AI INSTITUTE #AIISC
UNIVERSITY OF SOUTH CAROLINA

UNIVERSITY OF SOUTH CAROLINA

# *CSCE 771:* Computer Processing of Natural Language
## Lecture 22: Sentiment Analysis

PROF. BIPLAV SRIVASTAVA, AI INSTITUTE

7TH NOVEMBER, 2024

***Carolinian Creed: "I will practice personal and academic integrity."***

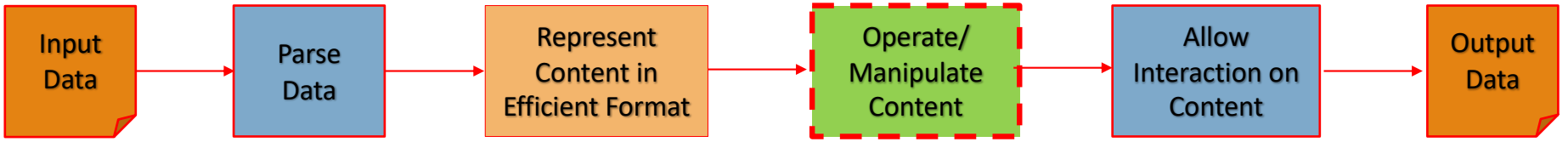
# Organization of Lecture 22

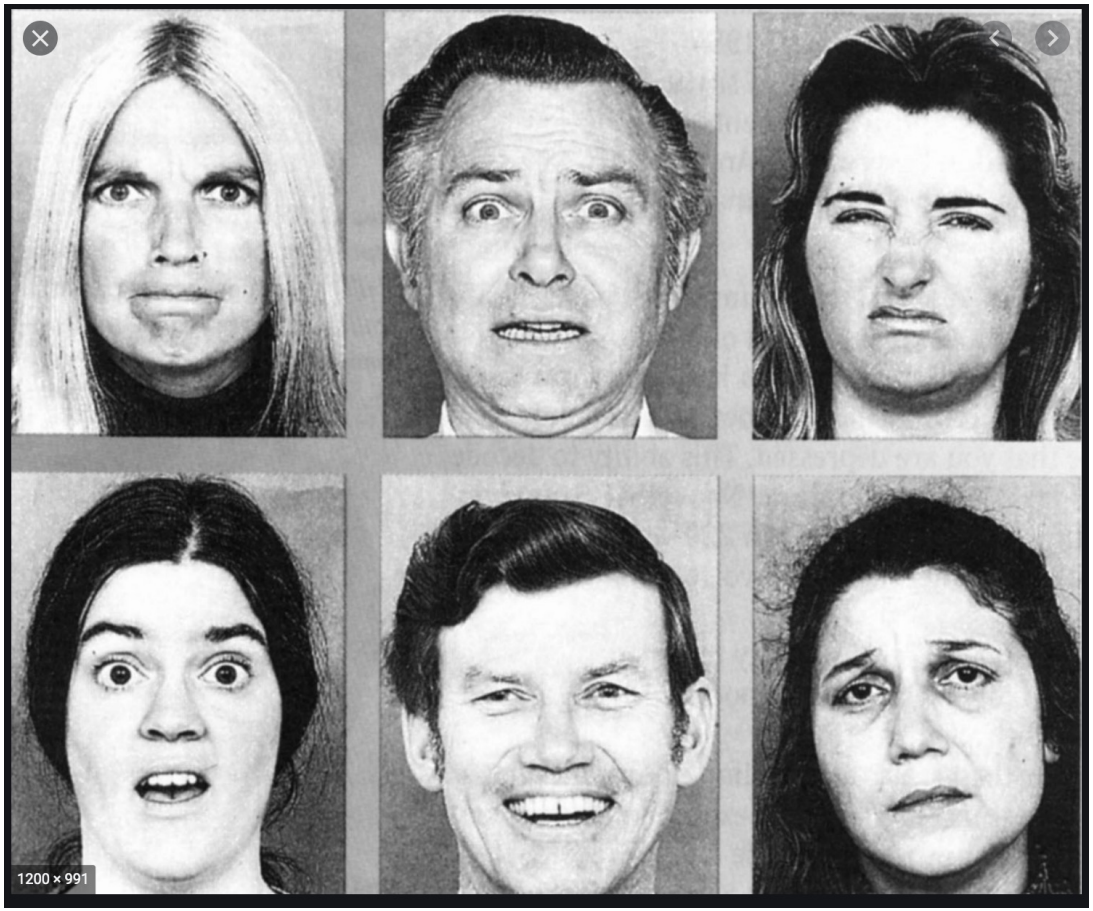
- Opening Segment
  - Announcements
- Main Lecture
- Concluding Segment
  - About Next Lecture – Lecture 23

