01 – Opinion mining, sentiment analysis
IA161 Advanced Techniques of Natural Language Processing

Z. Nevěřilová

NLP Centre, FI MU, Brno

September 15, 2021

Opinion mining, sentiment analysis

Example 1:

> So *boring*. *I enjoyed the first book* but *this one* really *didn't work* for me. The *story*, *characters*, and *relationships* all fell *flat*.

Example 2:

> *Lair of Dreams* like everything else Miss Bray writes is *mind-boggling*. It's *big*. It's *insanely atmospheric* and it's *creeptastic*.[1]

this book: boring	Lair of Dreams: mind-boggling
first book: enjoyed	LoD: big
this book: did not work	LoD: insanely atmospheric
story: flat	LoD: creeptastic
characters: flat	
relationships: flat	

---

[1] both examples from goodreads.com

Opinion mining, sentiment analysis

Opinion mining / sentiment analysis / (emotion AI):

Given a set of subjective texts that express opinions about a certain object, the purpose is to extract those attributes (features) of the object that have been commented on in the given texts and to determine whether these texts are positive, negative or neutral. [Dinu and Iuga, 2012]

Automatic opinion mining: why?

- many subjective texts exist
- mostly because of social media
  - ▸ people express their opinions in texts
  - ▸ one's opinions influence others' opinions
  - ▸ aggregation of opinions
- review sites influence customer behavior (decision making)
- framing in news ("freedom fighters" vs. "terrorists")
- emotions make part of a decision process (see [Minsky, 2007])

"Opinions" are key influencers of our behaviors. [Liu, 2012]

Opinion mining: related applications

- document sentiment classification:
  This document contains a lot of negative statements.
- sentence subjectivity classification:
  This sentence is objective.
- aspect-based opinion summarization/aggregation:
  Most customers of your company think that the communication is not good.
- mining comparative opinions:
  Many people think that iPhone is better than SG.
- utility or helpfulness of reviews:
  This review is useless.
- sarcasm detection:
  I truly love to spend a night in this hotel.
- toxic speech detection:
  No skills. Shut it down.
- cross-lingual opinion mining

Problem definition

What is an opinion?

- an evaluating proposition: *Linux is great.*
- a comparative proposition: *Linux is better than Windows.*

  An opinion is simply a positive or negative sentiment, view, attitude, emotion, or appraisal about an entity or an aspect of the entity from an opinion holder. [Liu, 2012]

entity *e* is a product, person, event, organization, or topic: iPhone, Madonna, Microsoft . . .
aspect *a* (feature) is a component of *e* or attribute of *e*: battery, price, appearance, communication skills . . .

Problem definition

opinion $= (e_j, a_{jk}, so_{ijkl}, h_i, t_l)$, where

- e_j is a target entity.
  named entity recognition

- a_{jk} is an aspect/feature of the entity e_j.
  information extraction

- so_{ijkl} is the sentiment value of the opinion from the opinion holder h_i on feature a_{jk} of entity e_j at time t_l.
  sentiment identification

- h_i is an opinion holder.
  information extraction

- t_l is the time when the opinion is expressed.
  information extraction

not just one problem
anaphora resolution + synonym matching

Problem granularity

Generally, find structure in unstructured data (text)

- document level opinion mining: *The document is negative.*
- sentence level: *The sentence is negative.*
- object/entity and feature/aspect level: *iPhone is expensive.*

Classification task:

- 2-classes: positive/negative
- 3-classes: positive/negative/neutral
- 5-classes . . .

A hard problem (sometimes)

- opinion mining in tweets is relatively easy (short texts, hashtags)
  usually 3-classes classification for each tweet
- opinion mining in reviews is harder but still the form contains aspects
  and the reviewer has to mark the review positive/negative
  usually 2-classes classification for each aspect (e.g. high price)
- opinion mining in discussions, comments, blogs is very hard

sentiment lexicon

evaluative words: nice, cool, shit, bad. . .
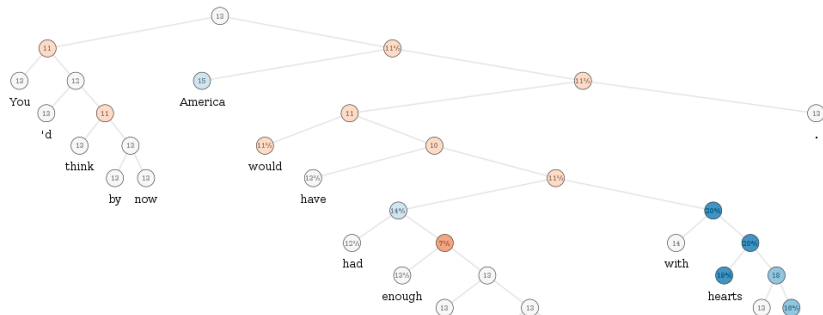SentiWordNet [Baccianella et al., 2010]
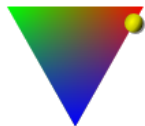
Positive: 0 Objective: 0.125 Negative: 0.875
blue = filled with melancholy and despondency

