Introduction to Intelligent Systems '18/'19
Lab 4

Instructions

You should hand in a brief report, in which you provide answers to all the problems, with the same structure of the present document (Assignment 1, Assignment 2, etc.). The code developed to solve the problems has to be included in the report, commented, and opportunely referenced.

The plots have to be self-consistent, with clear title, labels and captions. They have to be sufficiently commented in the text.

The grading will be based on the following aspects (ordered acc. to importance):

- completeness of the report and correctness/plausibility of the results
- readability, layout, formatting of the text
- appearance of the plot, e.g. axis labels, quality of the caption
- language/English issues

A grade equal to 10 can be achieved by solving all the problems correctly and matching the above conditions.

Assignment 1

Decision Trees.
A marine biologist gives you the following descriptions of various whale species:

Killer whale The fluke of this relatively small (6-8m) whale is not visible when it dives, but its tall and pointed dorsal fin is often clearly visible. You can also see the whale blow water quite often.

Beluga whale This whale can be difficult to spot as it does not show its fluke when diving, and does not have a dorsal fin.

Narwhal whale These very small whales usually do not grow above 5 meters and are known for their single, extraordinarily long tusk. Their fluke is clearly visible when they dive, and they do not have a dorsal fin.

Bowhead whale Much like the Narwhal, this whales fluke is visible when it dives, and it does not have a dorsal fin to show off. It is however, a lot larger, reaching sizes up to 20 meters.

Blue whale This whale is believed to be the largest animal ever to have existed. Growing over 30 meters long, its impressive fluke can be seen clearly when it dives. Its dorsal fin, although relatively small, is also often clearly visible.

a) Design a classification system based on a binary decision tree that can be used by whale spotters. Try to keep the queries as simple as possible, minimize the height of the tree, and explain why you chose this particular tree. Provide clear motivation of your design choices.

Assignment 2

Hierarchical clustering

The file dataAEX.mat contains 19 time series of 19 stocks whose names are specified in labelsAEX.mat. You can load the data of these files in your program using *importdata* or *load*. Create a dendrogram of the 19 stocks using the Matlab functions linkage and dendrogram.
Apply the complete linkage algorithm, using Euclidean distance between the time series. What considerations can you make by observing the dendrogram?