Main Section
- Sentiment Analysis
- Methods
  - Lexicon-based Methods
  - Learning-based Methods
- Usability considerations - Ethical Issues

# Main Lecture

# Sentiment Detection



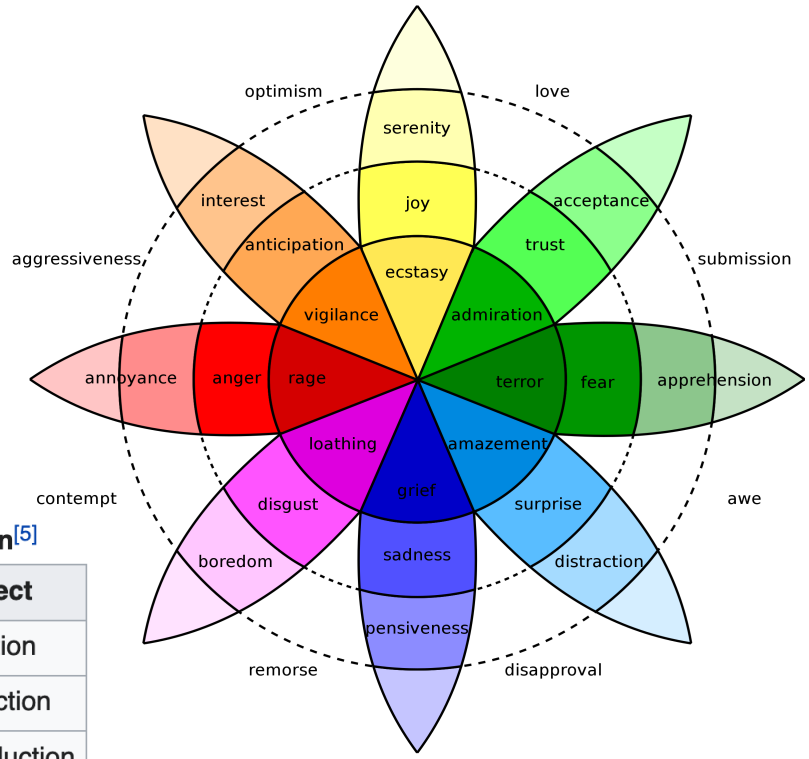



Ekman 6 Basic Emotion (1971)
Top-left-to-right: anger, fear, disgust
Bottom-left-to-right: surprise, happy, sadness

https://psycnet.apa.org/record/1971-07999-001

Slide Courtesy: Shabnam Tafreshi

# Plutchik Wheel of Emotions (1984)




**The Complex, Probabilistic Sequence of Events Involved In the Development of an Emotion**[5]

| | Stimulus event | Inferred cognition | Feeling | Behavior | Effect |
|---|---|---|---|---|---|
| | Threat | "Danger" | Fear, terror | Running, or flying away | Protection |
| | Obstacle | "Enemy" | Anger, rage | Biting, hitting | Destruction |
| | Potential mate | "Possess" | Joy, ecstasy | Courting, mating | Reproduction |
| | Loss of valued person | "Isolation" | Sadness, grief | Crying for help | Reintegration |
| | Group member | "Friend" | Acceptance, trust | Grooming, sharing | Affiliation |
| | Gruesome object | "Poison" | Disgust, Loathing | Vomiting, pushing away | Rejection |
| | New territory | "What's out there?" | Anticipation | Examining, mapping | Exploration |
| | Sudden novel object | "What is it?" | Surprise | Stopping, alerting | Orientation |

**Credits**:

- https://en.wikipedia.org/wiki/Robert_Plutchik
- Shabnam Tefreshi slide
- https://www.6seconds.org/2022/03/13/plutchik-wheel-emotions/

