CA 5314: Lab Assignment 0

Exercise 1

- Create an ndarray x of shape=(5,4) with random numbers
    - Each of the 5 rows represents a sample with 4 features
- Create a flat ndarray y of shape=(5,), whose elements are 0 and 1
    - Each element is the class label of the corresponding sample in x

Hint: use **`np.random.rand(...)`** and **`np.random.randint(...)`**

Exercise 2

- Define a function **`extract_subset(x, y, y0)`** that takes as input:
    - the feature matrix x, and the labels y from the previous exercise
    - a target class y0 **`(i.e., either y0=0 or y0=1)`**
- and returns a feature matrix containing only samples belonging to y0

Hint: use array indexing with y==y0

Exercise 3

- Define a function **`min_feature_values(x)`**that returns the minimum value of each feature in x
- Apply it on the previously extracted samples x0 of class 0

Hint: use `np.min()` with a proper axis value

Exercise 4

- Define a function **`make_gaussian_dataset (n0, n1, mu0, mu1)`**
- that generates a two-class Gaussian dataset in a bi-dimensional space
    - n0, n1 are the number of samples for class 0 and class 1
    - mu0, mu1 are the means of the two Gaussians (one per class)
- We consider only Gaussian distributions with covariance equal to the identity matrix here for simplicity
- The function returns the corresponding feature matrix and labels x, y

Hints:

- use **`np.random.randn(...)`** to generate random numbers from a standard Gaussian distribution, with zero mean and unit variance, and then transform them to have a different mean (repeat twice, one per class)
- use **`np.ones(...)`** and/or **`np.zeros(...)`** for class labels
- use **`np.vstack(...)`** and **`np.hstack(...)`** to concatenate arrays