



04 - SENTIMENT ANALYSIS

MACHINE LEARNING FOR NATURAL LANGUAGE PROCESSING, AIMS 2024

Dr. Elvis Ndah

OVERVIEW

1. What is sentiment analysis.
2. Why sentiment analysis.
3. Why is sentiment analysis hard.
4. Types of sentiment analysis
5. How sentiment analysis works

WHAT IS SENTIMENT ANALYSIS?

- **Sentiment analysis** refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials.
- Technique to analyzing textual data to determine whether it is positive, negative, or neutral (or other predefined class)

Alternatively:

- Opinion mining
- Sentiment mining
- Subjective detection
- social media analytics

WHAT IS SENTIMENT ANALYSIS?

- Identify the orientation of opinion in a piece of text



- Can be generalized to a wider set of emotions a lot more dimensions than just positive and negative

WHAT IS SENTIMENT ANALYSIS?

- Conceptual Challenge

- Which of the following sentences express sentiment?
- What is their sentiment polarity (positive/negative/neutral)

1. There was an earthquake in Buea.
2. The team failed to complete the physical challenge.
3. They said it would be great.
4. They said it would be great, and they were right.
5. They said it would be great, and they were wrong.
6. The party fat-cats are sipping their expensive imported wines.

WHY SENTIMENT ANALYSIS?



KNOWING SENTIMENT IS A
NATURAL ABILITY FOR HUMAN
BEING.



CAN A COMPUTER BE TRAINED TO
DO IT?



SENTIMENT ANALYSIS AIMS AT
GETTING SENTIMENT-RELATED
KNOWLEDGE ESPECIALLY FROM THE
HUGE AMOUNT OF TEXT FROM THE
INTERNET.



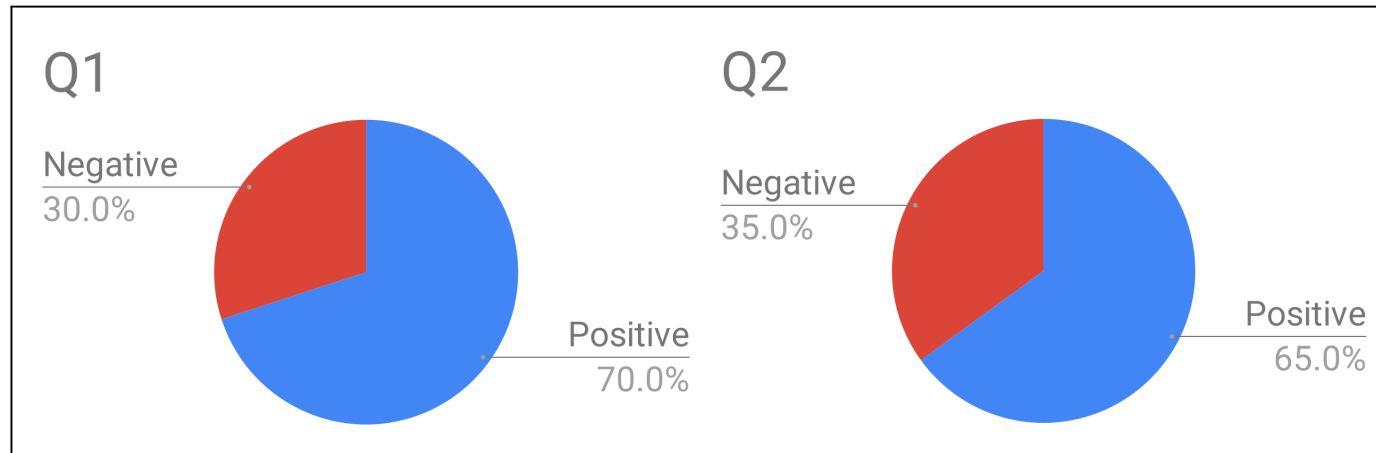
SENTIMENT ANALYSIS IS A MACHINE
LEARNING TECHNIQUE THAT USES
NLP TO IDENTIFY POSITIVE AND
NEGATIVE SENTIMENT IN TEXT.

WHY SENTIMENT ANALYSIS?

