

# Pensum STV2022

Solveig Bjørkholt\*

Martin Søyland†

## Bøker (må kjøpes)

- Grimmer J, Roberts ME, Stewart BM (2022). *Text as Data: A New, Framework for Machine Learning and the Social Sciences*. Princeton, University Press.

## Bøker (tilgjengelig online, kan kjøpes)

- Feldman R, Sanger J (2006). *The Text Mining Handbook: Advanced, Approaches in Analyzing Unstructured Data*. Cambridge University Press., <URL: <https://dl.icdst.org/pdfs/files/25a6d982ee80e1db7a4ebf7eeca4e0ec.pdf>>.
- Liu B (2012). *Sentiment Analysis and Opinion Mining*. Claypool, Publishers. <URL: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>>.
- Silge J, Robinson D (2017). *Text mining with R: A tidy approach.*, O'Reilly Media, Inc. <URL: <https://www.tidytextmining.com/>>.

## Bøker (anbefalt lesing, tilgjengelig online)

1. Jurafsky D, Martin JH (2021). *Speech and Language Processing*. Online, draft. <URL: <https://web.stanford.edu/~jurafsky/slp3/>>.
2. Wickham H (2016). *ggplot2: elegant graphics for data analysis.*, Springer. <URL: <https://ggplot2-book.org/>>.

## Artikler

- Barnes J, Touileb S, Øvrelid L, Velldal E (2019). “Lexicon information, in neural sentiment analysis: a multi-task learning approach.” In, *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, 175-186. <URL: <https://aclanthology.org/W19-6119>>.
- Benoit K, Matsuo A (2020). *spacyr: Wrapper to the ‘spaCy’ ‘NLP,’ Library*. R package version 1.2.1, <URL: <https://CRAN.R-project.org/package=spacyr>>.
- Benoit K, Watanabe K, Nulty P, Obeng A, Wang H, Lauderdale B, Lowe W, (2017). *quanteda: Quantitative Analysis of Textual Data*. R package, version 0.9.9-65, <URL: <http://quanteda.io>>.
- Blei DM (2012). “Probabilistic Topic Models.” *Communications of the, ACM*, 55(4), 77 - 84.
- Cooksey B (2014). “An introduction to APIs.” *Zapier, Inc.*. <URL: [https://cdn.zapier.com/storage/learn\\_ebooks/e06a35cfcf092ec6dd22670383d9fd12.pdf](https://cdn.zapier.com/storage/learn_ebooks/e06a35cfcf092ec6dd22670383d9fd12.pdf)>.
- D’Orazio V, Landis ST, Palmer G, Schrod P (2014). “Separating the, Wheat from the Chaff: Applications of Automated Document Classification, Using Support Vector Machines.” *Political Analysis*, 22(2), 224 -, 242. doi: 10.1093/pan/mpt030 (URL: <https://doi.org/10.1093/pan/mpt030>).

---

\*solveig.bjorkholt@stv.uio.no

†martin.soyland@stv.uio.no

- Denny MJ, Spirling A (2018). “Text Preprocessing For Unsupervised, Learning: Why It Matters, When It Misleads, And What To Do About It.” *Political Analysis*, 26(2), 168-189. <URL:, <https://doi.org/10.1017/pan.2017.44>>.
- Finseraas H, Høyland B, Søyland MG (2021). “Climate politics in hard, times: How local economic shocks influence MPs attention to climate, change.” *European Journal of Political Research*, 60(3), 738-747., <URL:, <https://ejpr.onlinelibrary.wiley.com/doi/abs/10.1111/1475-6765.12415>>.
- Høyland B, Søyland M (2019). “Electoral Reform and Parliamentary, Debates.” *Legislative Studies Quarterly*, 44 (4), 593-615. doi:, 10.1111/lsq.12237 (URL: <https://doi.org/10.1111/lsq.12237>).
- Jørgensen F, Aasmoe T, Husevåg AR, Øvrelid L, Velldal E (2019). “NorNE:, Annotating Named Entities for Norwegian.” <URL:, <https://arxiv.org/abs/1911.12146>>.
- Lauderdale BE, Herzog A (2016). “Measuring Political Positions from, Legislative Speech.” *Political Analysis*, 24 (3), 374-394. doi:, 10.1093/pan/mpw017 (URL: <https://doi.org/10.1093/pan/mpw017>).
- Laver M, Benoit K, Garry J (2003). “Extracting Policy Positions from, Political Texts Using Words as Data.” *American Political Science, Review*, 97(02), 311-331.
- Lowe W (2017). “Understanding Wordscores.” *Political Analysis*., 16(4), 356-371. doi: 10.1093/pan/mpn004 (URL:, <https://doi.org/10.1093/pan/mpn004>).
- Lucas C, Nielsen RA, Roberts ME, Stewart BM, Storer A, Tingley D, (2015). “Computer-Assisted Text Analysis for Comparative Politics.” *Political Analysis*, 23(2), 254-277. doi: 10.1093/pan/mpu019 (URL:, <https://doi.org/10.1093/pan/mpu019>).
- Muchlinski D, Siroky D, He J, Kocher M (2016). “Comparing Random Forest, with Logistic Regression for Predicting Class-Imbalanced Civil War, Onset Data.” *Political Analysis*, 24(1), 87-103. doi:, 10.1093/pan/mpv024 (URL: <https://doi.org/10.1093/pan/mpv024>).
- Pang B, Lee L, others (2008). “Opinion mining and sentiment analysis.” *Foundations and Trends® in information retrieval*, 2(1-2), 1-135., <URL: <https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>>.
- Peterson A, Spirling A (2018). “Classification Accuracy as a, Substantive Quantity of Interest: Measuring Polarization in Westminster, Systems.” *Political Analysis*, 26(1).
- Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, Albertson B, Rand DG (2014). “Structural Topic Models for Open-Ended, Survey Responses.” *American Journal of Political Science*, 58(4), 1064-1082. ISSN 0092-5853.
- Slapin JB, Proksch S (2008). “A Scaling Model for Estimating, Time-Series Party Positions from Texts.” *American Journal of Political, Science*, 52(3), 705-722. ISSN 1540-5907.
- Stortinget (2022). *Stortingets datatjeneste*. <URL:, <https://data.stortinget.no>>.
- Søyland M (2022). *storningscrape: Scrape and structure raw data from, the Norwegian parliament’s API*. <URL:, <https://github.com/martigso/storningscrape>>.
- Wickham H (2020). *httr: Tools for Working with URLs and HTTP*. R, package version 1.4.2, <URL:, <https://cran.r-project.org/web/packages/httr/vignettes/quickstart.html>>.
- Wilkerson J, Casas A (2017). “Large-Scale Computerized Text Analysis in, Political Science: Opportunities and Challenges.” *Annual Review of, Political Science*, 20(1), 529-544. <URL:, <https://doi.org/10.1146/annurev-polisci-052615-025542>>.