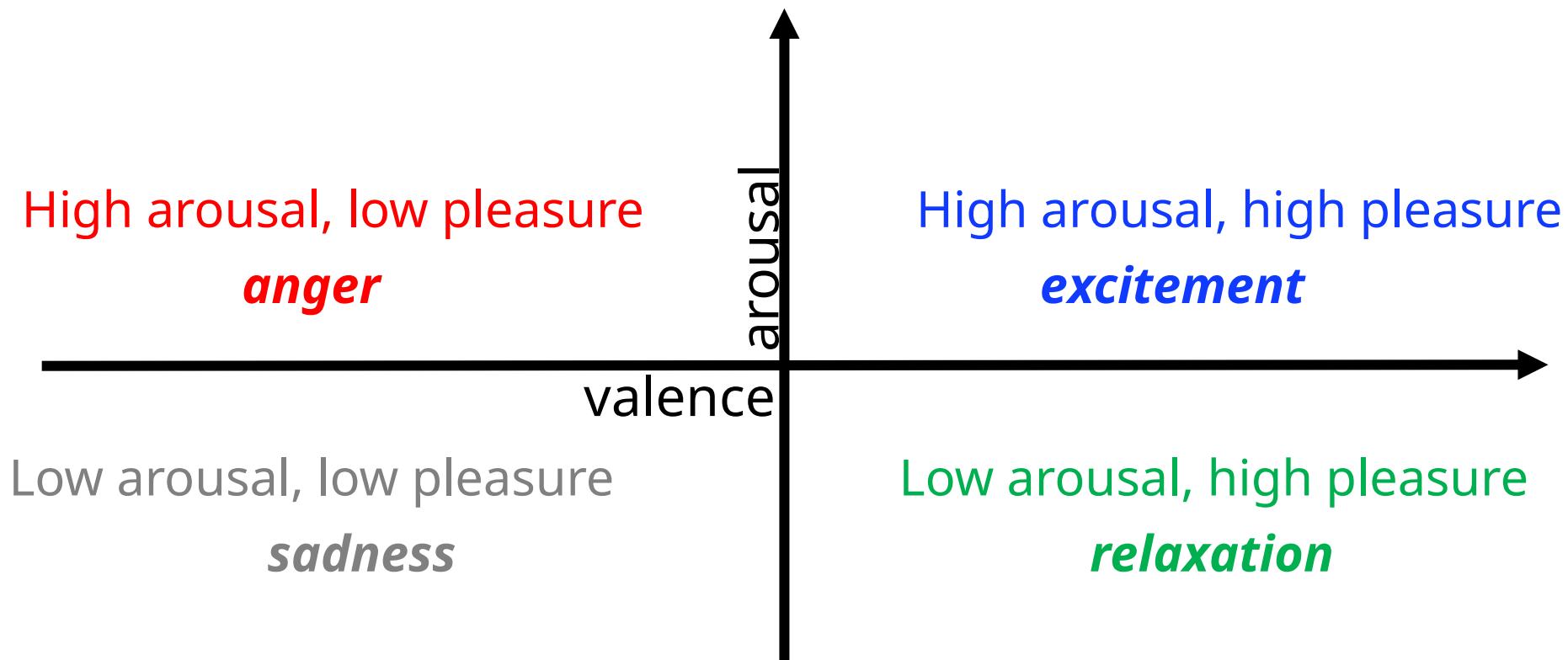
2-dimensinal models

- **Polarity + Intensity (=Valence)**
 - Good → moderately positive, superb → highly positive

Other 2- or 3-dimensional models (Russel '80)

- **Valence** → the pleasantness of the stimulus
- **Arousal** → the intensity of emotion provoked by the stimulus
- **[Dominance** → the degree of control exerted by the stimulus]

Example: Valence/Arousal Dimensions



Different Models for Different Tasks

- Basic Emotions are complex
- Most common sentiment analysis task: **Opinion mining**
 - Does not require complexity of basic emotions models
- Most sentiment analysis approaches utilize spatial models

Granularity Levels of Sentiment Analysis

- **Text-/document-level**

What is the general tone/ stance expressed by the author of a text?

- **Entity-level**

- What attitude/stance/opinion towards an entity is expressed in the document?
 - Example: Opinions on different bikes in a review of bikes.

- **Aspect-level**

- What attitude/stance/opinion towards aspect(s) of an entity is expressed?
 - Example: Opinions on different components of a bike.

- **Sentence-level**

- **Word-/Token-level**

Different
granularity levels
for different tasks

Analysing Expressed Sentiments: Sentiment Lexicons



HATRED or JEALOUSY wrinkles of forehead, if eye brows are sunk down & knit, if eye balls are half hid under eye brows, not turn towards of object, if nostrils are open & more markal than ordinary, & drawn backwad so as to make wrinkles in of cheeks, if mouth is so shut as to shew of teeth are closed, if corners of of mouth are drawn back & very much sunk down, if muscles of of jaw appear sunk.



D E S P A I R is extream, its motions are so likernise, the forehead wrinckes from the top to the bottom, the eye brows bend down over the eyes, & press one another on the sides of the nose, the eye seems to be on fire, & full of blood, the eye balls are distorted, hid under the eye brows, sparkling & unfixed, the eye lids are swelld & livid, the nostrils are large, open, & lifted up, the end of the nose sinks down, the myofls & veins are swelld, the upper part of the cheeks is large, markd & narrow towards the jaw, the mouth drawn backwards is more open at the sides than in the middle, the lower lip large & turned out, they gnash their teeth, they foam, the hair is strait and stands on end.

Sentiment Lexicons as Tools

- Sentiment lexicons (sometimes: dictionaries) contain, for a number of words/tokens (keys), information on emotional dimensions (values)
- Can be used to determine **sentimental connotation**
- Different **forms of sentiment lexicons:**
 - Lists of positive/negative words
 - List of words with polarity and intensity
 - List of positive negative n-grams
 - ...

