## **TEXT MINING**

#### LO9. SENTIMENT ANALYSIS AND STANCE DETECTION

SUZAN VERBERNE 2022



### **TODAY'S LECTURE**

- Quiz about week 8
- Sentiment analysis
  - Concepts of sentiment analysis
  - Labels and ordinal regression
  - Aspect-based sentiment analysis
  - Challenges
  - Evaluation
- Stance detection (by Myrthe Reuver)



- How is a topic defined in LDA?
  - a. As a cluster of words from the collection
  - b. As a cluster of documents from the collection
  - c. As a probability distribution over the collection's vocabulary
  - d. As a probability distribution over the collection's documents



- How is a topic defined in LDA?
  - a. As a cluster of words from the collection
  - b. As a cluster of documents from the collection
  - c. As a probability distribution over the collection's vocabulary
  - d. As a probability distribution over the collection's documents



> What is the role of the dirichlet prior in LDA?

- a. To initialize a random uniform distribution
- b. To initialize a random sparse distribution
- c. To optimize the number of topics in the collection
- d. To optimize the number of terms per topic



> What is the role of the dirichlet prior in LDA?

- a. To initialize a random uniform distribution
- b. To initialize a random sparse distribution
- c. To optimize the number of topics in the collection
- d. To optimize the number of terms per topic



- What is the difference between extractive and abstractive summarization?
  - a. Extractive summarization methods are feature-based and abstractive summarization methods are neural-network-based
  - Extractive summarization methods compose the summary from nuggets in the original text and abstractive summarization methods generate a new text
  - c. Extractive summarization methods summarize a whole document and abstractive summarization methods summarize a single sentence



- What is the difference between extractive and abstractive summarization?
  - a. Extractive summarization methods are feature-based and abstractive summarization methods are neural-network-based
  - b. Extractive summarization methods compose the summary from nuggets in the original text and abstractive summarization methods generate a new text
  - c. Extractive summarization methods summarize a whole document and abstractive summarization methods summarize a single sentence



- Why is the LEAD-3 baseline difficult to beat in commonly used benchmark sets?
  - a. Benchmark sets often consist of newspaper texts and in newspaper articles, the most important information is in the beginning of the text
  - b. Because the training data is not sufficiently diverse
  - c. Because the sentences with the highest centrality are the most representative of the text



- Why is the LEAD-3 baseline difficult to beat in commonly used benchmark sets?
  - a. Benchmark sets often consist of newspaper texts and in newspaper articles, the most important information is in the beginning of the text
  - b. Because the training data is not sufficiently diverse
  - c. Because the sentences with the highest centrality are the most representative of the text



- What is the common metric for evaluating automatic summarization? How is it defined?
  - a. BLEU: the precision of n-grams in the automatically generated summary compared to the reference summary
  - b. BLEU: the recall of n-grams in the automatically generated summary compared to the reference summary
  - c. ROUGE: the precision of n-grams in the automatically generated summary compared to the reference summary
  - d. ROUGE: the recall of n-grams in the automatically generated summary compared to the reference summary



- What is the common metric for evaluating automatic summarization? How is it defined?
  - a. BLEU: the precision of n-grams in the automatically generated summary compared to the reference summary
  - b. BLEU: the recall of n-grams in the automatically generated summary compared to the reference summary
  - c. ROUGE: the precision of n-grams in the automatically generated summary compared to the reference summary
  - d. ROUGE: the recall of n-grams in the automatically generated summary compared to the reference summary



#### Which of these statements is true?

- a. Extractive summarization leads to more fluent summaries than abstractive summarization
- b. Abstractive summarization can be modelled as a translation task
- c. Extractive summarization is always a supervised learning task



#### > Which of these statements is true?

- a. Extractive summarization leads to more fluent summaries than abstractive summarization
- b. Abstractive summarization can be modelled as a translation task
- c. Extractive summarization is always a supervised learning task