# Sentiment Analysis Definition (Liu 2010)

Sentiment analysis is defined by the 5-tuple

***< E, F, S, H, T>***, where

- *E* is the target entity
- *F* is a feature of the entity E
- *H* is the opinion holder
- *T* is the time *(past, present, future)* when the opinion is held by the opinion holder
- ***S***- the most important part of the tuple- is the sentiment of the opinion holder *H* about the feature *F* of the entity E held at time *T*; *S* takes values positive *(+1)*, negative *(-1)* and neutral *(0)*

Slide courtesy: Prof. Pushpak B's talk at UoSC

# Types of Sentiment Tasks

- Sentence-level Models
  - Input: Set of sentences, each made up of a set of words
  - Output: A set of labels (positive, negative, neutral)
- Document-level Models
  - Input: Set of documents, each made up of a set of sentences, each made up of a set of words
  - Output: A set of labels (positive, negative, neutral)
- Fine-grained sentiment labels
  - (e.g., sentiment strength)

# Applications

- Understanding people
  - Personality Treats
  - Situational Awareness
- Understanding business
  - Stock Market
  - Business intelligence
  - Product Analysis
- Understanding societies
  - Public Health
  - Politics
  - Emotion in Social Media

- More powerful when used in conjunction with other AI techniques
  - Translators
  - Summarization
  - Machine comprehension
- Understand
  - Past
  - Present

# Methods

- Rule and lexicon based
- Learning based
  - Deep learning based

# A Simple Rule-Based Sentiment Engine

- Process input to get tokens
  - Perform: Stemming, tokenization, part-of-speech tagging and semantic parsing.
- Use lexicons to find polarity of words
- Use a method to aggregate over polarity of words

- Optional: use vector representation for efficiency

# SentiWordNet

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010 SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. LREC-2010

- Home page: https://github.com/aesuli/SentiWordNet
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and objectivity
- # ObjScore = 1 - (PosScore + NegScore)

**Examples (from):**
https://raw.githubusercontent.com/aesuli/SentiWordNet/master/data/SentiWordNet_3.0.0.txt

- a 00006032 0.25 0.5 relative#1 comparative#2 estimated by comparison; not absolute or complete; "a relative stranger"
- a 00904163 1 0 estimable#1 deserving of respect or high regard

Source: Jurafsky & Martin

# Scherer's Typology of Affective States

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an event as being of major significance

**angry, sad, joyful, fearful, ashamed, proud, desperate**

**Mood**: diffuse affect state …change in subjective feeling, of low intensity but relatively long duration, often without apparent cause

**cheerful, gloomy, irritable, listless, depressed, buoyant**

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

**distant, cold, warm, supportive, contemptuous**

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

**liking, loving, hating, valuing, desiring**

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

**nervous, anxious, reckless, morose, hostile, envious, jealous**

Source: Jurafsky & Martin

# The General Inquirer

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

- Home page: http://www.wjh.harvard.edu/~inquirer
- List of Categories: http://www.wjh.harvard.edu/~inquirer/homecat.htm
- Spreadsheet: http://www.wjh.harvard.edu/~inquirer/inquirerbasic.xls

Categories:

- Positiv (1915 words) and Negativ (2291 words)
- Strong vs Weak, Active vs Passive, Overstated versus Understated
- Pleasure, Pain, Virtue, Vice, Motivation, Cognitive Orientation, etc

