

Sentiment/tone (Automated Content Analysis)

AUTHOR

Valerie Hase

KEYWORDS

sentiment analysis, emotions, dictionary, supervised machine learning

BRIEF DESCRIPTION

Sentiment/tone describes the way issues or specific actors are described in coverage. Many analyses differentiate between negative, neutral/balanced or positive sentiment/tone as broader categories, but analyses might also measure expressions of incivility, fear, or happiness, for example, as more granular types of sentiment/tone. Analyses can detect sentiment/tone in full texts (e.g., general sentiment in financial news) or concerning specific issues (e.g., specific sentiment towards the stock market in financial news or a specific actor).

The datasets referred to in the table are described in the following paragraph:

Puschmann (2019) uses four data sets to demonstrate how sentiment/tone may be analyzed by the computer. Using Sherlock Holmes stories (18th century, N = 12), tweets (2016, N = 18,826), Swiss newspaper articles (2007-2012, N = 21,280), and debate transcripts (2013-2017, N = 205,584), he illustrates how dictionaries may be applied for such a task. Rauh (2019) uses three data sets to validate his organic German language dictionary for sentiment/tone. His data consists of sentences from German parliament speeches (1991-2013, N = 1,500), German-language quasi-sentences from German, Austrian and Swiss party manifestos (1998-2013, N = 14,008) and newspaper, journal and news wire articles (2011-2012, N = 4,038). Silge and Robinson (2020) use six Jane Austen novels to demonstrate how dic-

tionaries may be used for sentiment analysis. Van Atteveldt and Welbers (2020) use state of the Union speeches (1789-2017, N = 58) for the same purpose. The same authors (van Atteveldt & Welbers, 2019) show based on a dataset of N = 2,000 movie reviews how supervised machine learning might also do the trick. In their Quanteda tutorials, Watanabe and Müller (2019) demonstrate the use of dictionaries and supervised machine learning for sentiment analysis on UK newspaper articles (2012-2016, N = 6,000) as well as the same set of movie reviews (n = 2,000). Lastly, Wiedemann and Niekler (2017) use state of the Union speeches (1790-2017, N = 233) to demonstrate how sentiment/tone can be coded automatically via a dictionary approach.

FIELD OF APPLICATION/THEORETICAL FOUNDATION

Related to theories of “Framing” and “Bias” in coverage, many analyses are concerned with the way the news evaluates and interprets specific issues and actors.

REFERENCES/COMBINATION WITH OTHER**METHODS OF DATA COLLECTION**

Manual coding is needed for many automated analyses, including the ones concerned with sentiment. Studies for example use manual content analysis to develop dictionaries, to create training sets on which algorithms used for automated classification are trained, or to validate the results of automated analyses (Song et al., 2020).

REFERENCES

Puschmann, C. (2019). Automatisierte Inhaltsanalyse mit R. Retrieved from <http://inhaltsanalyse-mit-r.de/index.html>



<https://doi.org/10.34778/1d>

© 2021, the authors. This work is licensed under the “Creative Commons Attribution – NonCommercial – NoDerivatives 4.0 International” license (CC BY-NC-ND 4.0)

- Rauh, C. (2018). Validating a sentiment dictionary for German political language—A workbench note. *Journal of Information Technology & Politics*, 15(4), 319–343.
- Silge, J., & Robinson, D. (2020). Text mining with R. A tidy approach. Retrieved from <https://www.tidytextmining.com/>
- Song, H., Tolochko, P., Eberl, J.-M., Eisele, O., Greussing, E., Heidenreich, T., Lind, F., Galyga, S., & Boomgaarden, H.G. (2020) In validations we trust? The impact of imperfect human annotations as a gold standard on the quality of validation of automated content analysis. *Political Communication*, 37(4), 550-572.
- van Atteveldt, W., & Welbers, K. (2019). Supervised Text Classification. Retrieved from https://github.com/ccs-amsterdam/r-course-material/blob/master/tutorials/r_text_ml.md
- van Atteveldt, W., & Welbers, K. (2020). Supervised Sentiment Analysis in R. Retrieved from https://github.com/ccs-amsterdam/r-course-material/blob/master/tutorials/sentiment_analysis.md
- Watanabe, K., & Müller, S. (2019). Quanteda tutorials. Retrieved from <https://tutorials.quanteda.io/>
- Wiedemann, G., Niekler, A. (2017). Hands-on: a five day text mining course for humanists and social scientists in R. Proceedings of the 1st Workshop Teaching NLP for Digital Humanities (Teach4DH@GSCL 2017), Berlin. Retrieved from <https://tm4ss.github.io/docs/index.html>

Table 1. Measurement of „Sentiment/Tone“ using automated content analysis.

Author(s)	Sample	Procedure	Formal validity check with manual coding as benchmark*	Code
Puschmann (2019)	(a) Sherlock Holmes stories (b) Tweets (c) Swiss newspaper articles (d) German Parliament transcripts	Dictionary approach	Not reported	http://inhaltsanalyse-mit-r.de/sentiment.html
Rauh (2018)	(a) Bundestag speeches (b) Quasi-sentences from German, Austrian and Swiss party manifestos (c) Newspapers, journals, agency reports	Dictionary approach	Reported	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BKBXWD

Author(s)	Sample	Procedure	Formal validity check with manual coding as benchmark*	Code
Silge & Robinson (2020)	Books by Jane Austen	Dictionary approach	Not reported	https://www.tidytext-mining.com/sentiment.html
van Atteveldt & Welbers (2020)	State of the Union speeches	Dictionary approach	Reported	https://github.com/ccs-amsterdam/r-course-material/blob/master/tutorials/sentiment_analysis.md
van Atteveldt & Welbers (2019)	Movie reviews	Supervised Machine Learning Approach	Reported	https://github.com/ccs-amsterdam/r-course-material/blob/master/tutorials/r_text_ml.md
Watanabe & Müller (2019)	Newspaper articles	Dictionary approach	Not reported	https://tutorials.quanteda.io/advanced-operations/targeted-dictionary-analysis/
Watanabe & Müller (2019)	Movie reviews	Supervised Machine Learning Approach	Reported	https://tutorials.quanteda.io/machine-learning/nb/
Wiedemann & Niekler (2017)	State of the Union speeches	Dictionary approach	Not reported	https://tm4ss.github.io/docs/Tutorial_3_Frequency.html

** Please note that many of the sources listed here are tutorials on how to conduct automated analyses – and therefore not focused on the validation of results. Readers should simply read this column as an indication in terms of which sources they can refer to if they are interested in the validation of results.*