Sentiment Analysis

Pablo Mosteiro (many slides by Ayoub Bagheri)

Recap

- What is a deep-learning word embedding?
- 2 Wooclap questions

The Little Prince example

This is a nice book for both young and old. It gives beautiful life lessons in a fun way. Definitely worth the money!

+ Educational

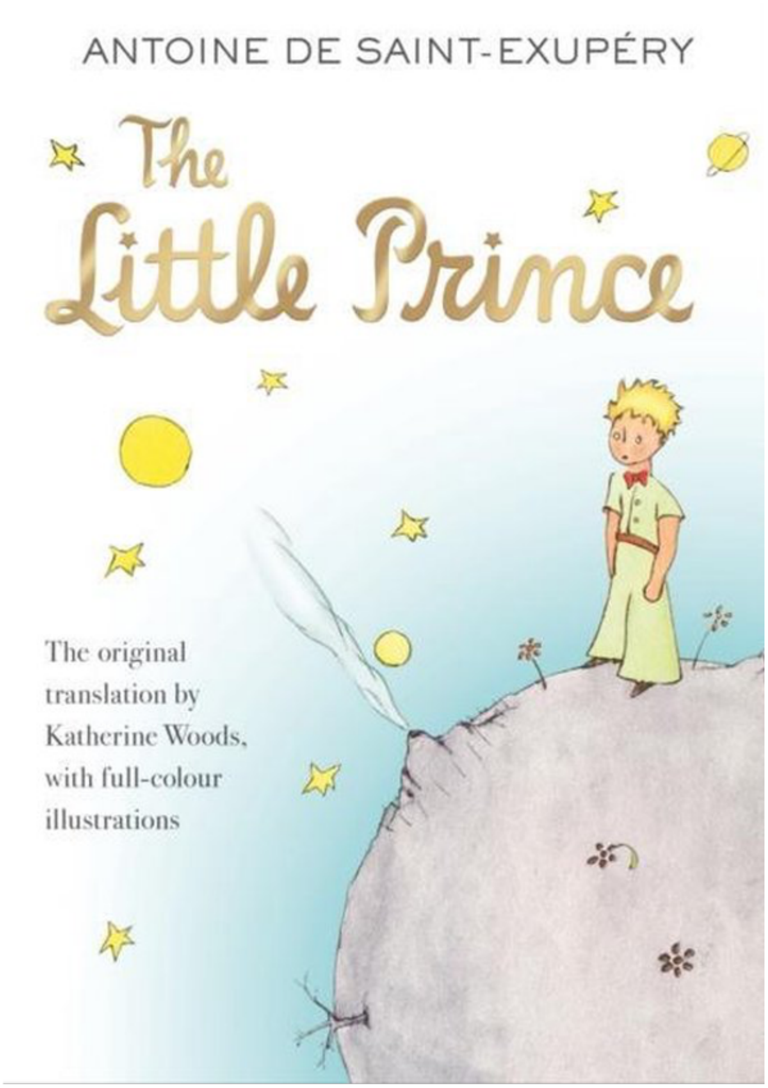
+ Fun

+ Price

Nice story for older children.

+ Funny

- Readability



Conceptual challenges

Wooclap time

Sentiment

- Sentiment =
  - Feelings, Attitudes, Emotions, Opinions
  - A thought, view, or attitude, especially one based mainly on emotion instead of reason
- Subjective impressions, not facts

Sentiment analysis

- Use of natural language processing (NLP) and computational techniques to automate the extraction or classification of sentiment from unstructured text
- Other terms
  - Opinion mining

- Sentiment mining
- Sentiment classification

Related tasks

- Subjectivity (neutral vs sentimental text)
- Emotion detection (e.g., happiness, anger, sadness)
- Stance detection (in favor or against)
- Reputation analysis
- Sarcasm/Irony detection
- Hate-speech

Sentiment analysis

- Can be applied in every topic & domain (non exhaustive list):
  - Examples?

Sentiment analysis

- Can be applied in every topic & domain (non exhaustive list):
  - Book: is this review positive or negative?
  - Humanities: sentiment analysis for German historic plays.
  - Products: what do people think about the new iPhone?
  - Blog: how are people thinking about immigrants?
  - Politics: who is going to win the election?
  - Social Media: what is the trend today?
  - Movie: is this review positive or negative (IMDB, Netflix)?
  - Marketing: how is consumer confidence? Consumer attitudes?
  - Healthcare: are patients happy with the hospital environment?

