# Lab 2a: Heart Disease / Orange

Use “Orange” in Anaconda to follow these steps. Insert the five screenshots at the marked positions. Upload the Word file with your screenshots to eCampus (preferably saved as PDF).

**Your name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. Load "Heart Disease" dataset. This dataset comes with Orange; you can use the “Datasets” widget to easily access it.
2. Check its data using a "Data Table" widget.
3. Draw a Scatter Plot. Configure the plot to find & show the most informative projection. Insert a screenshot of the scatter plot window showing the most informative projection:

[Insert Screenshot]

1. Use the "Test & Score" widget. Drag the "Datasets" output to the "Test & Score" input.
2. Create the "Tree", "Random Forest", "kNN", "SVM" and "Neural Network" classifiers and connect all their outputs to the input of "Test & Score"
3. Check the "Test & Score" results. Configure it to use 20 folds for stratified cross validation. Order the classifiers by their precision. Which classifier has the highest precision? Paste a screenshot.

[Insert Screenshot]

1. Connect the "Test & Score" output to a "Confusion Matrix". Take a screenshot of the confusion matrix of the classifier with the best precision.

[Insert Screenshot]

1. Use “Datasets” as input for a second “Tree (1)” classifier. Connect the new classifier to a “Tree Viewer” widget to visualize the decision tree. Configure the decision tree to use the following configuration, and then insert a screenshot of the resulting tree:
   1. Binary tree
   2. Min number of instances in leaves: 2
   3. Do not split subsets smaller than: 5
   4. Limit the maximum tree depth to: 3

[Insert Screenshot]

1. Take a screenshot of your completed canvas in orange that shows how you connected the widgets.

[Insert Screenshot]