

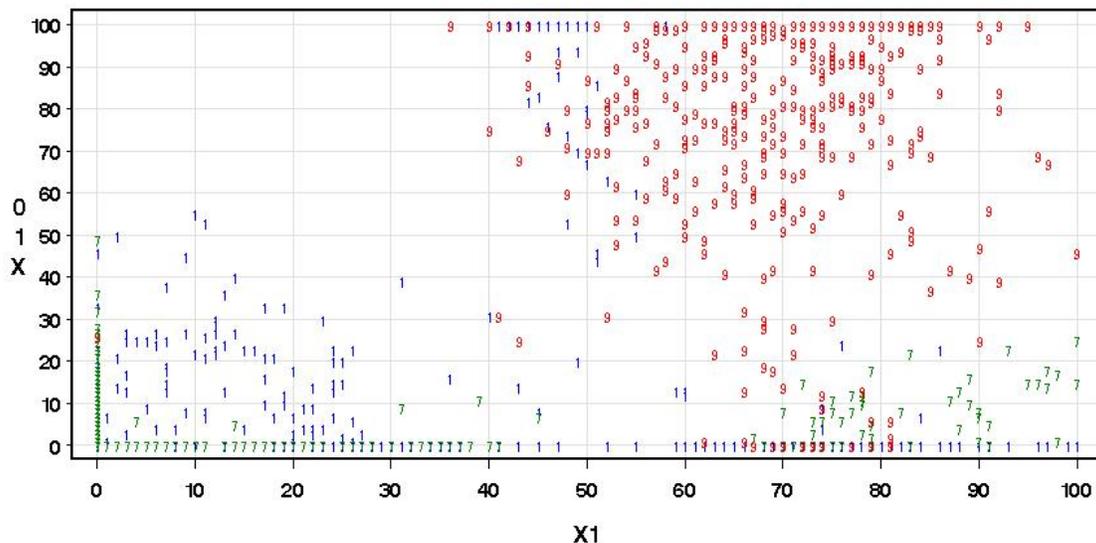
STA5703 Assignment #5

Handwriting recognition is a classic pattern recognition task. The cases are digits written on a pressure-sensitive tablet. The inputs are measures of position of the pen scaled to be between 0 and 100.

The exercise data is a subset of the pen-based recognition of handwritten digits data set available from the UCI repository (Blake et al. 1998). It only uses two of the 16 original inputs (x_1 and x_{10}) so that the decision boundaries can be visualized. The target is the true written digit 0, 1, ..., 9.

This exercise uses only a subset of data, that is, the digits 1, 7, and 9. In all, 1064 cases were extracted from the original test data set. The sample is roughly equally allocated across the three digit classes. It is further assumed that the proportion of cases from each class in the sample equals that of the intended population. If this is not the case, and if the posterior probability is required, the probabilities and the decision rule will need to be adjusted (using prior probabilities).

Figure 1 Hand Writing Data



Problem 1 (3 Point) NEURAL Network node can be used to fit a generalized linear model model. What are the settings to perform this task? Report the accuracy rate and the confusion matrix?

Problem 2 (2 Point) Change the setting to be MLP with 8 hidden units with adequate number of iterations. Report the accuracy rate and the confusion matrix?

STA5703 Assignment #5

Problem 3 (2 Point) Change the setting to unconstrained ordinary radial basis function (ORBFUN) with 8 hidden units and adequate number of iterations. Report the accuracy rate and the confusion matrix?

Problem 4 (5 Point) Perform unconstrained normalized radial basis function (NRBFUN) and then complete the following table.

Algorithm	Accuracy	Rank
<i>NRBF – Unconstrained</i>		
<i>NRBF – Equal Heights and Widths</i>		
<i>NRBF – Equal Heights</i>		
<i>NRBF – Equal Widths</i>		
<i>NRBF – Equal Volumes</i>		