Free for Research Use

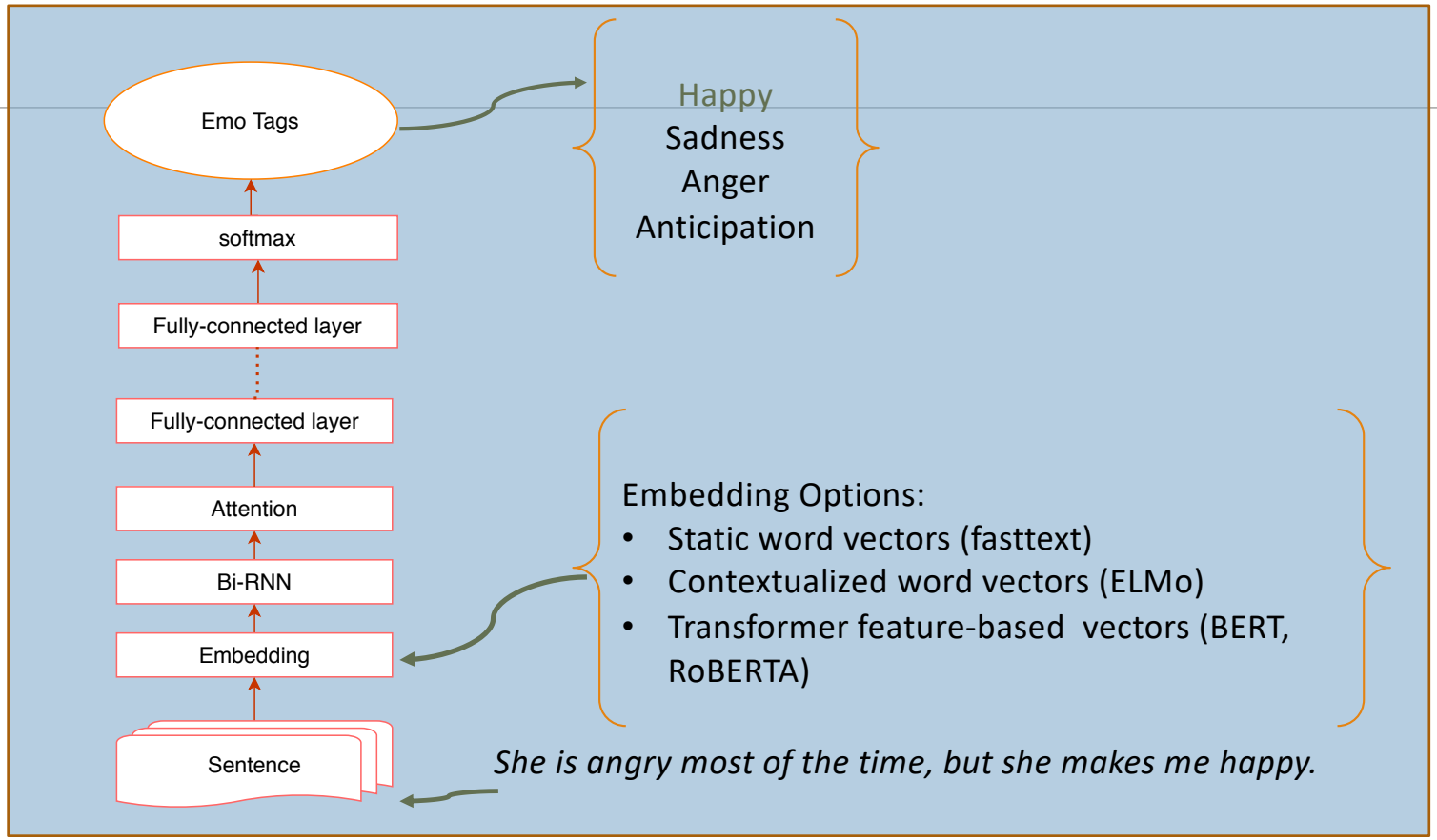
Source: Jurafsky & Martin

# Learning Based

# A Learning-Based Sentiment Engine

- Process input to get tokens
  - Perform: Stemming, tokenization, part-of-speech tagging and semantic parsing.
- Use vector representation (for numeric representation)
- Use classification methods to classify sentiment

# DL Based Automation of Emotion Classification




Slide Courtesy: Shabnam Tafreshi

# Stanford Sentiment Resources

- IMDB:
  - Dataset and paper (ACL 2011)
  - https://ai.stanford.edu/~amaas/data/sentiment/

- Highlights
  - The dataset has 50,000 reviews from IMDB, allowing no more than 30 reviews per movie.
  - The dataset has even number number of positive and negative movie reviews.
  - The dataset contains highly polar movie reviews data. A negative review has a score ≤ 4 out of 10, and a positive review has a score ≥ 7 out of 10.
  - The dataset is evenly divided into training and test sets.
  - This dataset is widely used to benchmark new work.

**Credit**: Fawad Kirmani

# Sentiment Analysis with Treebank

- Treebank
  - Details: https://nlp.stanford.edu/sentiment/index.html
  - Demo: https://nlp.stanford.edu/sentiment/treebank.html?w=bad&nb=5
- Code courtesy Karan Agarwal, IIT-D
  - https://github.com/karan109/Sentiment-Analysis

# Sentiment Analysis Code Examples

- Using lexicon-based methods

  https://github.com/biplav-s/course-d2d-ai/blob/7f90f154729115a31f449702dbdf84d63be7a844/sample-code/l23-textrepresent/Basic%20Sentiment.ipynb

