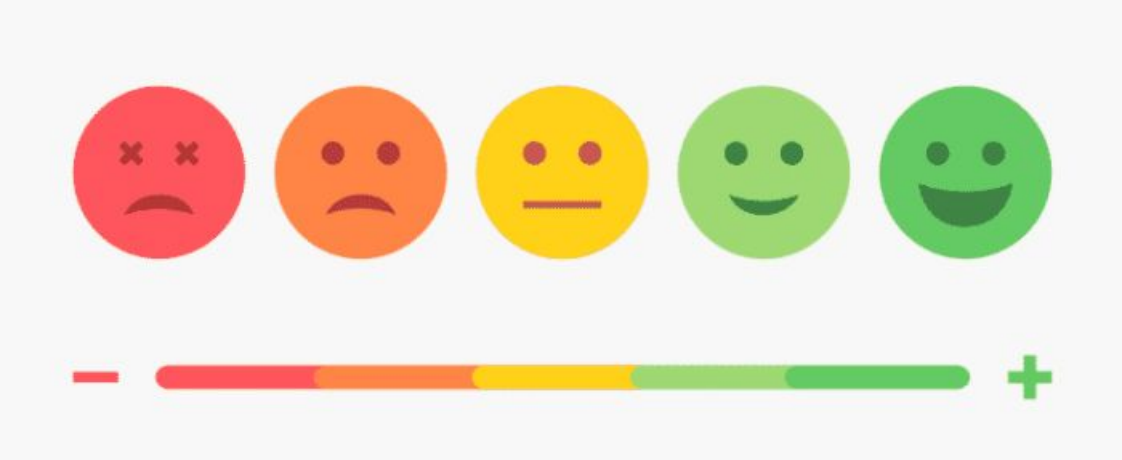
Sentiment Analysis



Done by:

PA29- Shrutika

PA49- Shreya Sundkar

PA62- Shriyash Shingare

PA67- Neel Khalade

What is Sentiment Analysis?

- Sentiment analysis refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.
- Sentiment analysis detects polarity (e.g. a positive or negative opinion) within text, whether a whole document, paragraph, sentence, or clause.
- Sentiment analysis tools allow businesses to identify customer sentiment toward products, brands or services in online feedback.

Types of Sentiment Analysis

1) Fine-grained Sentiment Analysis

Businesses can expand their polarity categories to include:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

Fine-grained sentiment analysis, could be used to interpret 5-star ratings in a review, for example:

- Very Positive = 5 stars
- Very Negative = 1 star

2) Emotion detection

- This type of sentiment analysis aims at detecting emotions, like happiness, frustration, anger, sadness, and so on. Many emotion detection systems use lexicons (lists of words and the emotions they convey) or complex machine learning algorithms.

3) Aspect-based Sentiment Analysis

- Usually, when analyzing sentiments of texts, let's say product reviews, you'll want to know which particular aspects or features people are mentioning in a positive, neutral, or negative way.

4) Multilingual sentiment analysis

- Multilingual sentiment analysis can be difficult. It involves a lot of preprocessing and resources. Most of these resources are available online (e.g. sentiment lexicons), while others need to be created (e.g. translated corpora or noise detection algorithms), but you'll need to know how to code to use them.
- Alternatively, you could detect language in texts automatically with a language classifier, then train a custom sentiment analysis model to classify texts in the language of your choice.

Sentiment Analysis Algorithms

Sentiment analysis uses various Natural Language Processing (NLP) methods and algorithms and the main types of algorithms used include:

1) Rule-based Approaches

Usually, a rule-based system uses a set of human-crafted rules to help identify subjectivity, polarity, or the subject of an opinion.

These rules may include various techniques developed in computational linguistics, such as:

- *Stemming*, *tokenization*, *part-of-speech tagging* and *parsing*.
- Lexicons (i.e. lists of words and expressions).

Here's a basic example of how a rule-based system works:

a) Defines two lists of polarized words (e.g. negative words and positive words).

