

Exercise Sheet 1

Due date April 28, 2025

1 Task: Importance of Explainable Artificial Intelligence

- a. What is explainable artificial intelligence (in general)?
- b. What are potential problems of black box machine learning models (in general)?
- c. What are applications of explainable artificial intelligence? Perform a web search and identify three *concrete* applications that were not already mentioned in the lecture and describe each application in a paragraph. For each application:
 - a) Provide a source (e.g., a website, blog post, scientific paper)
 - b) Explain how this application benefits from explainable AI: Who uses the system? Why is the user interested in explanations? What do explanations look like?

2 Task: Implement Linear Regression from Scratch

Write a Python class `MyLinearRegression` that implements linear regression from scratch with basic linear algebra operations from the numpy library without using a machine learning library.

- a. Your class should have the two methods:

- `fit(X_train, y_train)`
- `predict(X_test)`

The former trains the model and the latter makes predictions. `X_train/X_test` shall be numpy arrays of shape `(n_samples, n_features)` where `n_samples` can differ from the training to the test set, but `n_features` stays the same. Provide the code of your solution.

- b. Use the California Housing dataset from scikit learn¹ and split it into two parts: one training set and one testing set (recommendation: use 80% for training and 20% for testing)
- c. Train your linear regression model on the training dataset and make predictions both on the training and the test dataset. Report the mean squared error for the training set and for the test set.

¹https://scikit-learn.org/stable/modules/generated/sklearn.datasets.fetch_california_housing.html

- d. For comparison, apply the `LinearRegression` model from Scikit-Learn to the same dataset.²
- e. Compare whether you get the same results. If not, try to fix errors in your own implementation or try to adjust the settings of the Scikit-learn model (e.g., disable regularization, ...)

Helpful Numpy resources:

- Official tutorial: <https://numpy.org/devdocs/user/quickstart.html>
- Numpy tutorial on W3Schools: https://www.w3schools.com/python/numpy_intro.asp
- Advanced Numpy (exploring the internals): http://scipy-lectures.org/advanced/advanced_numpy/index.html

²https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html#sklearn.linear_model.LinearRegression