

EEL 6814

HMW #4

Due March 29, 2016

The goal of this homework is to find structure in clusters of two species (identified as 1 or 0) of rock crabs using anatomic measurements of front lip, rear width, length, width, and depth. Males and females have distinctive measurements, providing some extra structure to the data. The data set is composed of 200 specimens, 50 males and 50 females from the two species. The main goal is to use unsupervised methods to find structure in the input space, therefore, below you can only use the input measurements WITHOUT using the desired response labels.

Method 1.

Use projections to 2D subspaces using different PCA projections to cluster the input data. Design a simple vector quantizer (experiment with the number of clusters) to compare the performance of clustering on the PCA subspace with the MLP classification.

Method 2

Use a Self-Organizing Map (SOM) to cluster the data directly to a 2D output space. Experiment with the number of clusters. Since you have the desired response you can group the clusters to differentiate between the males and female. Evaluate also the performance of the SOM clustering versus the MLP classifier.

Method 3

Use a one hidden layer autoencoder (bottleneck layer with 2 PEs) with nonlinear PEs and compare performance with methods 1 and 2. Repeat with a 3 hidden layer of nonlinear PEs.

The class labels are also given for you to design a classifier and see how much the availability of the desired response helps in the definition of the class boundaries.