

BM20A6100 Advanced Data Analysis and Machine Learning

Exercise 1: Nonlinear Dimensionality Reduction

Deadline: 03.11.2025

Submitting the solutions

Create a new GitHub repository (or use an existing one if you have already made one for this course) and return the solutions by adding a link to the GitHub repository in Moodle. Make sure to name the files clearly.

You can either create a report PDF or utilize Markdown comment sections in a Jupyter notebook. Comments within the code are insufficient to act as answers to the questions.

1. Comparing linear and non-linear DR (4 points)

Compare PCA and t-SNE methods by visualizing Bike Sharing Rental dataset. Explore how the different features are shown in the DR components. Build a simple prediction model (for example, MLP or Random Forest) to predict the count of total rental bikes and compare the performance of the model with the different DR techniques.

2. Visualizing with SOM (3 points)

Visualize MNIST-784 handwritten digits dataset with SOM and discuss what you can learn from the visualization.