Additional references for interpretable ML and causal ML

- Balzer, Laura B., and Maya L. Petersen. "Invited commentary: machine learning in causal inference—how do I love thee? let me count the ways." *American Journal of Epidemiology* 190.8 (2021): 1483-1487.
- Díaz, Iván. "Machine learning in the estimation of causal effects: targeted minimum loss-based estimation and double/debiased machine learning." *Biostatistics* 21.2 (2020): 353-358.
- Du, Mengnan, Ninghao Liu, and Xia Hu. "Techniques for interpretable machine learning." *Communications of the ACM* 63.1 (2019): 68-77.
- Doshi-Velez, Finale, and Been Kim. "Towards a rigorous science of interpretable machine learning." *arXiv preprint arXiv:1702.08608* (2017).
- Freiesleben, Timo, et al. "Scientific inference with interpretable machine learning: Analyzing models to learn about real-world phenomena." *Minds and Machines* 34.3 (2024): 32.
- Kaddour, Jean, et al. "Causal machine learning: A survey and open problems." *arXiv preprint arXiv:2206.15475* (2022).
- Li, Sheng, and Zhixuan Chu, eds. *Machine Learning for Causal Inference*. Springer, 2023.
- Lipton, Zachary C. "The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery." *Queue* 16.3 (2018): 31-57.
- McConnell, K. John, and Stephan Lindner. "Estimating treatment effects with machine learning." *Health services research* 54.6 (2019): 1273-1282.
- Moccia, Chiara, et al. "Machine learning in causal inference for epidemiology." *European Journal of Epidemiology* (2024): 1-12.
- Molnar, Christoph, Giuseppe Casalicchio, and Bernd Bischl. "Interpretable machine learning–a brief history, state-of-the-art and challenges." *Joint European conference on machine learning and knowledge discovery in databases*. Cham: Springer International Publishing, 2020.
- Murdoch, W. James, et al. "Definitions, methods, and applications in interpretable machine learning." *Proceedings of the National Academy of Sciences* 116.44 (2019): 22071-22080.
- Naimi, Ashley I., Alan E. Mishler, and Edward H. Kennedy. "Challenges in obtaining valid causal effect estimates with machine learning algorithms." *American Journal of Epidemiology* 192.9 (2023): 1536-1544.
- Rose, Sherri, and Dimitris Rizopoulos. "Machine learning for causal inference in biostatistics." *Biostatistics* 21.2 (2020): 336-338.

- Rudin, Cynthia. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." *Nature machine intelligence* 1.5 (2019): 206-215.
- Rudin, Cynthia, et al. "Interpretable machine learning: Fundamental principles and 10 grand challenges." *Statistic Surveys* 16 (2022): 1-85.
- Schölkopf, Bernhard. "Causality for machine learning." *Probabilistic and causal inference: The works of Judea Pearl*. 2022. 765-804.
- Tursunalieva, Ainura, et al. "Making sense of machine learning: a review of interpretation techniques and their applications." *Applied Sciences* 14.2 (2024): 496.