CSI 635 Assignment #1: Supervised Learning

Due: 03/08/02
Marks: 17% of final grade

Implement a fully connected, feed forward neural network that uses SIGMOID units. The model should be learnt using a stochastic gradient descent backpropagation scheme. The neural network will need to perform both TRAIN and TEST functions. Limit the number of training epochs to stop over-fitting.

Testing Functionality

1. Show that your network is capable of learning the following functions: AND, OR, XOR. Present the network topology and final network weights.

Questions 2 through 7 refer to the recognition of hand-written characters data set available on the course web page on 02/22/02.

Evaluating Performance

2. Divide the training data into ten folds and perform 10-fold cross validation. This involves training the neural network on nine of the folds and using the remaining fold for testing. This is repeated so that each fold is used as a test set.
3. Describe the network architecture and why it was chosen.
4. List the training and test accuracies for each of the ten experiments and report your approximation of the learner bias, variance and MSE.
5. Try your neural network on the test set available on the course home page. For each instance, report the prediction of your neural network.

Understanding a Neural Network Model

5. For each of the ten folds, use C5.0 or C4.5 to train a model that mimics the Neural Network model. Note this involves training the neural network, then building a decision tree that predicts the neural network decision, not the dependent attribute.
6. For the C4.5/C5.0 experiments list the training set accuracy and test set accuracy for each fold. For the test set accuracy report the accuracy of predicting the neural network prediction and the dependent attribute value. Are the decision tree models fairly similar or different?

Hidden Layer Interpretation (Bonus Questions)

7. For each hidden layer node try to determine what latent concept (if any) it represents. Justify your conclusions with empirical analysis.

For the interested student

Try using TANH units
Experiment with differing topologies with respect to:
- Number of hidden layer nodes and hidden layers.
- Fully versus partial connection of neurons.
- Allowing neurons to connect to other neurons on the same level.
- Ordering of the training instances. What changes effect the overall accuracy?