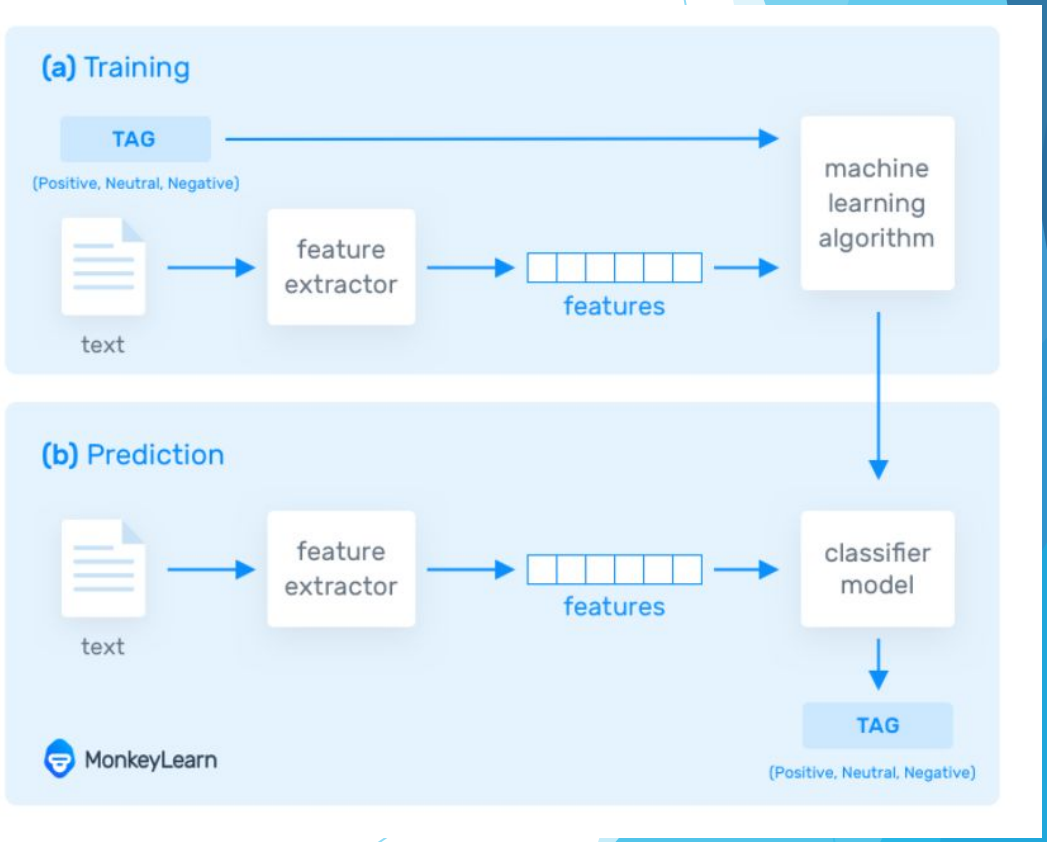
b) Counts the number of positive and negative words that appear in a given text.

c) If the number of positive word appearances is greater than the number of negative word appearances, the system returns a positive sentiment, and vice versa. If the numbers are even, the system will return a neutral sentiment.

2) Automatic Approaches

Automatic methods, contrary to rule-based systems, don't rely on manually crafted rules, but on machine learning techniques. A sentiment analysis task is usually modeled as a classification problem, whereby a classifier is fed a text and returns a category, e.g. positive, negative, or neutral.

- In the training process (a), our model learns to associate a particular input (i.e. a text) to the corresponding output (tag) based on the test samples used for training. The feature extractor transfers the text input into a feature vector. Pairs of feature vectors and tags (e.g. positive, negative, or neutral) are fed into the machine learning algorithm to generate a model.
- In the prediction process (b), the feature extractor is used to transform unseen text inputs into feature vectors. These feature vectors are then fed into the model, which generates predicted tags (again, positive, negative, or neutral).



- Classification Algorithms used:

The classification step usually involves a statistical model like-

a) Naïve Bayes
b) Logistic Regression
c) Support Vector Machines
d) Neural Networks

3) Hybrid Approaches

- Hybrid systems combine the desirable elements of rule-based and automatic techniques into one system. One huge benefit of these systems is that results are often more accurate.

Sentiment Analysis Applications:

1) Social Media Monitoring

- Analyze tweets and/or facebook posts over a period of time to detect sentiment of a particular audience
- Monitor social media mentions of your brand and automatically categorize by urgency
- Automatically route social media mentions to team members best fit to respond
- Automate any or all of these processes
- Gain deep insights into what's happening across your social media channels

2) Brand Monitoring

- Understand how your brand reputation evolves over time
- Research your competition and understand how their reputation also evolves over time.
- Identify potential PR crises and know to take immediate action. Again, prioritize what fires need to be put out immediately and what mentions can wait.
- Tune into a specific point in time. Again, maybe you want to look at just press mentions on the day of your IPO filing, or a new product launch. Sentiment analysis lets you do that.

3) Customer Feedback

- Use results of sentiment analysis to design better informed questions to ask on future surveys
- Understand the nuances of customer experience over time, along with why and how shifts are happening
- Respond more quickly to signals and shifts from customers

4) Customer Service

- Automate text classification all incoming customer support queries
- Route queries to specific team members best suited to respond
- Gain deep insights into what's happening across your customer support

5) Market Research

- Analyze product reviews of your brand and compare those with the competition
- Generate weekly, monthly, or daily reports – a sort of early-warning system
- Compare sentiment across international markets
- Analyze formal market reports or business journals for long-term, broader trends

Sentiment Analysis Challenges

Computer scientists have been trying to develop more accurate sentiment classifiers, and overcome limitations in recent years. Let's take a closer look at some of the challenges they face:

a) Subjectivity and Tone
b) Context and Polarity
c) Irony and Sarcasm
d) Comparisons
e) Emojis
f) Defining Neutral

Thank You!