The goal of this assignment is to implement the first supervised strategy of the class: Decision Trees. You will be applying this technique to different datasets and get a "feel" on how it works. You can work individually or discuss your work with a classmate. The submission must be individual and each student must work on her/his own datasets. If you work collaboratively please indicate the name of your classmate in your submission.

1. Identify three (3) datasets that you will work with. The conditions for the datasets are:
    1. The datasets have to contain at least 3,000 entries
    2. The datasets have to have at least 5 features/columns, aside from the label/class
2. Make a post under the corresponding discussion for repositories and include each dataset that you will work on.
3. For each dataset-algorithm run a full implementation:
    1. You can look for existing Python libraries and code examples as your baseline for the code. (make sure to include references to the original code)
    2. Adapt your baseline code to be able to load your dataset and print results and statistics
    3. For each implementation you should do the training and the testing phases. Document the code to explain how much data you use for training and for testing (in %s) and the criteria for dividing the dataset into these subsets.
4. Document your observations and any issues you faced while trying to implement the code.

Submission format:

a. Upload a text document in this folder containing:

1. Links to your posts for the datasets (1 pt)
2. Observations from the full run over the three datasets (1 pts per set)

b. Upload an independent Jupyter notebook for the code and output of each run/dataset (2 pt per set)