Final Project

In general, it is a good idea to keep a journal of what you do when experimenting with data (and the results), so you can go back to it later without having to re-run everything.

For these projects, use script files and plots to explain your answers- You can write results up by hand. Include enough information so that someone versed in neural nets can understand what you did, and can reproduce your results.

1 Project I

The goal: Given a collection of photographs of people, construct a classifier based on gender. Here are some specifics:

1. Separate the data into a "Training" set and a "Validation" set. For the training set, use 7 (out of 13) randomly chosen photos for each gender for a total of 14 in the training set.

2. Method 1: Subspace classifier. For this classifier, construct a best basis for each gender. The gender of a new face is defined to be the subspace that best matches the given face. You should run your algorithm several times using different choices for training, and then settle on the best set. For each set, use as few basis vectors as possible.

3. Method 2: RBF Classifier. In this case, we will input a face, and output the gender. Write the RBF script yourself using two centers and Gaussian transfer function. You may run this several times to see if you can find a best width for the two Gaussians, and also with different training sets to see if this value changes. Compare this output to Matlab's built-in RBF routine (it will use Orthogonal Least Squares).

4. What to turn in: Matlab files, plots and a summary of your conclusions (this can be handwritten)- You might keep a "journal" of the experiments and the results of what you try. You can decide what plots you want to use in the explanation.

2 Project 2

This is another classification task- Use the data in `heart1.mat`. This data was taken from cardiology patients. The first column of the input matrix is age at heart attack, the second column is a binary variable representing whether or not there was fluid around the heart. The third variable is a measure of contractility around the heart (lower numbers are increasingly abnormal). Numbers in columns 4-6 are also measures of contractility from the echocardiogram, the general idea being that large hearts are sick hearts.

Construct a feed forward neural network to classify patients. The target values are "Alive" or "Dead" (taken after a fixed time length of time). Use the smallest number of neurons you can- 12 neurons in the hidden layer will give good results.

Note that the neural network will produce real numbers. To see how the network is performing, get the output and plot it against the actual targets. You want to see a good separation of the data.

To examine how the network is performing, also inspect the weights of the neural network. With 12 neurons and 6 input dimensions, the matrix will be 12 x 6, and you can see it with the command `net.IW{1}` Comment on which variables seem to be standing out.

What to turn in: Matlab files, plots and a summary of your conclusions (this can be handwritten). You can decide what plots you want to use in the explanation.

Side Remark: Although this is a little morbid, Neural Network classifiers such as the one you're building are being widely used in medical questions such as this, and have been implemented in hardware chips and integrated in some medical equipment!