HW5 (ECE7252 Spring 2008)

1. Based on the prostate cancer data we did in HW3, let's design a nonlinear neural network to fit the data (remember the universal approximator we discussed in class). In this case you have 67 cases for training (not a lot) and 30 samples for evaluating the goodness-of-fit. Use RSS as a way to measure the fit in training and testing. Conduct at least the following analysis (you need at least five different choices ranging from a small number like 10, to a large number, like 100):

1) Fix the number of hidden units.
2) Pick a step size for backpropagation
3) Train a feedforward 3-layer perceptron (you can write your own code or get a tool from the web) by observing the RSS on the training data as a function of the number of iterations. Plot the behavior on a curve and see if the training process converges. Use different step sizes, can you see different convergence behavior?
4) Test the networks you obtained above on the 37 test cases, and compare them with the results you obtained in HW3 when linear regression methods were used.
5) Repeat steps 1) to 4) with different hidden units for the five sizes you picked.
6) Compare RSS for the training and test sets when different hidden units are used.

2. The MLP can also be used to design nonlinear classifiers. Use the zip code data (about 7000 training token) to design ANNs for handwritten digit classification. You can use RSS as a way to measure the fit in training and use error rates to measure performance in training and testing. Conduct at least the following analysis (you need at least five different choices ranging from a small number like 100, to a large number, like 500):

1) Fix the number of hidden units.
2) Pick a step size for backpropagation
3) Train a feedforward 3-layer perceptron by observing the RSS on the training data as a function of the number of iterations. Plot the behavior on a curve and see if the training process converges. Use different step sizes, can you see different convergence behavior?
4) Test the networks you obtained above on the test samples, and compare the error rates obtained on training and testing data. Are they very different? Your error rates can be listed on a per digit basis and the overall average error rates.
5) Repeat steps 1) to 4) with different hidden units for the five sizes you picked.
6) Compare errors for the training and test sets when different hidden units are used.