General Inquirer Lexicons

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

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

General Inquirer (GI) contains different lists of words of sentimental/qualitative dimensions:

- Positive/negative
- Cognitive orientation
- Words for yes/no
- Negations
- Various other category tags

General Inquirer Lexicons

MPQA Subjectivity Cues Lexicon

- a) Theresa Wilson, Janyce Wiebe, and Paul Hoffmann (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. Proc. of HLT-EMNLP-2005.
- b) Riloff and Wiebe (2003). Learning extraction patterns for subjective expressions. EMNLP-2003.

<http://mpqa.cs.pitt.edu/lexicons/>

- List of subjective words with 2 dichotomous dimensions:
 - Positive/negative
 - Strong subjectivity/weak subjectivity
- Word form
- Stemmed verbs
- 6885 words in total

MPQA Subjectivity Cues Lexicon

```
type=weaksubj    len=1 word1=solemn pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj  len=1 word1=solicitous pos1=anypos stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=solicitously pos1=anypos stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=solicitude pos1=noun stemmed1=n priorpolarity=positive
type=weaksubj    len=1 word1=solid pos1=adj stemmed1=n priorpolarity=positive
type=weaksubj    len=1 word1=solid pos1=noun stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=solidarity pos1=adj stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=solidarity pos1=noun stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=soliloquize pos1=verb stemmed1=y priorpolarity=neutral
type=strongsubj  len=1 word1=somber pos1=adj stemmed1=n priorpolarity=negative
type=strongsubj  len=1 word1=soothe pos1=verb stemmed1=y priorpolarity=positive
type=strongsubj  len=1 word1=soothingly pos1=anypos stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=sophisticated pos1=adj stemmed1=n priorpolarity=positive
type=strongsubj  len=1 word1=sore pos1=adj stemmed1=n priorpolarity=negative
type=strongsubj  len=1 word1=sore pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj  len=1 word1=sorely pos1=anypos stemmed1=n priorpolarity=negative
type=weaksubj    len=1 word1=sorenness pos1=noun stemmed1=n priorpolarity=negative
type=strongsubj  len=1 word1=sorrow pos1=noun stemmed1=n priorpolarity=negative
```

NRC Emotion Intensity Lexicon (NRC-EIL) and NRC Valence, Arousal, and Dominance (NRC-VAD) Lexicon

Saif M. Mohammad. In Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018), May 2018, Miyazaki, Japan.

<https://saifmohammad.com/WebPages/AffectIntensity.htm>

- Both available in 100 languages (Google translated!)
- NRC-EIL:
 - 8 Basic emotions (Tomkins)
 - 1 to 8 emotion scores per word
- NRC-VAD:
- Valence, Arousal, and Dominance scores for words

The NRC Emotion Intensity Lexicon (NRC-EIL)

hatred	Hass	anger	0.953
hateful	hasserfüllt	anger	0.940
terrorize	terrorisieren	anger	0.939
Violently	heftig	anger	0.938
infuriated	wütend	anger	0.938
furious	wütend	anger	0.929
enraged	wütend	anger	0.927
furiously	wütend	anger	0.927
screwyou	NO TRANSLATION	anger	0.924
murderer	Mörder	anger	0.922
fury	Wut	anger	0.922
execution	Ausführung	anger	0.917
angered	verärgert	tanger	0.916
savagery	Wildheit	anger	0.915
slaughtering	Schlachtung	anger	0.914
Veryangry	NO TRANSLATION	anger	0.913
fuckoff	Verpiss dich	anger	0.912
annihilation	Vernichtung	anger	0.912
assassinate	ermorden	anger	0.912

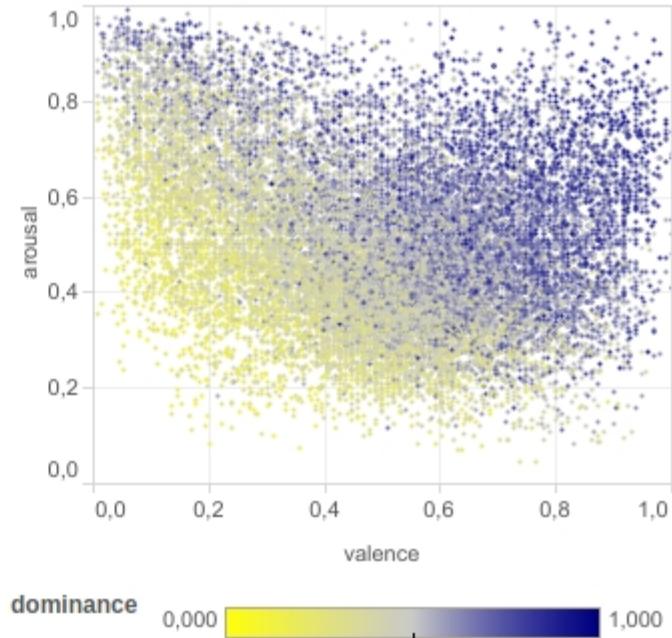
NRC Valence, Arousal, and Dominance (NRC-VAD) Lexicon

https://public.tableau.com/views/TheNRCValenceArousalandDominanceLexiconViz/Dashboard1?:embed=y&:embed_code_version=3&:loadOrderID=0&:display_count=y&publish=yes&:origin=viz_share_link

The NRC VAD Lexicon

term	z	valence	arousal	dominance
aaaaaaah		0,479	0,606	0,291
aaaah		0,520	0,636	0,282
aardvark		0,427	0,490	0,437
aback		0,385	0,407	0,288
abacus		0,510	0,276	0,485
abalone		0,500	0,480	0,412
abandon		0,052	0,519	0,245
abandoned		0,046	0,481	0,130
abandonment		0,128	0,430	0,202
abashed		0,177	0,644	0,307
abate		0,255	0,696	0,604
abatement		0,388	0,338	0,336
abba		0,562	0,500	0,480
abbey		0,580	0,367	0,444
abbot		0,427	0,321	0,483
abundance		0,521	0,375	0,320

valence vs. arousal



Domain Specific Language

- Different domains have different languages
 - Use of emojis, technical terms, slang,
- Can be used to determine sentimental connotation on a granularity level of choice
- Different forms of sentiment lexicons:
 - Lists of positive/negative words
 - List of words with polarity and intensity
 - List of positive negative n-grams
 - ...