# **SENTIMENT ANALYSIS**



Suzan Verberne 2022

#### **SENTIMENT ANALYSIS AND OPINION MINING**

- Booming research area since 2000s
- Connected to the rise of social media: the place for expressing personal and subjective opinions
- Lists of different tasks in applications: opinion mining, sentiment classification, sentiment mining, subjectivity analysis, affect analysis, emotion detection, stance detection, ...
- The term 'sentiment analysis' is over-used. Always ask the client what they really want



#### **SENTIMENT RELATED TASKS**

#### Examples:

- Aggregate positive and/or negative opinions from customer reviews, e.g. products, accommodations, movies
- Compare aspects of products using customer reviews
- Analyse the sentiment on social media related to political issues
- Analyse the sentiment on social media related to companies or products
- Detect how opinions change over time (e.g. towards companies or political topics)



#### **LEVELS OF SENTIMENT ANALYSIS**

Level of sentiment analysis	Type of task
<b>Document</b> level: a sentiment score for the complete text (e.g. product review or Tweet)	Classification
Sentence level: a sentiment score per sentence (e.g. because a review is longer and may address multiple aspects)	Classification
Entity and Aspect level: relate the sentiment to features of a product, event or entity	Extraction and classification



# **SENTIMENT LABELS**



Suzan Verberne 2022

#### SENTIMENT CLASSIFICATION TASKS

Sentiment classes: negative, positive, neutral

- Alternatives:
- Objective versus Subjective
- Emotion classes: joy, fear, anger, sadness, disgust, shame, and guilt
- Stance: pro, con, neutral (or favor, against, neutral)



#### **ORDINAL LABELS**

Sentiment classes: negative, positive, neutral

Common for sentiment: ordinal scales (e.g. product reviews)

very bad	bad	neutral	good	very good
-2	-1	0	+1	+2
1	2	3	4	5

Ordinal variable: variable with values that are categorical but have an order (e.g. very bad, bad, neutral, good, very good)



#### **EXAMPLE ORDINAL SENTIMENT**

Sentence	Sentiment (-2,+2)
today on the advice of Paulien with me for relaxation therapy because of anxiety and fatigue	-2
It goes pretty well.	+1
There is a decrease of the complaints with respect to the previous treatment.	+1
it goes reasonably well, still notices muscle pain in the thigh and now considerable bruising by the knee	+1
Pain in the right ankle since Friday (outside), painful today with cycling and walking. No trauma moment.	-1
goes well, no details. Wednesday to doctor	+1
Going has a small tip laterally thigh which is a bit painful / nagging	-2



Suzan Verberne 2022

#### **ORDINAL LABELS**

- Misclassifying an entity with true label '-2' as '-1' is a smaller mistake than labeling an entity with true label '-2' as '+2'
- Ordinal regression: learn a model to predict class labels on an ordinal scale
  - Variant of regression for ordinal variables
  - A problem between regression and classification ("ordinal classification")



#### **ORDINAL REGRESSION**

$$P(y \le j | \theta_j, w, X) = \frac{1}{1 + e^{-(\theta_j - Xw)}}$$

sigmoid function (like in logistic regression)

y : target variable

 $\geq \theta_j$ : threshold for class j

- > X : input instances
- $\succ$  w : weights to be learned





https://towardsdatascience.com/deep-ordinal-logistic-regression-1afd0645e591

# ASPECT-BASED SENTIMENT ANALYSIS



Suzan Verberne 2022

#### **ANNOTATION ON THE ENTITY/ASPECT LEVEL**



\*\*\* Reviewed in the United Kingdom on 10 November 2021 Colour Name: Space Grey | Size Name: 64GB | Style Name: WiFi | Verified Purchase

It's an IPad, not sure what else to say. Reasonably priced given other options has all the usual apple functions and quality. Quick response and good battery life and hopefully they won't stop updates on this one too soon. Quick and easy delivery as always.