- NLP is essential to help business analysis textual data for:
 - Brand monitoring
 - Social Media Monitoring (user sentiments from feedback),
 - Market research,
 - Understand customers' needs
 - Customer services
- **Examples:**
 - Product sold on Amazon allow customers to write reviews
 - Sentiment analysis can help quickly analyse 10k+ product reviews
 - Extract insights for decision making
 - Get insights to determine if the customers are happy with your product or customer service.

WHY IS SENTIMENT ANALYSIS HARD

- There are lots of application of sentiment analysis in business today.
- Businesses often expect charts for decision making.



- However this does not provide relevant information for decision making.
- Such distributions often hide the underlying relevant insights

WHY IS SENTIMENT ANALYSIS HARD

Social media analytics: are we nearly there yet?

Businesses have been trying to crack sentiment analysis and social reach metrics for years, but how close are they to turning social analytics into the gold mine it was always meant to be?

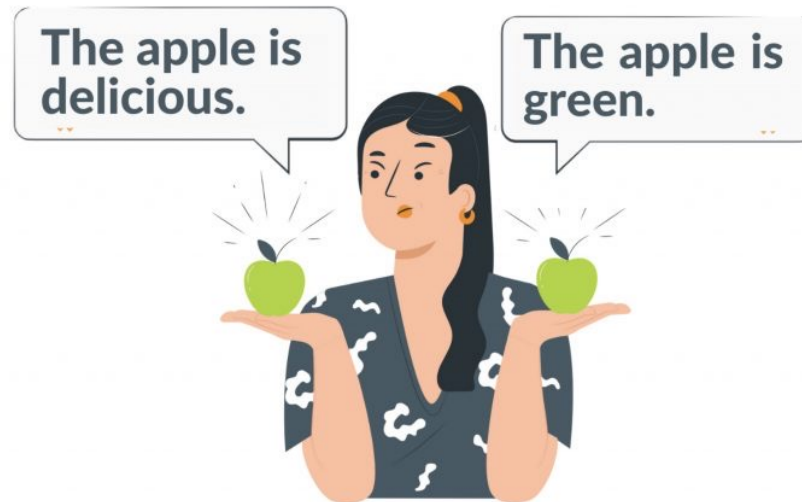
[. . .]

"Anyone who says they're getting better than 70% [today] is lying, generally speaking", said Halstead.

"There has been a clear shift in the last three years - the difficulty with sentiment analysis really is about understanding the context of it, and the tech definitely has got better. We're starting to bridge the gap, and we're way beyond word lists now", said Halstead.

WHY IS SENTIMENT ANALYSIS HARD - SUBJECTIVITY

- Sentiment analysis is one of the most difficult task in NLP.
- **Text Subjectivity and Tone:** textual data is usually of two types:
 - Subjective: often contains sentiment
 - Objective: does not often contain sentiment



WHY IS SENTIMENT ANALYSIS HARD - POLARITY

- **Text Polarity and Context**
 - Analysing text with context increases accuracy
 - Computers learn to analyse context only when they are specifically included
- **Question framing:**
 - *What is it that you like about the game?*
 - Positive context in the question
 - *What is it that you don't like about the game?*
 - Negative context



WHY IS SENTIMENT ANALYSIS HARD – SARCASM

- Irony and Sarcasm in Text
 - Very difficult sentiments for computers to detect
 - People communicate negative sentiments through positive words

What's the sentiment of this tweet?



Based on the context the author is praising the skills

WHY IS SENTIMENT ANALYSIS HARD - ORIENTATION

Opinion orientation

- change according to valence shifters (e.g., negation: *not* etc.)
- **But** clauses (“the pictures are good, but the battery life ...”)
- Dictionary-based: Use semantic relations (e.g., synonyms, antonyms)