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
- Comparative opinions: ?

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
  - E.g.?
- Comparative opinions: Comparison of more than one entity.
  - E.g., "iPhone is better than Blackberry."

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
  - "The touch screen is really cool."
- Comparative opinions: Comparison of more than one entity.
  - E.g., "iPhone is better than Blackberry."

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
  - Direct opinions:
    * "The touch screen is really cool."
  - Indirect opinions:
    * Example?
- Comparative opinions: Comparison of more than one entity.
  - E.g., "iPhone is better than Blackberry."

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
  - Direct opinions:
    * "The touch screen is really cool."
  - Indirect opinions:
    * "After taking the drug, my pain has gone."
- Comparative opinions: Comparison of more than one entity.
  - E.g., "iPhone is better than Blackberry."

Opinion types

- Regular opinions: Sentiment/opinion expressions on some target entities
  - Direct opinions:
    * "The touch screen is really cool."
  - Indirect opinions:
    * "After taking the drug, my pain has gone."
      · Positive or negative? About what/whom?
- Comparative opinions: Comparison of more than one entity.
  - E.g., "iPhone is better than Blackberry."

Practical definition

- An opinion is a quintuple (entity, aspect, sentiment, holder, time) where
  - entity: target entity (or object).
  - aspect: aspect (or feature) of the entity.
  - sentiment: +, -, or neu, a rating, or an emotion.
  - holder: opinion holder.
  - time: time when the opinion was expressed.

Sentiment analysis tasks

- Simplest task:
  - Is the attitude of this text positive or negative?
- More complex:
  - Is the attitude of this text positive, negative or neutral?
  - Label the attitude of this text from 1 to 5
- Advanced:
  - Detect the target, source, or complex opinion types
  - Implicit opinions or aspects

Document sentiment analysis

- Classify a document (e.g., a review) based on the overall sentiment of the opinion holder
  - Classes: Positive, negative (possibly neutral)
    - Neutral means no sentiment expressed
    - "I believe he went home yesterday."
    - "I bought an iPhone yesterday"
- An example review:
  - "I bought an iPhone a few days ago. It is such a nice phone, although a little large. The touch screen is cool. The voice quality is great too. I simply love it!"
  - Classification: positive or negative?
- It is basically a text classification problem

Sentence sentiment analysis

- Classify the sentiment expressed in a sentence
  - Classes: positive, negative (possibly neutral)
- But bear in mind
  - Explicit opinion: "I like this car."
  - Fact-implied opinion: "I bought this car yesterday and it broke today."
  - Mixed opinion: "Apple is doing well in this poor economy"

Challenges

- Think Pair Share

Challenges

- Hard to do with bag of words
- Must consider other features due to…
  - Subtlety of sentiment expression
    - irony (What a great car, it stopped working in the second day.)
    - expression of sentiment using neutral words (The concert didn't meet my expectations.)
  - Domain/context dependence
    - words/phrases can mean different things in different contexts and domains (long queue vs long battery life)
  - Effect of syntax on semantics (Negation)

Methods for sentiment analysis

- **Lexicon-based methods**
  - Dictionary based: Using sentiment words and phrases (e.g., good, wonderful, awesome, troublesome, cost an arm and leg)
  - Corpus-based: Using co-occurrence statistics or syntactic patterns embedded in text corpora
- **Supervised learning methods**: to classify reviews into positive and negative.
  - Traditional Machine Learning: Naïve Bayes, Support Vector Machine
  - Deep learning: BERT, GPT

Lexicon-based Methods

Sentiment and other lexicons

- Lists of words that are associated with sentiment scores
- Can have binary scores (1, -1) or intensity scores (from 0 to 1)
- Positive/negative polarity, emotions, affective states, negation lists

```
brainwashing    -3
brave   2
breakthrough    3
breathtaking    5
bribe   -3
bright  1
brightest       2
brightness      1
brilliant       4
brisk   2
broke   -1
broken  -1
```

Basic Lexicon Approach |

- Detect sentiment in two independent dimensions:
  - Positive: {1, 2,... 5}
  - Negative: {-5, -4,... -1}
- Example: "He is brilliant but boring"
  - Overall sentiment = ?