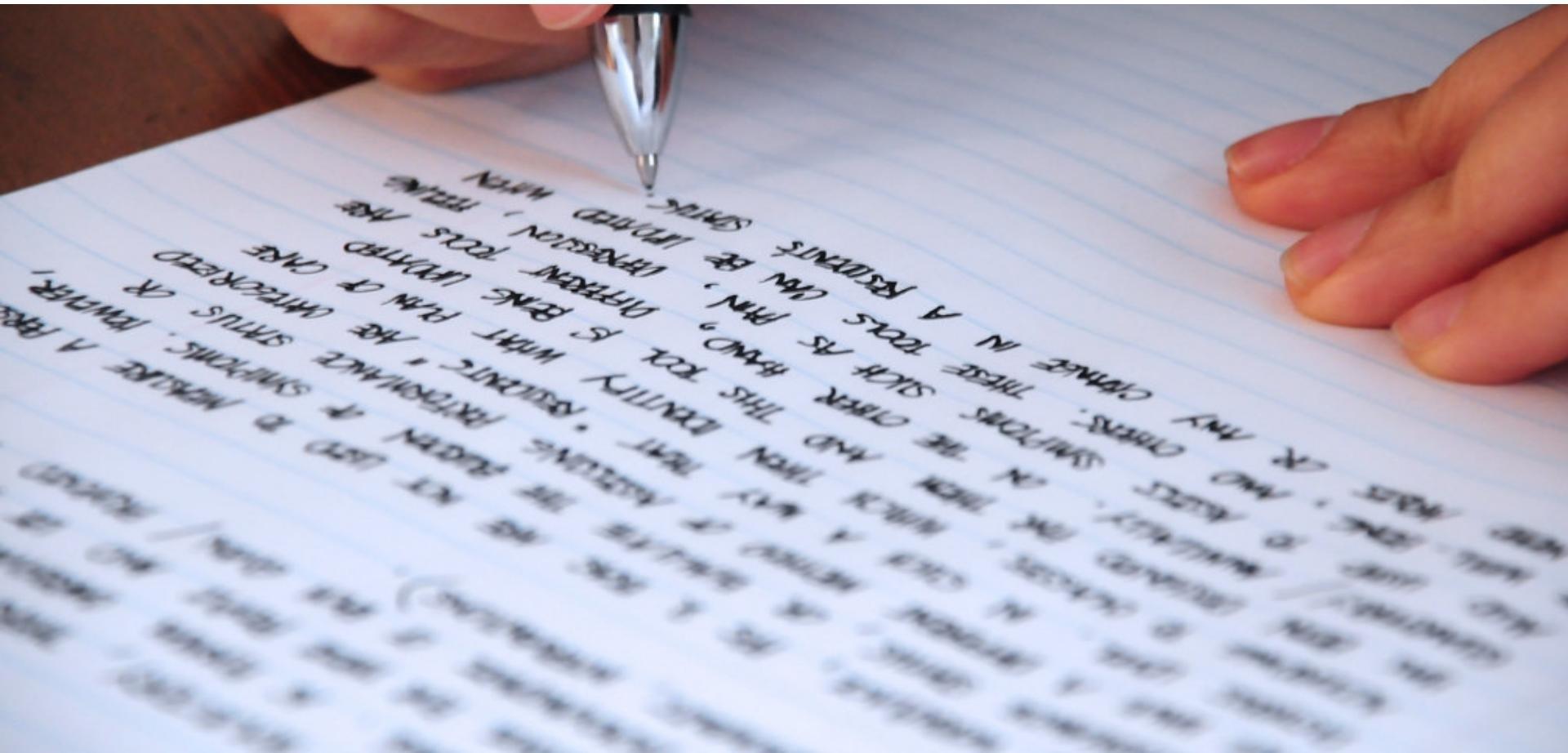
VADER Sentiment Lexicon

Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

<https://github.com/cjhutto/vaderSentiment>

- „specifically attuned to **sentiments expressed in social media**“
- Negations, contractions, punctuation, word-shape (insane vs INSANE), and modifier aware
- Emoticons and utf-8 encoded emojis
- sentiment-laden slang words as words (e.g., 'sux') or as modifiers such as 'uber' or 'friggin' or 'kinda'
- sentiment-laden initialisms and acronyms (for example: 'lol')
- **Included in NLTK**

Building Lexicons: Human Labels



Method: Human Annotation

- Recruit „novice“ crowdworkers or experts for annotation
- Annotators label candidate words or random words
- Example: NRC Valence, Arousal, Dominance (VAD) lexicon, dimension valence:
 - Q1. Which of the four words below is associated with the MOST happiness / pleasure / positiveness / satisfaction / contentedness / hopefulness OR LEAST unhappiness / annoyance / negativeness / dissatisfaction / melancholy / despair?
 - vacation, consolation, whistle, torture
- → different heuristics to derive lists

