Multivariate Behavioral Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

AI-Assisted Systematic Reviewing: Selecting Studies to Compare Bayesian Versus Frequentist SEM for Small Sample Sizes

Gerbrich Ferdinands

To cite this article: Gerbrich Ferdinands (2021) AI-Assisted Systematic Reviewing: Selecting Studies to Compare Bayesian Versus Frequentist SEM for Small Sample Sizes, Multivariate Behavioral Research, 56:1, 153-154, DOI: 10.1080/00273171.2020.1853501

To link to this article: https://doi.org/10.1080/00273171.2020.1853501

Published online: 02 Dec 2020.

MULTIVARIATE BEHAVIORAL RESEARCH
2021, VOL. 56, NO. 1, 153–154
https://doi.org/10.1080/00273171.2020.1853501

Routledge
Taylor & Francis Group

Check for updates

AI-Assisted Systematic Reviewing: Selecting Studies to Compare Bayesian Versus Frequentist SEM for Small Sample Sizes

Gerbrich Ferdinands

Department of Methodology and Statistics, Utrecht University, The Netherlands

Smid et al. (2020) conducted a systematic review to characterize the performance of Bayesian and frequentist estimation for SEM with small sample sizes. After manually screening 5050 studies, only 27 were selected to answer their research question.

Conducting a systematic review requires great screening effort. This screening effort makes synthesis of evidence an extremely challenging task. A potential reduction in workload is offered by an open source AI-aided screening tool: Active learning for Systematic Reviews (ASReview; van de Schoot et al., 2020). In ASReview, the researcher screens abstracts in interaction with an active learning model. Based on the researcher's decisions (*relevant* versus *irrelevant*), the model iteratively updates its relevancy predictions for the remaining abstracts. By prioritizing articles that are most likely to be relevant (i.e. certainty-based active learning) ASReview minimizes the number of articles to be screened by the researcher, while still identifying the majority of relevant articles.

The process of manually screening and automatically prioritizing publications leads to a set of relevant publications. As an illustration, ASReview was applied to the full set of 5050 studies identified by Smid et al. (2020). Desirable performance was defined as maximizing identification of the 27 relevant articles originally identified by Smid et al., while minimizing the number of articles to be screened by the researcher. The relevancy predictions were made by an active learning model using either naïve Bayes or logistic regression as the classifier. For the first prediction, ASReview requires a few example articles. ASReview was applied 27 times for each classifier, using every relevant article as an example article once, paired with one random irrelevant article.

As shown in Figure 1, both the Bayesian and logistic regression models found more than 80% of relevant publications after screening only 10% of all publications. They identified 95% of relevant publications after screening about 20% of all publications.

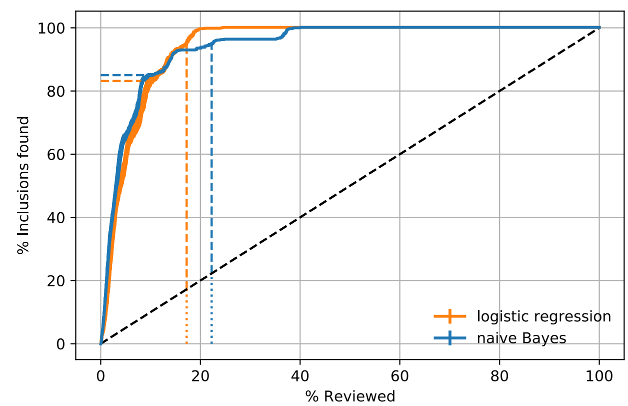

Figure 1. Recall curves for the search by Smid et al. (2020). *Note.* The curves indicate when the models find relevant articles during the simulated screening process. Example articles are subtracted so that the curve starts at (0, 0). The diagonal line indicates the rate of finding relevant articles when articles are screened at random. The horizontal lines indicate the proportion of relevant articles found after screening 10%. The vertical lines indicate when 95% of the relevant articles are identified.

In the current study, the two active learning models reduced the screening effort by 78–82%, depending on the classifier. In conclusion, relevant publications can be detected much earlier when an active learning model dictates the order in which articles are screened in systematic reviews.

Acknowledgments

I thank my SMEP sponsor, Rens van de Schoot and Sanne Smid for providing the data.

References

Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian versus frequentist estimation for

CONTACT Gerbrich Ferdinands g.ferdinands@uu.nl Department of Methodology and Statistics, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands.

structural equation models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 131–161. https://doi.org/10.1080/10705511.2019.1577140

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. (2020). *ASReview: Open source software for efficient and transparent active learning for systematic reviews*. http://arxiv.org/abs/2006.12166