Basic Lexicon Approach |

- Detect sentiment in two independent dimensions:
  - Positive: {1, 2,... 5}
  - Negative: {-5, -4,... -1}
- Example: "He is brilliant but boring"
  - Sentiment('brilliant') = +4
  - Sentiment('boring') = -2
  - Overall sentiment = +2

LIWC (Linguistic Inquiry and Word Count) | Tausczik and Pennebaker (2011)

- 2,300 words, >70 classes
- Affective Processes
  - negative emotion (bad, weird, hate, problem, tough)
  - positive emotion (love, nice, sweet)
- Cognitive Processes
  - Tentative (maybe, perhaps, guess), Inhibition (block, constraint)
- Pronouns, Negation (no, never), Quantifiers (few, many)

VADER Sentiment Analysis | Hutto and Gilbert (2014)

- VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool designed specifically for social media text. Contains a pre-built lexicon of words that are associated with sentiment scores ranging from -4 to +4
- Five generalizable heuristics based on grammatical and syntactical cues:
  - Punctuation: "The food here is good!!!" vs "The food here is good."
  - Capitalization: "The food here is GREAT!" vs "The food here is great!"
  - Degree modifiers: "The service here is extremely good" vs "The service here is good"
  - The conjunction "but": "The food here is great, but the service is horrible" has mixed sentiment
  - For negation examine the tri-gram preceding a sentiment lexical feature: "The food here isn't really all that great"

SentiWordNet | Esuli and Sebastiani (2006)

- https://github.com/aesuli/SentiWordNet
- All WordNet synsets automatically annotated for degrees of positivity, negativity, and objectivity
- [estimable(J,3)] "may be computed or estimated"

$$\text{Pos} \quad 0 \quad \text{Neg} \quad 0 \quad \text{Obj} \quad 1$$

- [estimable(J,1)] "deserving of respect or high regard"

$$\text{Pos} \quad .75 \quad \text{Neg} \quad 0 \quad \text{Obj} \quad .25$$

How to measure polarity of a phrase?

- Positive phrases co-occur more with "excellent"
- Negative phrases co-occur more with "poor"
- But how to measure co-occurrence?

Pointwise Mutual Information

- PMI between two words:
  - How much more do two words co-occur than if they were independent?

$$PMI(word_1, word_2) = log_2 \frac{P(word_1, word_2)}{P(word_1)P(word_2)}$$

How to estimate PMI

- P(word) estimated by `hits(word)/N`
- P(word1,word2) by `hits(word1 NEAR word2)/N^2`

$$PMI(word_1, word_2) = log_2 \frac{hits(word_1 \text{ NEAR } word_2)}{hits(word_1) hits(word_2)}$$

Does phrase appear more with “poor” or “excellent”?

Polarity(*phrase*) = PMI(*phrase*, “excellent”) - PMI(*phrase*, “poor”)

Lexicon-based methods in summary

- Intuition
  - Start with a seed set of words (“good”, “poor”)
  - Find other words that have similar polarity:
    * Using “and” and “but”
    * Using words that occur nearby in the same document
    * Using synonyms and antonyms
    * Using rules based on punctuation, emoticons

Lexicon-based methods in summary (contd)

- Advantages and Disadvantages:
  - Think Pair Share

Lexicon-based methods in summary (contd)

- Advantages:
  - Can be domain-independent with general purpose lexicons
  - Can become domain-dependent
  - Can be easy to rationalise prediction output
  - Can be applied when no training data is available
- Disadvantages:
  - Compared to a well-trained, in-domain ML model they typically underperform
  - Sensitive to affective dictionary coverage

Supervised Methods

Basic steps

- Pre-processing and tokenization
- Feature representation
- Feature selection
- Classification
- Evaluation

Sentiment tokenization issues

- Deal with HTML and XML markup
- Twitter mark-up (names, hash tags)
- Capitalization (preserve for words in all caps)
- Phone numbers, dates
- Emoticons
- Useful code:
  - Christopher Potts sentiment tokenizer
  - Brendan O’Connor twitter tokenizer

Potts emoticons

```
[<>]?                          # optional hat/brow
[:;=8]                         # eyes
[\-o\*\']?                     # optional nose
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
|                              #### reverse orientation
[\)\]\(\[dDpP/\:\}\{@\|\\]     # mouth
[\-o\*\']?                     # optional nose
[:;=8]                         # eyes
[<>]?                          # optional hat/brow
```