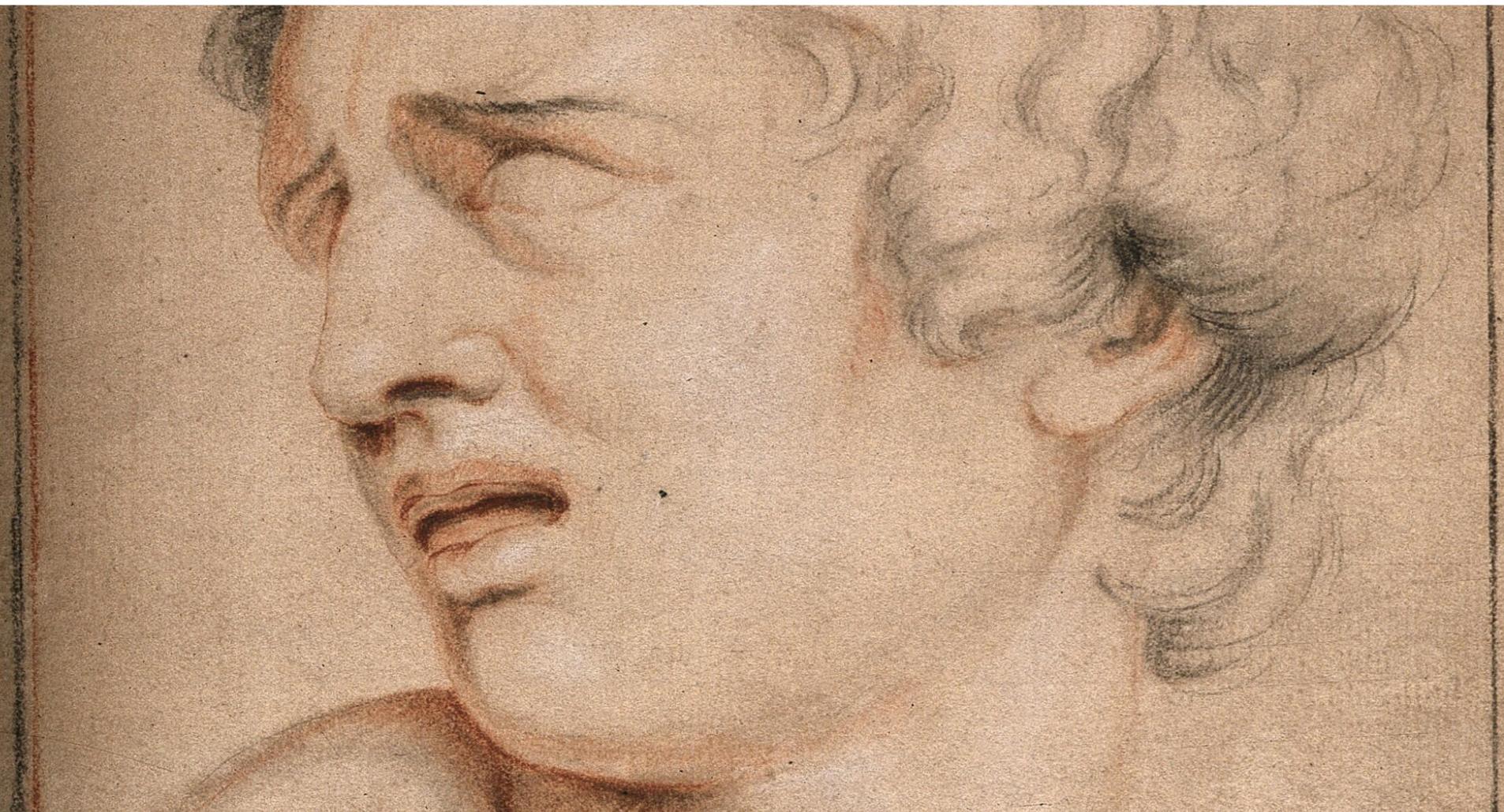
Issues with Human Labeled Lexicons

- Native speakers
- Clear task description
- Susceptible for bias
 - Cultural bias
 - Personal bias
 - Questionnaire bias
 - Response bias
- Different levels of measurement
 - Dichotomous
 - Ordinal (Likert)



Manual annotation is
complicated and
expensive!

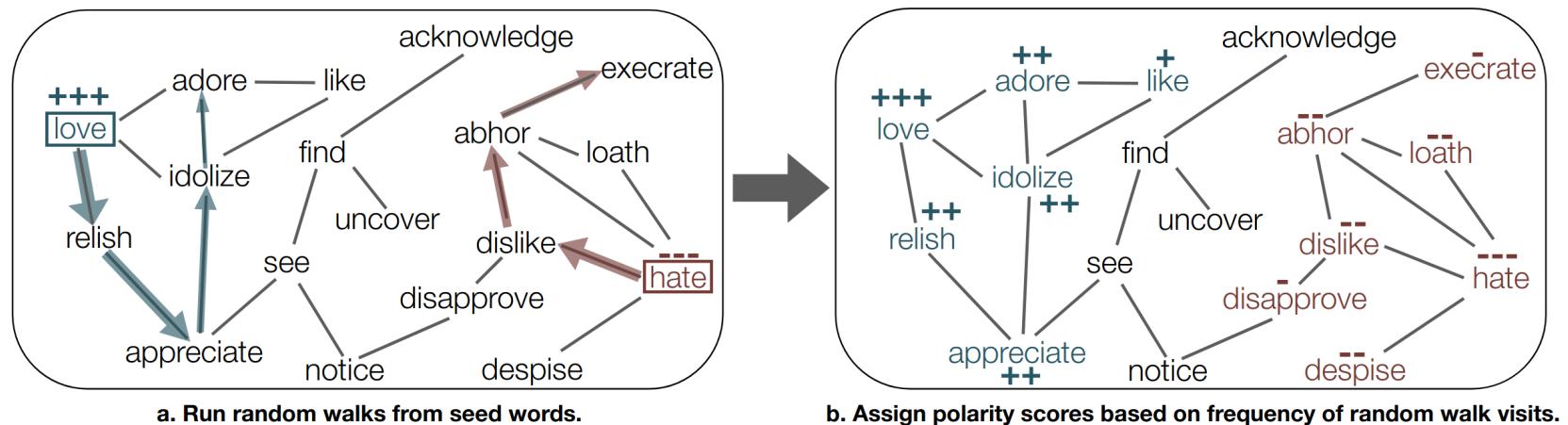
Building Lexicons: Automated Induction of Affect Lexicons



Head of a man expressing acute pain. Pencil drawing after Charles Le Brun. Wellcome Collection. Public Domain Mark

Semi-supervised Approaches

- Main concept: Generate lexicons based on similarity to positive and negative seedwords



