**HIGH-DIMENTIONAL ANALYSIS**

**---**

**Programming Exercise – 02**

Deadline: 23h59 22/05/2022

Submitting via email: [dxtien95@gmail.com](mailto:dxtien95@gmail.com)

**PROBLEM:**

Apply Linear Discriminant Analysis algorithm on the Iris flower dataset. Make a comparison between LDA and PCA projected samples.

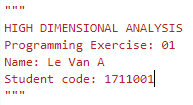
**Note:** Do LDA step by step, available tool for PCA is accepted.

Dataset description:

* The famous “Iris” dataset that has been deposited on the UCI machine learning repository (<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data>).
* It contains measurements for 150 iris flowers from three different species (*setosa, versicolor* and *virginica*).
* The four features of the dataset are *sepal length, sepal width, petal length* and *petal width*

After finishing your task, please write a short report or a summary (pdf file) to explain your answers, ideas and the way your code works.

**NOTICE**

1. Please send the two files (coding and report file) before the due date. Or send the jupyter-notebook file (ipynb, html, pdf) or google-colab link.
2. The mail subject and the folder’s name would be *[HDA2020\_PEXX\_Name\_StudentCode]*, where PE means Programming Exercise. For example: *HDA2020\_PE01\_Le\_Van\_A\_1711001* or *HDA2020\_PE01\_LeVanA\_1711001*.
3. Inside the coding file, there should be a brief introduction (as example below). 
4. There is **NO** acceptance for **cheating** or **copying**.

**TUTORIAL**

Export html file from jupyter-notebook