The danger of stemming

- The Porter stemmer identifies word suffixes and strips them off.
- But:
  - objective (pos) and objection (neg) -> object
  - competence (pos) and compete (neg) -> compet

Features for supervised learning

- The problem has been studied by numerous researchers.
- Key: feature engineering. A large set of features have been tried by researchers. E.g.,
  - Terms frequency and different IR weighting schemes
  - Part of speech (POS) tags
  - Opinion words and phrases
  - Negations
  - Stylistic
  - Syntactic dependency

Negation

Add NOT_ to every word between negation and following punctuation:

```
didn’t like this movie , but I
```

```
didn’t NOT_like NOT_this NOT_movie but I
```

Challenges of negation

- “terrible” vs “wasn’t terrible”
  - The movie was terrible
  - The movie was bad but wasn’t that terrible as they said
- The degree of the intensity shift varies from term to term for both positive and negative terms

Supervised sentiment analysis | Kiritchenko et al. (2014)

- A supervised statistical text classification approach based on surface, semantic, and sentiment features.
- For negation: estimate sentiment scores of individual terms in the presence of negation
- One lexicon for words in negated contexts and one for words in affirmative

Supervised sentiment analysis | Kiritchenko et al. (2014)

- Features:
  - ngrams
  - character ngrams
  - all-caps: the number of tokens with all characters in upper case
  - POS
  - the number of negated contexts
  - sentiment lexicons
  - the number of hashtags, punctuation, emoticons, elongated words
- Classifier: linear-kernel SVM

Supervised sentiment analysis | Kiritchenko et al. (2014)

Feature group	**Examples**
word ngrams	*grrreat*, *show*, *grrreat_show*, *miss_NEG*, *miss_NEG_the*
character ngrams	*grr*, *grrr*, *grrre*, *rrr*, *rrre*, *rrrea*
all-caps	all-caps:1
POS	POS_N:1 (nouns), POS_V:2 (verbs), POS_E:1 (emoticons), POS_,:1 (punctuation)
automatic lexicon features	HS_unigrams_positive_count:4, HS_unigrams_negative_total_score:1.51, HS_unigrams_POS_N_combined_total_score:0.19, HS_bigrams_positive_total_score:3.55, HS_bigrams_negative_max_score:1.98
manual lexicon features	MPQA_positive_affirmative_score:2, MPQA_negative_negated_score:1, BINGLIU_POS_V_negative_negated_score:1
punctuation	punctuation_!:1
emoticons	emoticon_positive:1, *emoticon_positive_last*
elongated words	elongation:1
clusters	*cluster_11111001110*, *cluster_10001111*

Table 6: Examples of features that the system would generate for message "GRRREAT show!!! Hope not to miss the next one :)". Numeric features are presented in the format: <feature_name>:<feature_value>. Binary features are italicized; only features with value of 1 are shown.

Supervised sentiment analysis

- Advantages
  - Lead to better performance compared to lexicon based approaches
  - The output can be explained (most of the times)
- Disadvantages
  - They need training data
  - They can't capture the context
  - Based on feature engineering that is a tedious task
  - Not good performance in multiclass classification

Deep Learning

Sentiment-specific word embedding | Tang et al. (2014)

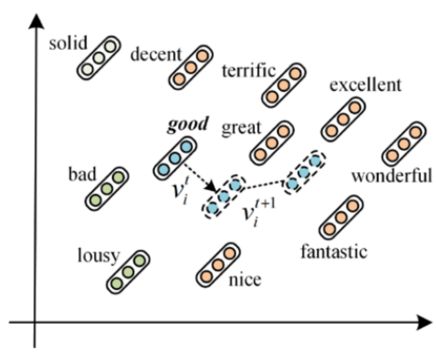
- Continuous word representations model the syntactic context of words but ignore the sentiment of text
- Good vs bad: They will be represented as neighboring word vectors
- Solution: Learn sentiment specific word embedding, which encodes sentiment information in the continuous representation of words

Word vector refinement | Yu et al. (2017)

- Start with a set of pre-trained word vectors and a sentiment lexicon
- Calculate the semantic similarity between each sentiment word and the other words in the lexicon based on the cosine distance of their pre-trained vectors
- Select top-k most similar words as the nearest neighbors and re-rank according to sentiment scores

Word vector refinement | Yu et al. (2017)

- Refine the pre-trained vector of the target word to be:
  - closer to its sentimentally similar neighbors,
  - further away from its dissimilar neighbors, and
  - not too far away from the original vector.



