Explainable Artificial Intelligence   **Paderborn University**
Summer Semester 2025   Dr. Stefan Heindorf

Exercise Sheet 4

Due date   June 16, 2025

1 Task: LIME

a. Apply LIME[1] to the bike rental dataset (e.g., using the preprocessed dataset and the random forest regression model from the previous exercise sheet). Experiment with different models (e.g., multilayer perceptron vs. random forest) and with different kernel widths. How do the explanations differ?

b. Train a classifier (instead of a regression model) on the bike rental dataset (e.g., a random forest classifier predicting whether more or less than 4.500 bikes will be rented on a given day). Pick two examples from the dataset and use LIME to explain the predictions for the two examples.

2 Task: Shapley Values

a. For a set of d players in a cooperative game with characteristic function $val : 2^d \mapsto \mathbb{R}$, the Shapley value of player $j \in \{1, \ldots, p\}$ is defined as

$$\phi_j(val) := \sum_{S \subseteq \{1, \ldots, p\} \setminus \{j\}} \frac{|S|!(p - |S| - 1)!}{p!} (val(S \cup \{j\} - val(S)))$$

Prove the efficiency property of Shapley values, i.e., prove

$$\sum_{j=1}^{p} \phi_j(val) = val(\{1, \ldots, p\}).$$

Hint: Recall how the definition of the Shapley values was derived via permutations in the lecture. For each permutation, compute the sum of the marginal contributions of all players.

b. Train a random forest on the bike rental dataset and apply an explainer from the SHAP library to it (e.g., exact, sampling, kernel, tree). Output the Shapley values in form of a table and experiment with different visualizations.[2]

---

[1] https://github.com/marcotcr/lime
[2] https://shap.readthedocs.io/en/latest/api.html#plots

c. In general, what are the similarities and differences between the four different explainers (exact,[3] sampling,[4] kernel,[5] tree[6])? Consult the SHAP documentation and/or perform a web search. Provide sources for your findings.

---

[3] `https://shap.readthedocs.io/en/latest/example\_notebooks/api\_examples/explainers/Exact.html`

[4] `https://shap.readthedocs.io/en/latest/generated/shap.explainers.Sampling.html`

[5] `https://shap.readthedocs.io/en/latest/example\_notebooks/tabular\_examples/model\_agnostic/Simple\%20Kernel\%20SHAP.html`

[6] `https://shap.readthedocs.io/en/latest/generated/shap.explainers.Tree.html`