Supervised sentiment analysis: Overview

Christopher Potts

Stanford Linguistics

CS224u: Natural language understanding

Overview of this unit

1. Sentiment as a deep and important NLU problem
2. General practical tips for sentiment analysis
3. The Stanford Sentiment Treebank (SST)
4. The DynaSent dataset
5. sst.py
6. Methods: hyperparameters and classifier comparison
7. Feature representation
8. RNN classifiers

Associated materials

1. Code
   a. `sst.py`
   b. `sst_01_overview.ipynb`
   c. `sst_02_hand_build_features.ipynb`
   d. `sst_03_neural_networks.ipynb`

2. Homework and bake-off: `hw_sentiment.ipynb`

3. Core reading: Socher et al. 2013; Potts et al. 2020

4. Auxiliary readings: Pang and Lee 2008; Goldberg 2015

Conceptual challenges

Which of the following sentences express sentiment? What is their sentiment polarity (pos/neg), if any?

1. There was an earthquake in California.
2. The team failed to complete the challenge. (We win/lose!)
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.
7. Oh, you're terrible!
8. Here's to ya, ya bastard!
9. Of 2001, "Many consider the masterpiece bewildering, boring, slow-moving or annoying, . . . "
10. long-suffering fans, bittersweet memories, hilariously embarrassing moments, . . .

Sentiment analysis in industry

Real-world performance falls short

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.

EMOTION AI TECHNOLOGY HAS GREAT PROMISE (WHEN USED RESPONSIBLY)

Affective computing knows how you feel. Sorta.

Stephen Gossett

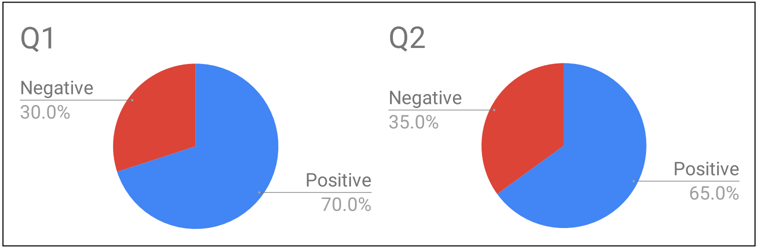
January 7, 2021

Updated: March 2, 2021

Burn-Murdoch 2013; Gossett 2020

Lots of applications, but what's the real goal?

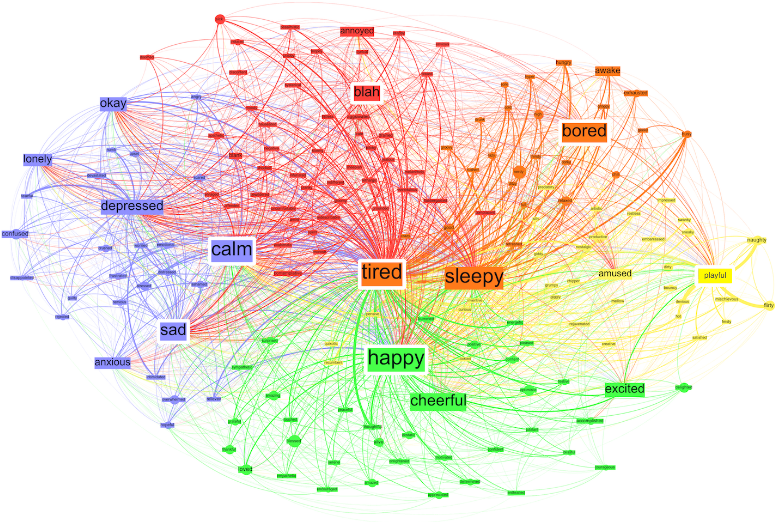
Many business leaders think they want this:



When they see it, they realize that it does not help them with decision-making. The distributions (assuming they are accurately measured) are hiding the phenomena that are actually relevant.