Sentiment analysis with BERT | Devlin et al. 2019

- Sentiment analysis was one of the tasks in the BERT paper

System	MNLI-(m/mm)	QQP	QNLI	SST-2
	392k	363k	108k	67k
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.8	90.4
OpenAI GPT	82.1/81.4	70.3	87.4	91.3
$\text{BERT}_{\text{BASE}}$	84.6/83.4	71.2	90.5	93.5
$\text{BERT}_{\text{LARGE}}$	**86.7/85.9**	**72.1**	**92.7**	**94.9**

Pre-trained models on SA

https://huggingface.co/blog/sentiment-analysis-python

- Twitter-roberta-base-sentiment is a roBERTa model trained on ~58M tweets and fine-tuned for sentiment analysis (https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment)
- SST-2 BERT: Fine-tuned on the Stanford Sentiment Treebank (SST-2) which consists of sentences from movie reviews. The model is well-suited for general sentiment analysis tasks. (https://huggingface.co/distilbert-base-uncased-finetuned-sst-2-english)
- Bert-base-multilingual-uncased-sentiment is a model fine-tuned for sentiment analysis on product reviews in six languages: English, Dutch, German, French, Spanish and Italian (https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment)
- Distilbert-base-uncased-emotion is a model fine-tuned for detecting emotions in texts, including sadness, joy, love, anger, fear and surprise (https://huggingface.co/bhadresh-savani/distilbert-base-uncased-emotion)

Bias in Sentiment Analysis

Bias in sentiment analysis | Kiritchenko and Mohammad (2018)

- Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems (Kiritchenko & Mohammad, *SEM 2018)
- Are systems that detect sentiment biased?
- Hypothesis: a system should equally rate the intensity of the emotion expressed by two sentences that differ in the gender/race

Bias in sentiment analysis | Kiritchenko and Mohammad (2018)

Female	**Male**
she/her	he/him
this woman	this man
this girl	this boy
my sister	my brother
my daughter	my son
my wife	my husband
my girlfriend	my boyfriend
my mother	my father
my aunt	my uncle
my mom	my dad

African American		European American	
Female	**Male**	**Female**	**Male**
Ebony	Alonzo	Amanda	Adam
Jasmine	Alphonse	Betsy	Alan
Lakisha	Darnell	Courtney	Andrew
Latisha	Jamel	Ellen	Frank
Latoya	Jerome	Heather	Harry
Nichelle	Lamar	Katie	Jack
Shaniqua	Leroy	Kristin	Josh
Shereen	Malik	Melanie	Justin
Tanisha	Terrence	Nancy	Roger
Tia	Torrance	Stephanie	Ryan

Bias in sentiment analysis | Kiritchenko and Mohammad (2018)

- > 75% of systems mark one gender/race with higher intensity scores than other
- more widely prevalent for race than for gender
- impact on downstream applications?

Bias in sentiment analysis

What about biases in LLMs?

- DistilBERT base uncased finetuned SST-2:
  - "This movie was filmed in France" -> ?
  - "This movie was filmed in Afghanistan" -> ?

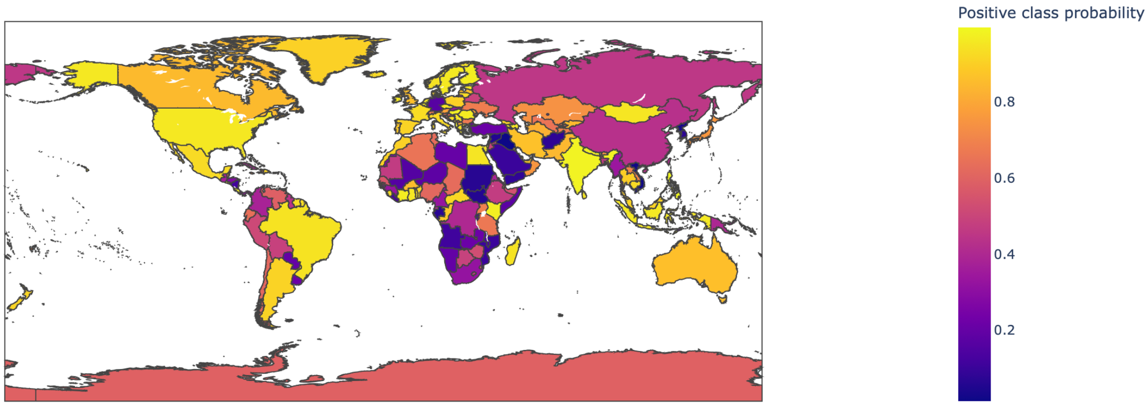
Bias in sentiment analysis

What about biases in LLMs?