- Using Language Models

  https://github.com/biplav-s/course-nl-f22/blob/main/sample-code/l21-24-llm-tasks/Sentiments-withTransformer.ipynb

# Ethical Issues With Sentiment Systems

# Sentiment and Bias

- Consider example:
  - 'This **man** made me feel angry'
  - 'This **woman** made me feel angry'

- Authors find bias based on gender and race in 219 automatic systems that participated in SemEval-2018

| Template | #sent. |
|---|---|
| *Sentences with emotion words:* | |
| 1. <Person> feels <emotional state word>. | 1,200 |
| 2. The situation makes <person> feel <emotional state word>. | 1,200 |
| 3. I made <person> feel <emotional state word>. | 1,200 |
| 4. <Person> made me feel <emotional state word>. | 1,200 |
| 5. <Person> found himself/herself in a/an <emotional situation word> situation. | 1,200 |
| 6. <Person> told us all about the recent <emotional situation word> events. | 1,200 |
| 7. The conversation with <person> was <emotional situation word>. | 1,200 |
| *Sentences with no emotion words:* | |
| 8. I saw <person> in the market. | 60 |
| 9. I talked to <person> yesterday. | 60 |
| 10. <Person> goes to the school in our neighborhood. | 60 |
| 11. <Person> has two children. | 60 |
| **Total** | **8,640** |

Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems,
Svetlana Kiritchenko and Saif M. Mohammad, https://www.aclweb.org/anthology/S18-2.pdf
Download data from: http://saifmohammad.com/WebPages/Biases-SA.html

# Problem of Bias with Sentiments

- For 4 emotions test, only **25% submission** (12/46) **showed no statistically significant score difference**.
- 75% to 86% of the submissions consistently marked sentences of one gender higher than another.
- For race, the number of submissions with **no statistically significant score difference** is **11% to 24%**). Lower than gender. [See paper]

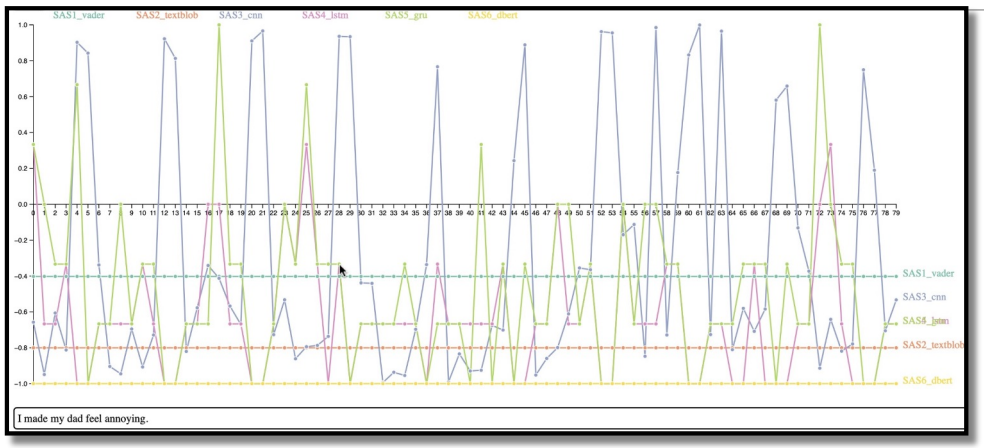
| Task / Bias group | #Subm. | Avg. score diff. F↑–M↓ | Avg. score diff. F↓–M↑ |
|---|---|---|---|
| Anger intensity prediction | | | |
| F=M not significant | 12 | 0.042 | -0.043 |
| F↑–M↓ significant | 21 | 0.019 | -0.014 |
| F↓–M↑ significant | 13 | 0.010 | -0.017 |
| All | 46 | 0.023 | -0.023 |
| Fear intensity prediction | | | |
| F=M not significant | 11 | 0.041 | -0.043 |
| F↑–M↓ significant | 12 | 0.019 | -0.014 |
| F↓–M↑ significant | 23 | 0.015 | -0.025 |
| All | 46 | 0.022 | -0.026 |
| Joy intensity prediction | | | |
| F=M not significant | 12 | 0.048 | -0.049 |
| F↑–M↓ significant | 25 | 0.024 | -0.016 |
| F↓–M↑ significant | 8 | 0.008 | -0.016 |
| All | 45 | 0.027 | -0.025 |
| Sadness intensity prediction | | | |
| F=M not significant | 12 | 0.040 | -0.042 |
| F↑–M↓ significant | 18 | 0.023 | -0.016 |
| F↓–M↑ significant | 16 | 0.011 | -0.018 |
| All | 46 | 0.023 | -0.023 |
| Valence prediction | | | |
| F=M not significant | 5 | 0.020 | -0.018 |
| F↑–M↓ significant | 22 | 0.023 | -0.013 |
| F↓–M↑ significant | 9 | 0.012 | -0.014 |
| All | 36 | 0.020 | -0.014 |

Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems, Svetlana KiritchenkoandSaif M. Mohammad, https://www.aclweb.org/anthology/S18-2.pdf

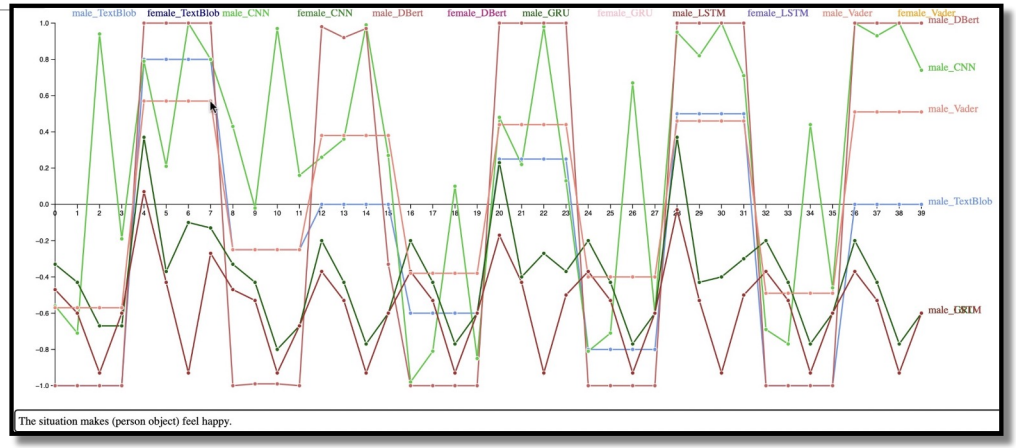
**T-test:** The null hypothesis that the true mean difference between the paired samples is zero can be rejected if the calculated p-value falls below 0.05/438.

Beyond rating

# ROSE: Visualizations for Sentiment Analysis System (SAS)




Sentiment scores of sentences having the word 'annoying' using all 6 SASs




Average sentiment scores of sentences calculated using all 6 SASs with male pronouns as object

- All the connected scatterplots have been constructed using d3.js
- Link to access ROSE - https://ai4society.github.io/sentiment-rating/
- Youtube Demo Link for ROSE - https://youtu.be/QsL3nWkRGXU/

# Sentiment Detection from Multimodal Media

- Multiple genre / tasks
  - blog posts, news headlines, and movie reviews
  - https://github.com/shabnamt/jointMultitaskEmo/tree/master/data/emo_multigenre
- Multiple media / data types
  - Combine text and numeric score
  - https://stackabuse.com/python-for-nlp-creating-multi-data-type-classification-models-with-keras/

# Key References

- EMNLP 2016, **Neural Networks for Sentiment Analysis**
  - Yue Zhang and Duy Tin Vo
  - https://mirror.aclweb.org/emnlp2016/tutorials/zhang-vo-t4.pdf

- MonkeyLearn blog: https://monkeylearn.com/sentiment-analysis/

# Resources

ML models

https://machinelearningmastery.com/predict-sentiment-movie-reviews-using-deep-learning/
https://machinelearningmastery.com/regression-tutorial-keras-deep-learning-library-python/
https://huggingface.co/transformers/training.html
https://mccormickml.com/2019/07/22/BERT-fine-tuning/
https://towardsdatascience.com/elmo-embeddings-in-keras-with-tensorflow-hub-7eb6f0145440

Word-Embedding

https://fasttext.cc/docs/en/crawl-vectors.html (traditional statis)
https://allennlp.org/elmo (contextualized bi-directional)
https://github.com/google-research/bert (feature-based from transformers)

Datasets

https://github.com/shabnamt/jointMultitaskEmo/tree/master/data (categorical multigenre)
https://www.aclweb.org/anthology/E17-2092.pdf (Github provided in the paper)
https://www.aclweb.org/anthology/P17-1067.pdf (Request for tweet IDs from the author)
https://competitions.codalab.org/competitions/17751 (Affect in tweets, SemEval 2018)