#### Assumption:

Review texts contain sentiment or opinions that someone expresses at some place and time about aspects of entities



#### **ANNOTATION ON THE ENTITY/ASPECT LEVEL**

Aspect-based sentiment analysis: find Quintuple (E, A, S, H, C)

- E: opinion target (entity, event or topic)
- A: aspect or feature of E
- S: sentiment/opinion content (sentiment score of A)
- H: opinion holder
- C: context; time and location of the expression



#### **EXAMPLE OF A PRODUCT REVIEW**



#### \*\*\*

Reviewed in the United Kingdom on 10 November 2021

Colour Name: Space Grey | Size Name: 64GB | Style Name: WiFi | Verified Purchase

It's an IPad, not sure what else to say. Reasonably priced given other options has all the usual apple functions and quality. Quick response and good battery life and hopefully they won't stop updates on this one too soon. Quick and easy delivery as always.

#### > (E, A, S, H, C)



Suzan Verberne 2022

#### **EXAMPLE OF A PRODUCT REVIEW**



t is what it is

Reviewed in the United Kingdom on 10 November 2021

Colour Name: Space Grey | Size Name: 64GB | Style Name: WiFi | Verified Purchase

It's an IPad, not sure what else to say. Reasonably priced given other options has all the usual apple functions and quality. Quick response and good battery life and hopefully they won't stop updates on this one too soon. Quick and easy delivery as always.

- (E, A, S, H, C)
- (ipad, price, reasonable, girlsandshoes, (UK, 10 November 2021))
- (ipad, battery life, good, girlsandshoes, (UK, 10 November 2021))



#### H: opinion holder

- In tweets and reviews: usually the author
  - username, e.g. *girlsandshoes*
- In news: needs to be extracted from the text
  - Named Entity Recognition needed
  - "Look what other countries are doing now about adopting a booster campaign virtually for everybody. I think if we do that ... by the spring we can have pretty good control of this," Fauci said.



- C: context: time and location of the expression
  - In tweets and reviews: date/location stamp
    - E.g. United Kingdom, 10 November 2021
  - In news: needs to be extracted from the text:
    - Time expression Recognition or/and Geolocation classication needed
    - On funding for adaptation, the UN secretary general, António Guterres, said on Thursday during the UN summit in Egypt: "We need a global surge in adaptation investment to save millions of lives from climate carnage."



#### E: entity, event or topic

- In reviews: given by the metadata (product, movie, book, hotel)
  - E.g. 2021 Apple iPad (10.2-inch iPad, Wi-Fi, 64GB) Space Grey (9th Generation)
- In news or tweets: needs to be extracted from the text:
- Named Entity Recognition / Event Detection
- Say about the #fairphone what you like, the fact that I can easily replace parts with only a screwdriver is a great.
- day one of leids ontzet: currently waiting in the rain to go to this music thing i dont know anything about. still on 0 drinks.



A: aspect or feature of the entity that the sentiment is about

- Information Extraction and aspect categorization needed
- Aspects are domain dependent and even product dependent

Location	4.8	Battery life	8,0	Power	<b>\$ \$ \$ \$</b> \$ \$ 6,0
Cleanliness	4.5	Wearing comfort	8,0	Drilling performance	8,0
Service	4.4	Design	8,0	Build guality	00000 8,0
Value	4.0	Options	00000 8.0		

- Can be challenging to extract
  - e.g. the room was dirty → aspect 'cleanliness'



#### **KNOWLEDGE BASES FOR SENTIMENT**

- It helps to have a product database. Why?
  - 1. To know which products exist (someone might mention a different product in the review)
  - 2. To know which aspects a given product type has (a drill does not have cleanliness as relevant aspect)
  - > This facilitates aspect extraction (know what to look for in the text)



#### S: sentiment of A

- We are looking for sentiments for a given aspect
- Once the entity (product of event) is known (e.g. iphone)
- > and the aspect is known (e.g. *battery life*)
  - we can classify the sentiment of the sentence(s) describing the aspect
    - e.g. the battery runs down very fast. disappointing



# **CHALLENGES**



Suzan Verberne 2022

#### EXAMPLE

"This past Saturday, I bought a Samsung phone and my girlfriend bought a HTC phone with 3D touch. We called each other when we got home. The voice on my phone was not so clear, worse than my previous phone. The battery life was long. My girlfriend was quite happy with her phone. I wanted a phone with good sound quality. So my purchase was a real disappointment. I returned the phone yesterday."