A hard problem (sometimes) II

evaluative word	aspect	sentiment
thin	phone	good
thin	steak	bad
high	value	good
high	price	bad
flat	story	bad
flat	phone	good

Sentiment analysis methods: supervised machine learning

1. get example data with labels
2. extract features from the data, i.e. convert the documents to feature vectors
3. train the parameters (choose an algorithm: SVM, Naive Bayes, Neural Networks ...)
4. test the model

Sentiment analysis methods: supervised machine learning

[Dinu and Iuga, 2012] report best results on Naive-Bayes with tokens as features and bigrams as features
[Liu, 2012] reports best results with SVM on balanced (English) data
From approx. 2016, non-English SA performed using automatic translation.

Sentiment analysis methods: deep learning

Simple use of word embeddings is questionable, since context vectors do not distinguish polarity (e.g. *good* and *bad* occur in similar contexts and thus have similar vectors.
[Ma et al., 2018] LSTM with two-level attention (target-level + sentence-level)
SA is sometimes solved using multi-task oriented methods: Autoencoders (BERT), Autoregressive models, or combination (XLNet, [Yang et al., 2020])

Datasets

Lexicons (Word lists)

- SentiWordNet
- afinn
- Subjectivity Lexicon
- Bing Liu's Lexicon

Texts

- IMDB Movie Reviews
- Sentiment140
- OpinRank Review Dataset
- Toxic Comment Classification Challenge

Sentiment analysis methods: state-of-the-art results

- on political tweets, [Maynard and Funk, 2012]: 78% precision and 47% recall
- on document level (movie reviews)[Richa Sharma and Jain, 2014]: 63% accuracy and 70% recall
- sentiment embeddings [Tang et al., 2016]: outperform `word2vec` by about 6 percentage points,
  F1 of Twitter Sentiment Classification on SemEval Datasets:
  pos/neg class: 86.6%
  pos/neg/neu class: 67.5%
  hybrid ranking model (neural net catching context and sentiment) + text features (word n-grams, character n-grams, . . .)
- a survey on using deep learning for sentiment analysis: [Zhang et al., 2018]
- XLNet [Yang et al., 2020] solves SA together with other tasks (multi-task): 97% accuracy on SST-2 (binary classification).

References I

📄 Baccianella, S., Esuli, A., and Sebastiani, F. (2010).
Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining.
In Chair), N. C. C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

📄 Dinu, L. P. and Iuga, I. (2012).
The Naive Bayes classifier in opinion mining: In search of the best feature set.
In Gelbukh, A., editor, *Computational Linguistics and Intelligent Text Processing*, volume 7181 of *Lecture Notes in Computer Science*, pages 556–567. Springer Berlin Heidelberg.

References II

📄 Liu, B. (2012).
Sentiment analysis and opinion mining.
Synthesis Lectures on Human Language Technologies, 5(1):1–167.

📄 Ma, Y., Peng, H., and Cambria, E. (2018).
Targeted aspect-based sentiment analysis via embedding
commonsense knowledge into an attentive lstm.
Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).

📄 Maynard, D. and Funk, A. (2012).
Automatic detection of political opinions in tweets.
In García-Castro, R., Fensel, D., and Antoniou, G., editors, *The
Semantic Web: ESWC 2011 Workshops*, volume 7117 of *Lecture
Notes in Computer Science*, pages 88–99. Springer Berlin Heidelberg.

References III

Minsky, M. (2007).
The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind.
SIMON & SCHUSTER.

Richa Sharma, S. N. and Jain, R. (2014).
Opinion mining of movie reviews at document level.
International Journal of Information Theory, 3(3):13–21.

Tang, D., Wei, F., Qin, B., Yang, N., Liu, T., and Zhou, M. (2016).
Sentiment embeddings with applications to sentiment analysis.
IEEE Transactions on Knowledge and Data Engineering, 28(2):496–509.

References IV

📄 Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., and Le, Q. V. (2020).
Xlnet: Generalized autoregressive pretraining for language understanding.

📄 Zhang, L. J., Wang, S., and Liu, B. (2018).
Deep learning for sentiment analysis: A survey.
Wiley Interdiscip. Rev. Data Min. Knowl. Discov., 8.