Corpus-based orientation

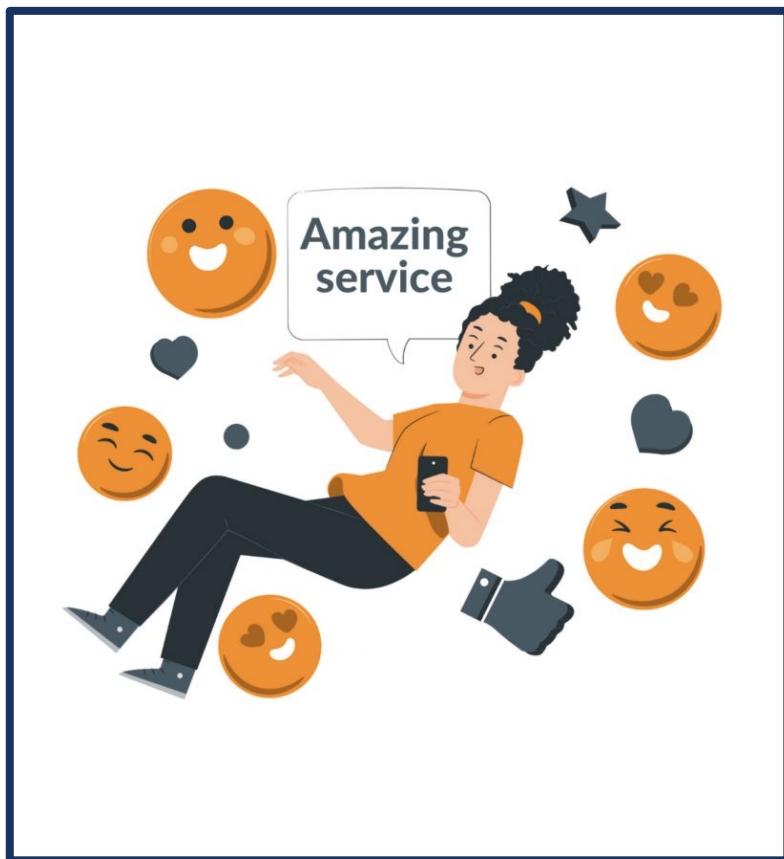
- learn from labelled examples
- Disadvantage: need these (expensive!)
- Advantage: domain dependence

WHY IS SENTIMENT ANALYSIS HARD - COMPARISON

- **Text Comparison**
 - One main challenge in general NLP is to capture comparison in text



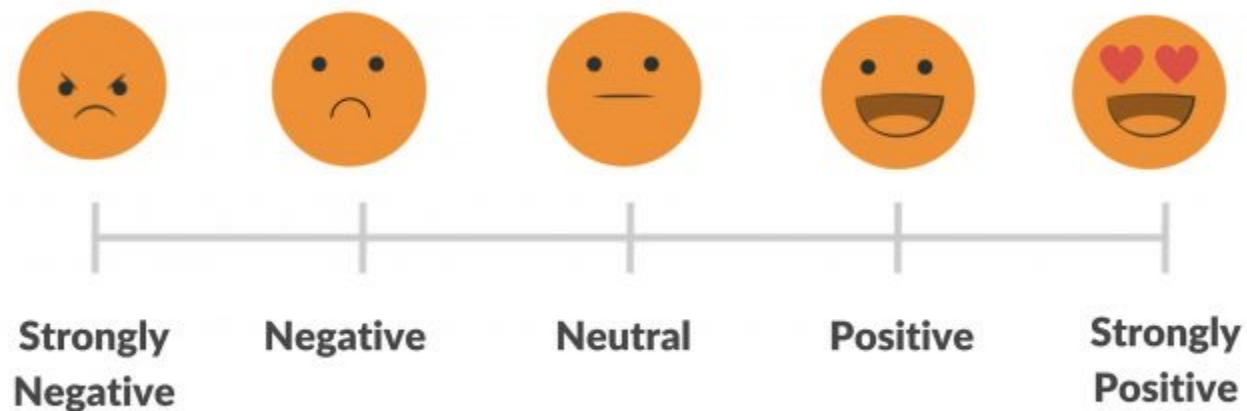
WHY IS SENTIMENT ANALYSIS HARD – EMOJIS



- **Emojis in Text**
 - Emojis play crucial role in expressing sentiments
 - Increasing popularity in the use of emojis to communicate
 - Emojis are replacing slangs
 - The need for character and word level consideration to account for emoji

TYPES OF SENTIMENT ANALYSIS

1. **Fine-grained sentiment analysis:** looks at text polarity to get insights on sentiments expressed



TYPES OF SENTIMENT ANALYSIS

2. Emotion detection: focus on detecting emotions such as happiness, frustration, anger, sadness



■ Example:

- **Negative:** Your customer support is soo **bad**, I **regret** buying your product.
- **Positive:** This is **bad-ass**, I **regret** not trying it earlier.