Affective computing

Affective dimensions, relations, and transitions



(Sudhof et al. 2014)

Related tasks in affective computing

With selected papers that make excellent entry points because of their positioning and/or associated public data:

- 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)

Our primary datasets

1. Ternary formulation of the Stanford Sentiment Treebank (SST-3; Socher et al. 2013)
2. The DynaSent dataset (Potts et al. 2020)
3. Our bakeoff data: dev/test splits from SST-3 and from a new (unreleased) corpus of sentences from restaurant reviews
4. Ternary sentiment throughout:
   a. Positive
   b. Negative
   c. Neutral

References I

Pranav Anand, Marilyn Walker, Rob Abbott, Jean E. Fox Tree, Robeson Bowmani, and Michael Minor. 2011. Cats rule and dogs drool!: Classifying stance in online debate. In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, pages 1–9, Portland, Oregon. Association for Computational Linguistics.

Luke Breitfeller, Emily Ahn, Aldrian Obaja Muis, David Jurgens, and Yulia Tsvetkov. 2019. Finding microaggressions in the wild: A case for locating elusive phenomena in social media posts. In *Proceedings of 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing*. Association for Computational Linguistics.

John Burn-Murdoch. 2013. Social media analytics: Are we nearly there yet? *The Guardian*.

Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. Anyone can become a troll: Causes of trolling behavior in online discussions. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1217–1230, New York, NY, USA. ACM.

Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. In *Proceedings of the 2013 Annual Conference of the Association for Computational Linguistics*, pages 250–259, Stroudsburg, PA. Association for Computational Linguistics.

Gabriel Doyle, Dan Yurovsky, and Michael C. Frank. 2016. A robust framework for estimating linguistic alignment in twitter conversations. In *Proceedings of the 25th International World Wide Web Conference*, WWW '16, pages 637–648, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Matthew Gentzkow, Jesse M. Shapiro, and Matt Taddy. 2019. Measuring group differences in high-dimensional choices: Method and application to congressional speech. Ms, Stanford University, Brown University and Amazon.

Yoav Goldberg. 2015. A primer on neural network models for natural language processing. Ms., Bar Ilan University.

Stephen Gossett. 2020. Emotion AI has great promise (when used responsibly). *Built In Blog*.

Mikhail Khodak, Nikunj Saunshi, and Kiran Vodrahalli. 2017. A large self-annotated corpus for sarcasm. *arXiv preprint arXiv:1704.05579*.

Vlad Niculae, Srijan Kumar, Jordan Boyd-Graber, and Cristian Danescu-Niculescu-Mizil. 2015. Linguistic harbingers of betrayal: A case study on an online strategy game. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1650–1659, Beijing, China. Association for Computational Linguistics.

Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang. 2016. Abusive language detection in online user content. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 145–153, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1):1–135.

Christopher Potts, Zhengxuan Wu, Atticus Geiger, and Douwe Kiela. 2020. DynaSent: A dynamic benchmark for sentiment analysis. *arXiv preprint arXiv:2012.15349*.

References II

Reid Pryzant, Richard Diehl Martinez, Nathan Dass, Sadao Kurohashi, Dan Jurafsky, and Diyi Yang. 2020. Automatically neutralizing subjective bias in text. In *Proceedings of AAAI*.

Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic models for analyzing and detecting biased language. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1650–1659, Sofia, Bulgaria. Association for Computational Linguistics.

Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Stroudsburg, PA. Association for Computational Linguistics.

Moritz Sudhof, Andrés Gómez Emilsson, Andrew L. Maas, and Christopher Potts. 2014. Sentiment expression conditioned by affective transitions and social forces. In *Proceedings of 20th Conference on Knowledge Discovery and Data Mining*, pages 1136–1145, New York. ACM.

Zijian Wang and Christopher Potts. 2019. TalkDown: A corpus for condescension detection in context. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3702–3710, Stroudsburg, PA. Association for Computational Linguistics.