**Exercises Week 1**

1. Install Python and Jupyter notebook on your computer. You will also need a few standard python libraries such as pandas, scipy, numpy.
2. Using the pandas library load the data from ‘mixture.csv’ (from the moodle page). The simplest function to do that is ‘read\_csv’ implemented in pandas. This creates a *DataFrame* (you can think of this as an Excel Sheet with column names and row names)*.* This is the training data shown in the Least Squares and Nearest Neighbor lectures and obtained from the *Elements of Statistical Learning* book (<https://web.stanford.edu/~hastie/ElemStatLearn/>).

Use the ‘scatterplot’ function built in pandas to visualize the data, use different colors for the two classes.

1. Program a function NearestNeighbor in python. This function should take as input the training data (i.e., the DataFrame containing the data from mixture.csv) and provide the class prediction for a given X1 and X2 coordinate based on the single Nearest Neighbor (1-NN) using Euclidean distance.
2. Use your NearestNeighbor function to make predictions for all entries in mixture\_test.csv and plot the result (again using the scatterplot function).
3. Challenge: implement a NearestNeighbor function that takes the desired neighborhood size as an input variable. Use it to make predictions based on the 15 Nearest Neighbors (15-NN) for mixture\_test.csv and plot the result.