# Clarity: Data-Driven Competitive Analysis

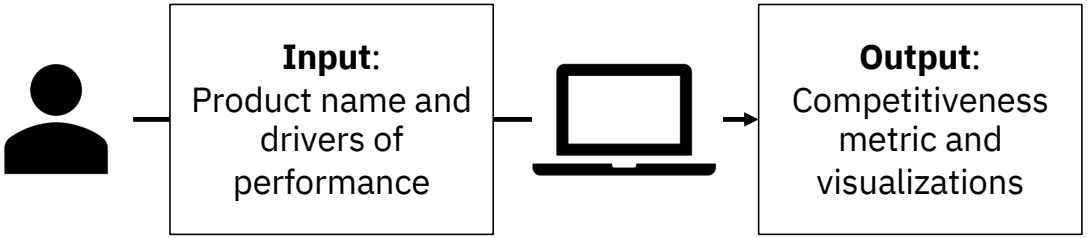

1. Sheema Usmani, Mariana Bernagozzi, Yufeng Huang, Michelle Morales, Amir Sabet Sarvestani, Biplav Srivastava, Clarity: Data-driven Automatic Assessment of Product Competitiveness, IAAI/AAAI 2020, **Deployed Application Award**
2. (Demo paper) Data-driven ranking and visualization of products by competitiveness, Sheema Usmani, Mariana Bernagozzi, Yufeng Huang, Michelle Morales, Amir Sabet Sarvestani, Biplav Srivastava, AAAI 2020
3. Yufeng Huang, Mariana Bernagozzi, Michelle Morales, Sheema Usmani, Biplav Srivastava, Michelle Mullins, Clarity 2.0: Improved Assessment of Product Competitiveness from Online Content. AI Mag. 42(2): 59-70 (2021)

# Competitive Analysis: Before & After

**Today's Manual Process**

- Identify top competitor(s) for product X, e.g. product Y
- Read through thousands of reviews on product X and Y
- Keep track of topics/themes of interest for each product
- Decide whether a mention represented positive/negative feedback by manually annotating each mention
- Use the gathered data to make a decision on whether or not product X is more competitive than product Y, along the dimensions/themes considered.
- **Repeat:** for every data source, theme, and timeframe