Which challenges for sentiment analysis do you identify in this review text?



#### EXAMPLE

This past Saturday, I bought a Samsung phone and my girlfriend bought a HTC phone with 3D touch. We called each other when we got home. The voice on my phone was not so clear, worse than my previous phone. The battery life was long. My girlfriend was quite happy with her phone. I wanted a phone with good sound quality. So my purchase was a real disappointment. I returned the phone yesterday."

- comparison: 2 opinion holders, 2 targets
- different aspects with different sentiments
- co-references
- factual statement expressing sentiment
- sentiment words not expressing sentiment



#### **CHALLENGES OF SENTIMENT ANALYSIS**

- 1. Sentiment words do not always express a sentiment
  - Detecting an entity+sentiment is not enough:

"Can you tell me which Sony camera is good?"

"If I can find a good camera in the shop, I will buy it."

However, excluding questions and conditionals would be too easy:

"Does anyone know how to repair this terrible printer?"

"If you are looking for a good car, get an Audi Passat"



#### **CHALLENGES OF SENTIMENT ANALYSIS**

- 2. Sentiment words are ambiguous, context- and domain dependent
  - "This video game is totally unpredictable" vs
  - "This car is totally unpredictable"

- 3. Sarcasm
  - "Great headphones if you enjoy the noises of other people"



#### **CHALLENGES OF SENTIMENT ANALYSIS**

- 4. Objective sentences that express sentiments
  - "After sleeping on the mattress for two days, a valley has formed in the middle."
  - "The washing machine uses a lot of water."



# **EVALUATION**



Suzan Verberne 2022

- In case of discrete labels: Positive(P) Neutral(U) Negative(N)
  - Precision<sub>pos</sub> =
  - Recall<sub>pos</sub> =

		Actual			
		Pos Neu Neg			
ted	Pos	PP	PU	PN	
edic	Neu	UP	UU	UN	
Pr	Neg	NP	NU	NN	



Suzan Verberne 2022

- In case of discrete labels: Positive(P) Neutral(U) Negative(N)
  - Precision<sub>pos</sub> = PP / (PP+PU+PN)
  - > Recall<sub>pos</sub> = PP / (PP+UP+NP)

		Actual		
		Pos Neu Neg		
ted	Pos	PP	PU	PN
edic	Neu	UP	UU	UN
Pr	Neg	NP	NU	NN



- In case of discrete labels: Positive(P) Neutral(U) Negative(N)
  - Precision<sub>pos</sub> = PP / (PP+PU+PN)
  - Recall<sub>pos</sub> = PP / (PP+UP+NP)

$$F_1^{\text{pos}} = 2 * (P_{\text{pos}} * R_{\text{pos}}) / (P_{\text{pos}} + R_{\text{pos}})$$

Average F-score is computed on positive and negative labels only:

		Actual		
	Pos		Neu	Neg
ted	Pos	PP	PU	PN
edic	Neu	UP	UU	UN
Pr	Neg	NP	NU	NN

$$F_{1}^{PN} = \frac{F_{1}^{Pos} + F_{1}^{Neg}}{2}$$



#### > In case of regression:

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$



Suzan Verberne 2022



SUZAN VERBERNE 2022



#### HOMEWORK

#### Read:

Reuver et al. (2021) Is Stance Detection Topic-Independent and Crosstopic Generalizable? – A Reproduction Study

Exercise week 9: BERTviz (Visualize Attention in NLP Models)

- Follow the tutorial on: <u>https://github.com/jessevig/bertviz</u>
- Make sure you understand what the steps mean
- There is no assignment on this, but you can choose to use it in the final assignment



#### **AFTER THIS LECTURE...**

- > You can define sentiment classification as ordinal regression tasks
- > You can define the components of aspect-based sentiment in texts
- > You can explain the challenges of sentiment analysis
- You can correctly evaluate sentiment classification methods