Semi-supervised: Semantic Axis

- Main concept: Generate lexicons based on similarity to positive and negative seedwords

- 1) Generate **list of seed words****
- 2) Generate **word embeddings** (f.e. Word2Vec)**
- 3) Optional: fine tune embeddings with domain-specific corpus**
- 4) Calculate pole centroids and semantic axis:**

$$\mathbf{V}^+ = \frac{1}{n} \sum_1^n E(w_i^+) \quad \mathbf{V}^- = \frac{1}{m} \sum_1^m E(w_i^-) \quad \mathbf{V}_{axis} = \mathbf{V}^+ - \mathbf{V}^-$$

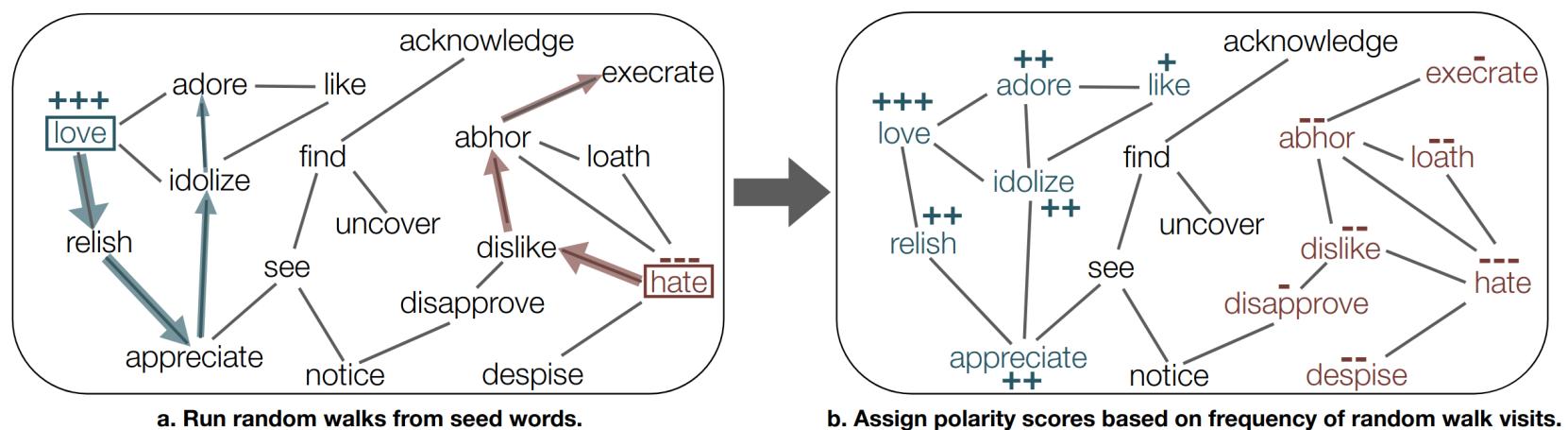
- 5) For each candidate word, calculate distance to poles:**

$$\text{score}(w) = (\cos(E(w), \mathbf{V}_{axis}))$$

$$= \frac{E(w) \cdot \mathbf{V}_{axis}}{\|E(w)\| \|\mathbf{V}_{axis}\|}$$

Semi-supervised: Label Propagation

- Main concept: Generate **domain specific** lexicons based on **random walk** through **graph** of embeddings, starting from **seed words**
- Travel to probability derived from similarity
- Assign confidence based on “visits in N travels”



Semi-supervised Approaches: Problems

- Different domains have different positive and negative words
- Approaches:
 - Full range of general seed words, trust system
 - Domain-specific embeddings
 - Domain-specific selection seed words:

Domain	Positive seeds	Negative seeds
General	good, lovely, excellent, fortunate, pleasant, delightful, perfect, loved, love, happy	bad, horrible, poor, unfortunate, unpleasant, disgusting, evil, hated, hate, unhappy
Twitter	love, loved, loves, awesome, nice, amazing, best, fantastic, correct, happy	hate, hated, hates, terrible, nasty, awful, worst, horrible, wrong, sad
Finance	successful, excellent, profit, beneficial, improving, improved, success, gains, positive	negligent, loss, volatile, wrong, losses, damages, bad, litigation, failure, down, negative

Other Semi-supervised Approaches

- Based on other similarity measures:
 - Morphological similarity (f.e. common stems)
 - Synonyms and antonyms (WordNet)

- → Sentiwordnet:

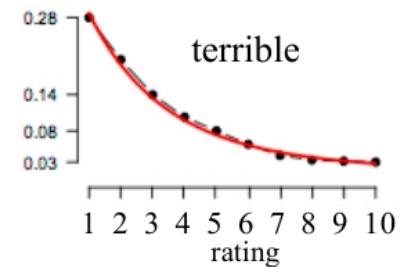
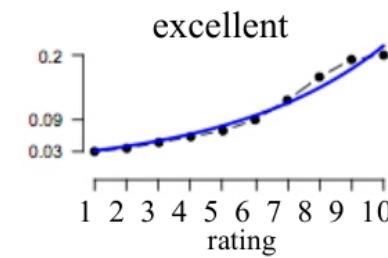
Synset		Pos	Neg	Obj
good#6	‘agreeable or pleasing’	1	0	0
respectable#2 honorable#4	good#4 estimable#2	0.75	0	0.25
estimable#3 computable#1	‘may be computed or estimated’	0	0	1
sting#1 burn#4 bite#2	‘cause a sharp or stinging pain’	0	0.875	.125
acute#6	‘of critical importance and consequence’	0.625	0.125	.250
acute#4	‘of an angle; less than 90 degrees’	0	0	1
acute#1	‘having or experiencing a rapid onset and short but severe course’	0	0.5	0.5

Supervised Learning of Word Sentiment

- Approach: Derive positive and negative words from corpus of labelled texts
- Corpora: Reviews of movies (IMDb), products (Amazon), places (Yelp, Google), ...
- Derive rating (Potts Score):