- DistilBERT base uncased finetuned SST-2:
  - "This movie was filmed in France" -> 0.89
  - "This movie was filmed in Afghanistan" -> 0.08

Bias in sentiment analysis

- "This movie was filmed in {country_name}"



From Aurélien Géron colab

Summary

Summary

- Sentiment analysis
- Lexicon-based methods
- Learning-based methods
- Sentiment-aware word embeddings
- Bias

Resources

- Crawl your own data from Twitter:
  - https://developer.twitter.com/en/docs/twitter-api
- SemEval Datasets: 2012-now
  - https://semeval.github.io/
- Stanford Sentiment Treebank:
  - https://nlp.stanford.edu/sentiment/
- Sanders Corpus:
  - https://github.com/zfz/twitter_corpus
- IMDB movie reviews (50K)
  - https://ai.stanford.edu/~amaas/data/sentiment/
- Datasets from Bing Liu's group:
  - https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
- Amazon review data
  - https://nijianmo.github.io/amazon/index.html
- iSarcasm
  - https://github.com/dmbavkar/iSarcasm

Lexicons and tools

- VADER (Hutto and Gilbert, 2014)
  - https://github.com/cjhutto/vaderSentiment
- LIWC
  - https://www.liwc.app/
- Bing Liu

- https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
- Multi-Perspective Question Answering - MPQA (Wiebe et al., 2005)
  - https://mpqa.cs.pitt.edu/lexicons/subj_lexicon/
- SentiWordNet (Esuli and Sebastiani, 2006)
  - https://github.com/aesuli/SentiWordNet
- NRC Lexicons
  - http://saifmohammad.com/WebPages/lexicons.html
- AFFINN (Nielsen, 2011)
  - https://github.com/fnielsen/afinn

Tutorials

- Sentiment analysis in huggingface
- Sentiment analysis with BERT

References

- Bagheri, A., Saraee, M, and De Jong, F. 2013. Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews. Knowledge-Based Systems. 52, 201–213
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
- Esuli, A. and Sebastiani, F. 2006. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)
- Ghanem, B., Rosso, P., & Rangel, F. 2020. An emotional analysis of false information in social media and news articles. ACM Transactions on Internet Technology (TOIT), 20(2), 1-18
- Giachanou, A., Rosso, P., & Crestani, F. 2019. Leveraging emotional signals for credibility detection. In Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval. pp. 877-880
- Hu, M., & Liu, B. 2004. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 168-177

References (Contd)

- Hutto, C. and Gilbert, E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media, vol. 8, no. 1, pp. 216-225
- Kiritchenko, S., and Saif M. 2018. Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems. In Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics, pp. 43-53. 2018
- Kiritchenko, S., Zhu, X., & Mohammad, S. M. 2014. Sentiment analysis of short informal texts. Journal of Artificial Intelligence Research, 50, 723-762
- Narayanan, R., Liu, B., & Choudhary, A. 2009. Sentiment analysis of conditional sentences. In Proceedings of the 2009 conference on empirical methods in natural language processing. pp. 180-189
- Nielsen, F. A. 2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages 718 in CEUR Workshop Proceedings 93-98
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., & Qin, B. 2014. Learning sentiment-specific word embedding for twitter sentiment classification. In ACL. pp. 1555-1565

References (Contd)

- Tausczik, Y. R., & Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of language and social psychology, 29(1), 24-54

- Turney, P. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. pp. 417-424
- Vosoughi, S., Roy, D., & Aral, S. 2018. The spread of true and false news online. science, 359(6380), 1146-1151
- Wiebe, J., Wilson, T., & Cardie, C. 2005. Annotating expressions of opinions and emotions in language. Language resources and evaluation, 39(2), 165-210
- Yu, L. C., Wang, J., Lai, K. R., & Zhang, X. 2017. Refining word embeddings for sentiment analysis. In Proceedings of the 2017 conference on empirical methods in natural language processing. pp. 534-539

Practical 8