

Explainable Machine Learning

Exercise 1

1. What is the meaning of explainable machine learning?
2. Why do we need to understand the decision-making mechanism of a machine learning algorithm?
3. What are the differences between explainable models and interpretable models? (give two examples for each type.)
4. What are the differences between global explainability methods and local explainability methods? (give two examples for each type.)
5. What are the differences between model-agnostic methods and model-specific methods? (give two examples for each type.)
6. What are the differences between Ad-hoc methods and post hoc methods?
7. What are the differences between data modality agnostic methods and data modality specific methods?

**** Please upload your solution to the course page in the ecampus within a week. if you have any questions, don't hesitate to get in touch with me at aemam@uni-bonn.de**