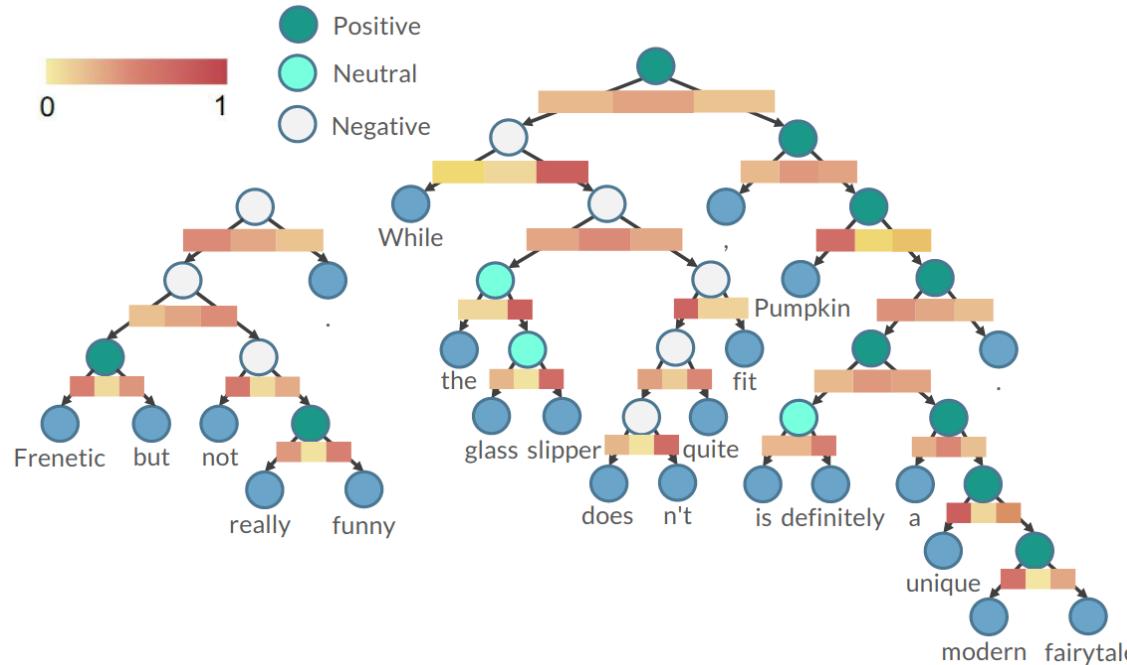
$$P(w|c) = \frac{\text{count}(w, c)}{\sum_{w \in C} \text{count}(w, c)}$$

$$\text{PottsScore}(w) = \frac{P(w|c)}{\sum_c P(w|c)}$$



Other Supervised Machine Learning Approaches

- Recently very successful: **Language model** based sentiment classification
- Learn sentiment connotations from training data
- For example: Modified BERT: SentiBERT



Using Lexicons for Sentiment Recognition



Two heads, both in states of fear. Engraving after C. Le Brun. Wellcome Collection. Public Domain Mark

Using Sentiment Analysis: Sentences and Texts

- 1) Ratio of positive to negative words
- 2) Sum of sentiment scores of all words
- 3) Rule-based approaches (e.g. VADER)
 - Includes negations, punctuation, emojis etc.

Extra: Building a Platform-specific Emoji Sentiment Lexicon



Plattform: Twitch.tv

www.twitch.tv

<https://www.twitch.tv/creatorcamp/de-de/learn-the-basics/emotes/>

- (Gaming) streaming platform
- Platform specific emojis: Twitch Emotes

:-)



Concept of the Approach

- Multimodal emojis carry **nonverbal functions** of non-formal chat communication (whatsapp, twitch, etc.)
 - Via iconographic properties
- Problem: Platforms with **high number of custom emojis** (> 300.000 twitch emotes)
- Utf-8 emojis:
 - Often not supported
 - Most sentiment scores older than 5 years
- Gold Standard: Emoji Sentiment Ranking ESR
[http://kt.ijs.si/data/Emoji sentiment ranking/](http://kt.ijs.si/data/Emoji_sentiment_ranking/)

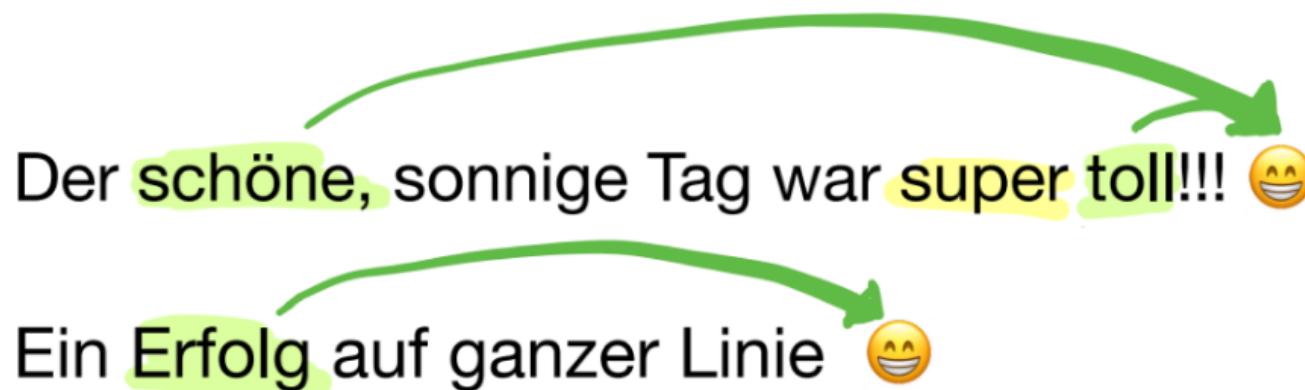
Goal: Semi-supervised creation of domain specific emoji lexicons

Solution: Cooccurrence with Sentiment Words

- Emojis have 3 main functions:

