Exercise 1 - deadline 19.03.2021 23:59

Build and use a classifier to predict the income for a person based on census data.
Data can be found here:
https://archive.ics.uci.edu/ml/machine-learning-databases/adult/
You have both the training dataset and the testing dataset. Do not use the test dataset in any way to build the classifier.

The goals of the exercise:

1. Build your first machine learning exercise
2. Understand how to load data, understand the features, type and role of attributes.
3. Understand what missing values are, and simple ways to handle them (ignoring them).
4. Basic preparation of data.
5. Create a classification model, using two techniques. (e.g. decision trees and naive bayes)
6. Evaluate a model using the test data.
7. Use the model to predict the class for new data.

The final submission should contain (at least) one script to contain all the goals mentioned before.

Optional goals:

1. Improve the efficiency of the chosen model. Describe the steps you did and the improvement.
2. Save the model and load it in a different script where you use the model to make predictions. In other words, separate the model creation script from the model utilisation script.

Important Note:
The submission should be done individually. Please name your script/notebook accordingly. At the beginning of your script / notebook give the name(s) of all students who worked on the exercise. You may work in small groups (of 2-3) but all students must be confident with the work of others. They should understand and be able to answer detailed questions of the implementation.