Assignment 6: Cross Validation

Implement 5-fold cross validation for LVQ1 training using the data set of the last assignment (LVQ1).

First, shuffle the data set randomly once, then split the data set into 5 disjoint subsets, each containing 20% of the data. Take care that the correct labels are still assigned to the data points.

Perform LVQ1 training for up to, say, 100 epochs in each fold and determine the training error and the test error at the end of each training process. Initialize each prototype in a randomly selected point of its class. Compute the average errors and their standard deviations over the 5 folds. Plot the results as a function of K (a useful command: **errorbar(....)**).

You should hand in, at least:

- your matlab code
- a structured report with a short introduction to the problem and a brief description of your computer experiments
- the results for K=1,2,3,4,5 prototypes per class (plots as described above)
- a brief discussion of your results

Further suggestions (bonus):

- consider more (even larger) values of K
- consider class 1 and class 2 errors separately
- repeat the 5-fold cross validation for several randomized splits into subsets, perform an additional average over these random realizations
- perform n-fold cross-validation for several values of n and compare the results.