- 1) Express sentiment
- 2) Amplify sentiment
- 3) Reverse sentiment

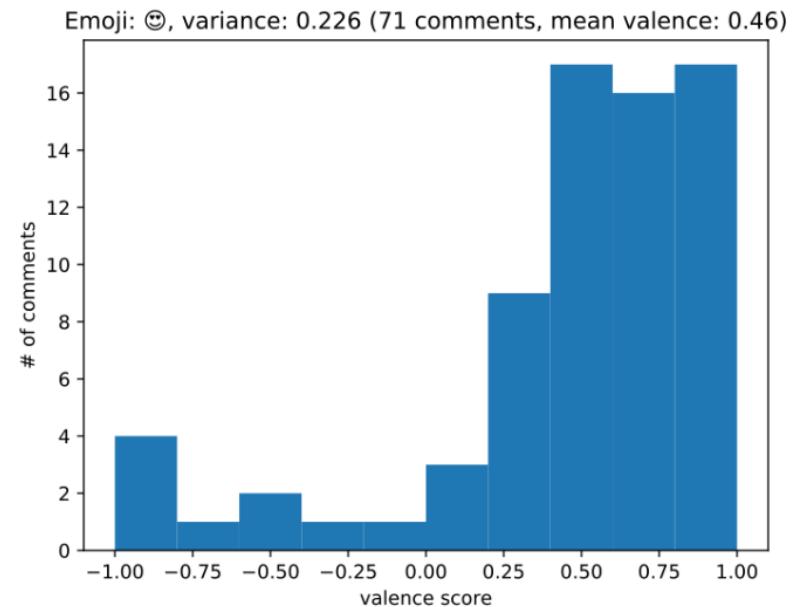
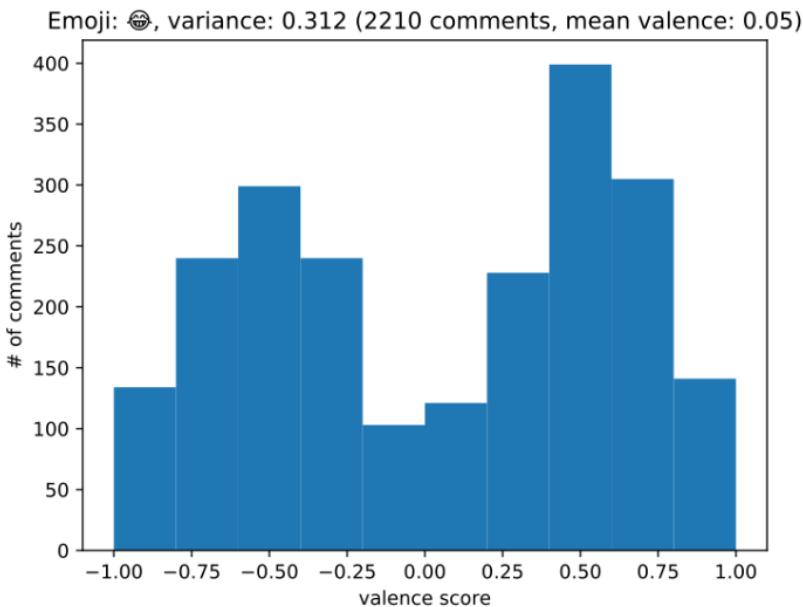
→ Derive emoji sentiment by assuming expression function: **Occurrence in chat messages with positive VADER score**



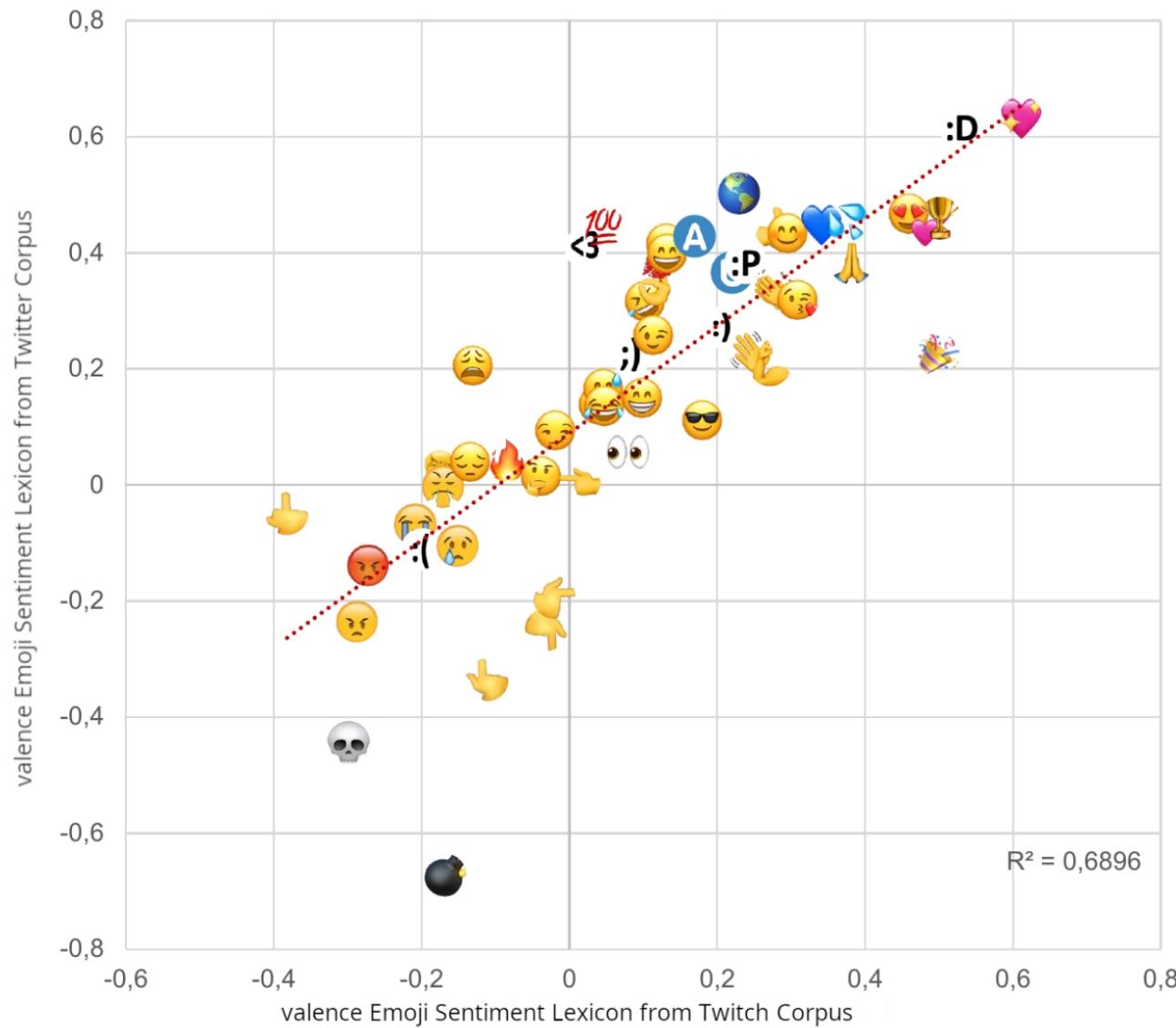
Development of an Emoji Sentiment Lexicon

