Computer Convergence Application
Homework 1

…

Due date: **4/14 23:59**

Late penalty: 10%p deduction on each day passed

How to submit:

- Mail to stj@snu.ac.kr
- Subject should begin with [CCA-HW1]

What to submit:

- Runnable python code either in Colab address, .ipynb file, or .py file
- Report should not exceed 2 pages

…

1. Load breast_cancer datasets from dataset module in sklearn library. What are the number of samples and the number of features?

2. Split the dataset into training and test sets. 20 samples are designated at test samples. Train LASSO regression model with training data. Evaluate the model with test data, using mean squared error as an evaluation metric.

3. Train and test ridge regression model under same condition. Compare mean squared errors.

4. Draw barplots for coefficients of variables in 2 and 3.

5. Split the features into half to make two datasets for CCA. Run CCA on these datasets with (n_comps=2). Plot CCA results of two datasets.

6. Draw t-SNE plot using the whole breast_cancer dataset. How are malignant and benign classes clustered?