

Exercise 1: Quiz

- (a) Yes or No: Shapley values and SHAP values are different names for the same concept.
- (b) What are SHAP value functions v in contrast to SHAP values ϕ ?
- (c) Does the dependence structure in the DGP influence the SHAP result?
- (d) In what sense are local, post-hoc interpretation techniques helpful in auditing AI systems?
- (e) In which way does the SHAP feature importance differ from the permutation feature importance?
- (f) In which way does the SHAP dependency plot differ from the ICE plot?

Exercise 2: Application

Make yourself familiar with the method Shapley in the iml package. Answer the following questions for your chosen dataset

- (a) Choose 3-4 observations of your dataset which might be worth to explain via Shapley values. Explain how you chose these observations.
- (b) Calculate and visualize the Shapley values for these observations and interpret your results.
- (c) Sample 100 observations from your dataset at hand and calculate the Shapley values for these observations
 - Calculate the Shapley feature importance for your sample. How does it differ from the PFI?
 - Create a Shapley dependency plot for the 6 features you have chosen in the feature effects exercise and compare your plots with the PDPs.