- Corpus 1 (utf-8 emojis): 1.5m tweets
- Corpus 2 (twitch emotes): 2.8m chat messages
-
- Edited VADER Lexicon: Emojis and emoticons removed
- Derive emoji sentiment scores by **mean sentiment value** of all texts containing the emoji
- Standard deviation and minimum cases for validity measure
- Two modes → expression of irony or sarcasm

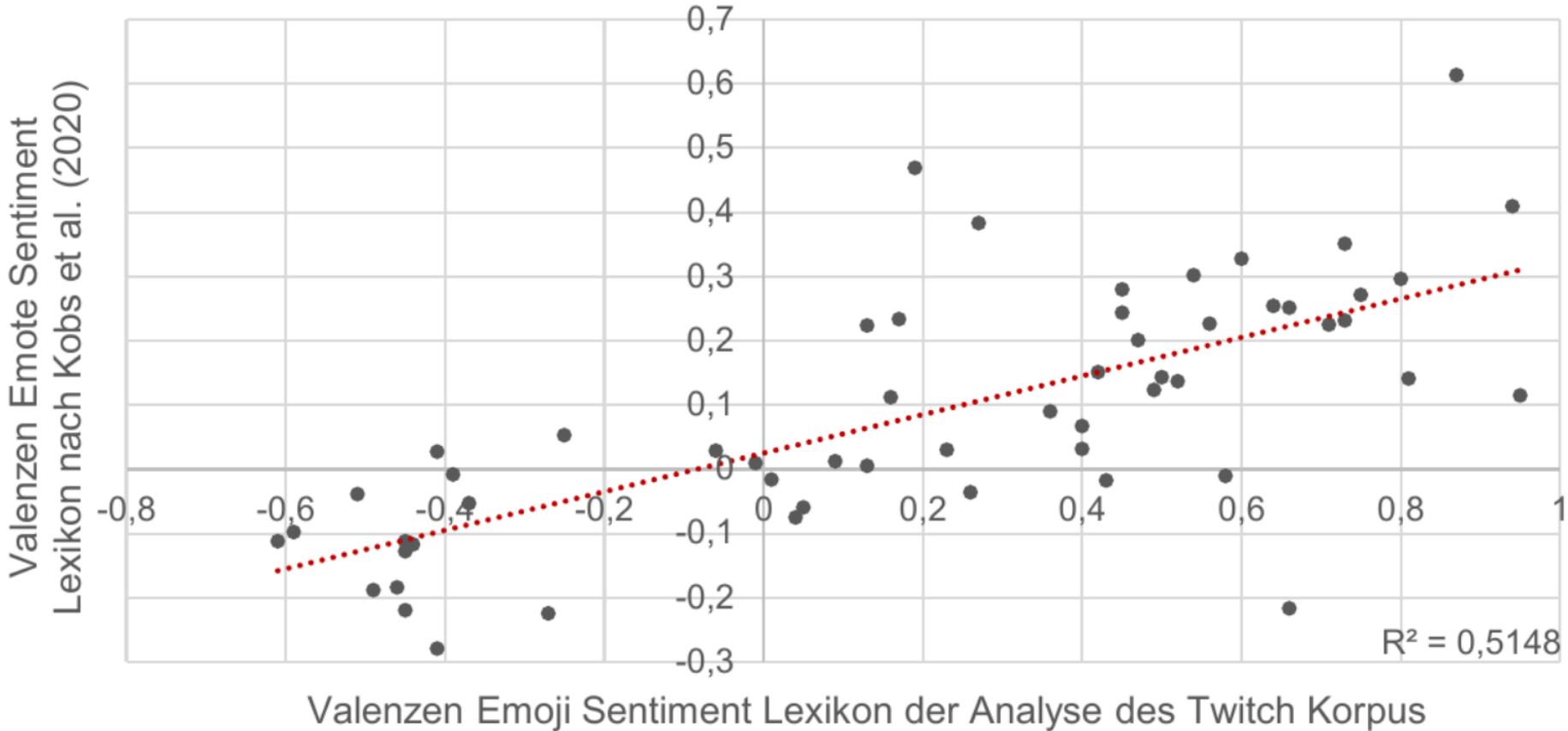
Distribution of Valence Scores



Results: UTF-8 Emojis



Results: Twitch Emotes



Conclusion

- Automatically derived sentiment scores for close to 200 twitch emotes and 200 utf-8 emojis
- The results have significantly (<0.01) high correlation with human labels
- Use with VADER: Doubles amount of text with sentiment in twitch corpus
- Possible follow-up studies:
 - 1) Bigger corpus, bigger lexicon: “ESR 2.0”
 - 2) Study on effects in sentiment analysis
 - 3) Domain specific lexica with domain specific corpora