TYPES OF SENTIMENT ANALYSIS

- **3.Aspect-based Sentiment Analysis:** when businesses want to know aspect of their products that are often discussed

The **mountain** hill hotel was **good**, but the **food** was really **bad**.

Aspect
expression of
location



Positive
sentiment
expression



Aspect
expression of
food



Negative
sentiment
expression



HOW SENTIMENT ANALYSIS WORKS



Sentiment analysis uses natural language processing to interpret human language and machine learning to identify the emotions expressed in textual data.



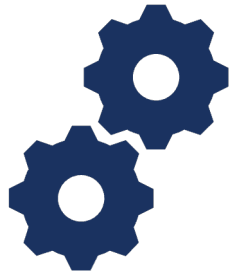
Three main approach to develop a sentiment analysis model (depends on the volume of data and target accuracy).

Rule-Based: performs sentiment analysis based on predefined rules

Automatic: leverage machine learning techniques to learn sentiments from data.

Hybrid: combines both rule-based and automatic analysis approaches.

HOW SENTIMENT ANALYSIS WORKS – HYBRID



Bring together elements of rule-based and automatic systems.

Start with automatic

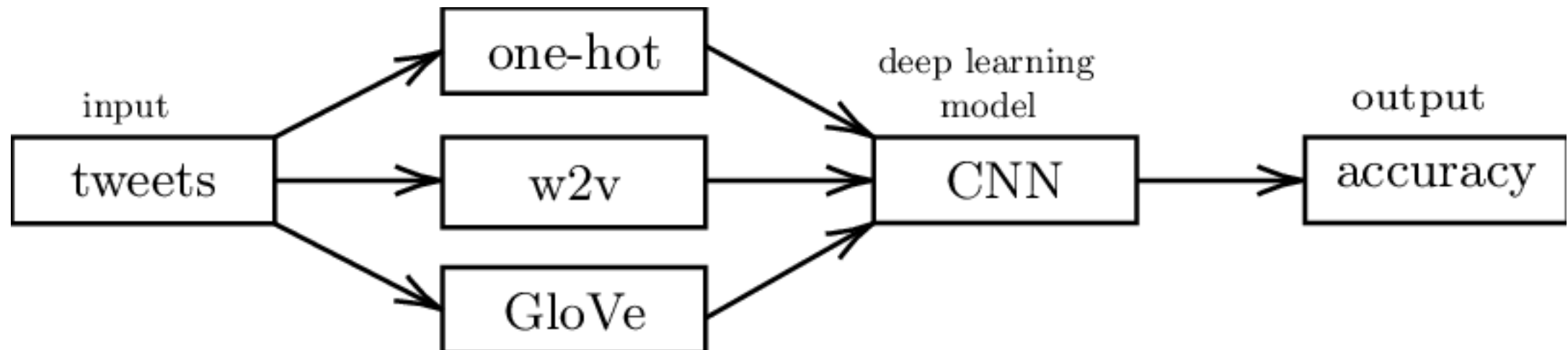
Use rule based to fine tune output

Vice versa



A major benefit of these methods is that they usually give more precise results.

HOW SENTIMENT ANALYSIS WORKS – DL METHODOLOGY



REFERENCES

Subjectivity	(Pang and Lee 2008)
Bias	(Recasens et al. 2013; Pryzant et al. 2020)
Stance	(Anand et al. 2011)
Hate-speech	(Nobata et al. 2016)
Microaggressions	(Breitfeller et al. 2019)
Condescension	(Wang and Potts 2019)
Sarcasm	(Khodak et al. 2017)
Deception and betrayal	(Niculae et al. 2015)
Online trolls	(Cheng et al. 2017)
Polarization	(Gentzkow et al. 2019)
Politeness	(Danescu-Niculescu-Mizil et al. 2013)
Linguistic alignment	(Doyle et al. 2016)