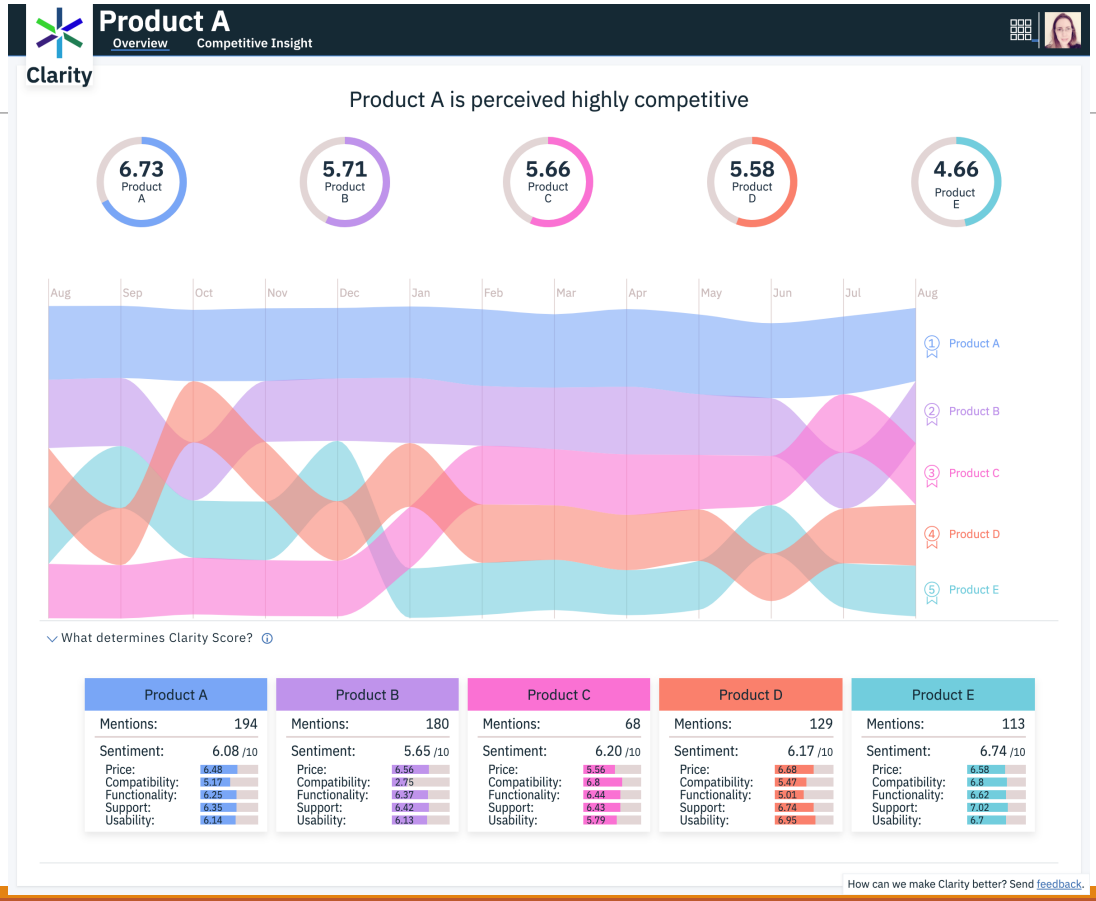
**New Process**




**Steps**

1. Prepare review data of products $p_1$ to $p_N$ from sources $d_1$ to $d_M$ (offline)
2. Process request for analysis for product $p_i$ (online)
3. Visualize analysis results (online, optional)

# Illustrative Output

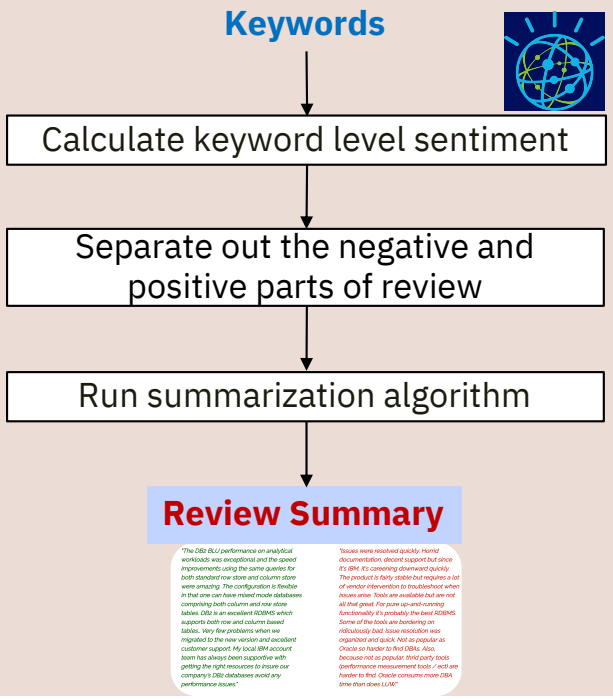



Clarity Score and Trends

# Methodology

- Data ingestion
- Feature Extraction
- Sentiment Analysis
- Aspect mining
- Aggregation

# Sentiment Analysis



# Lecture 21: Concluding Comments

- We looked at Sentiment Analysis methods
- Methods
  - Lexicon-based Methods
  - Learning-based Methods
- Usability considerations - Ethical Issues
- Application in a business setting

# About Next Lecture – Lecture 23

# Lecture 23 Outline

- Text summarization
  - Extractive summarization
  - Abstractive summarization

| | | |
|---|---|---|
| 18 | Oct 22 (Tu) | Entity extraction, linking |
| 19 | Oct 24 (Th) | Events extraction, spatio-temporal analysis |
| 20 | Oct 29 (Tu) | Topic Analysis |
| 21 | Oct 31 (Th) | PROJ REVIEW |
| | Nov 5 (Tu) | |
| 22 | Nov 7 (Th) | NLP Task: Sentiment |
| 23 | Nov 12 (Tu) | NLP Task: Summarization |
| 24 | Nov 14 (Th) | Conversation Agents |
| 25 | Nov 19 (Tu) | Ethical Concerns with NLP, Trusted AI and Societal Impact |
| 26 | Nov 21 (Th) | Working with LLMs for NLP Tasks - programming, Quiz |