Methods of machine learning

Exercise sheet V

June 11st, 2025

In this exercise class we apply PCA to the classical *digits dataset.*

1. Task.

a) Load the data, display the first digit in the dataset and use `StandardScaler` function from `sklearn.preprocessing` to standardize the data.

b) Apply a PCA model to the standardized dataset and plot the first four principal components.

c) Plot the individual explained variances and the cumulative explained variances of the principal components.

2. Task.

a) Create a PCA model with $n = 20$ components, apply it to the digits dataset, and reconstruct the transformed digits.

b) Compare the original digits and the reconstructed digits.

c) Calculate the mean square error (MSE) between the original images and the reconstructed images with the increasing number of principal components and plot the MSE with the increasing number of principal components.