

Reading list

Session 7

Magnus Nielsen

Please note that you are **not** expected to have read the reading materials before the classes and should be able to follow along without having done so.

The readings for this session are in Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning: Limitations and Opportunities. [fairmlbook.org.](https://fairmlbook.org/), which is available online.

Read chapter 3, *Classification*, focusing on:

- Introduction
- Statistical non-discrimination criteria
- Independence
- Separation
- Sufficiency
- How to satisfy a non-discrimination criterion

Chapter 1, *Introduction*, and 2, *When is automated decision making legitimate*, are also worthwhile, but you should focus on Chapter 3.