

Reading list

Session 6

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.

To prioritize, I recommend you read the introductions, which give a short overview of the chapters and the methods. Otherwise, focusing on the intuition and broad strokes in the beginning of the sections is recommended.

The readings are all from Molnar, C. (2022). Interpretable machine learning, <https://christophm.github.io/interpretable-ml-book/>, which is available online

- Chapter 3, Interpretability
 - 3, Introduction
 - 3.1, Importance of interpretability
 - 3.3, Scope of interpretability
 - 3.6, Human-friendly explanations
- Chapter 5, Interpretable Models
 - 5, Introduction
 - 5.1, Linear Regression
 - 5.4, Decision Tree
- Chapter 8, Global Model-Agnostic Methods
 - 8, Introduction
 - 8.1, Partial Dependence Plot
 - 8.5, Permutation Feature Importance
- Chapter 9, Local Model-Agnostic Methods
 - 9, Introduction
 - 9.1, Individual Conditional Expectation (ICE)
 - 9.2, Local Surrogate (LIME)
 - 9.5, Shapley Values
 - 9.6, SHAP (SHapley Additive exPlanations)