Neural Networks, Fall 2004
G.G.Lendaris

PRELIMINARY INFORMATION re. PATTERN RECOGNITION COURSE PROJECT and start-up assignment.

PART A of Course Project: November 10, 2004 DUE: Wed., November 17, 2004

This exercise is intended as start-up preparation for a larger-scale set of experiments to be carried out over the remaining weeks of the term after this preliminary set is due in.

The data is related to automatic land-use classification via photographic imagery of Phoenix, Arizona, taken by NASA from the satellite called SKYLAB II in 1974. There are two pairs of.nna files: each pair contains a data set for Training and a data set for Testing (Generalization), and each pair provides a different number of inputs for the neural net: 194 and 54. In the 194 input case, the data represent "raw measurements" from a preprocessor; in the 54 input case, the data represent "features" that were extracted from the 194 measurements. There are 292 total events in each set, approximately 150 are in the Training set, and the remainder are in the Test (Generalize) set.

A number of experiments have previously been carried out with this data with neural nets of the "backpropagation" variety. Some results will be described in class.

For your "course project" each of you is to carry out some experiments with this data on another of the neural network paradigms [see bottom of this sheet].

For this week, however, your task is to *familiarize yourself with the data set* and various of the NeuralWorks processes you will need to carry out your experiments. Data available on Web site.

Create a multi-layer, feedforward network for the **54 input data set**; the size of the input layer will be 54 elements [NOTE: ALLWAYS STUDY YOUR .nna FILES. IN THIS CASE, FIELDS 1, 56, 62 AND 63 CONTAIN IDENDIFICATION DATA; TELL NWORKS TO START LOOKING IN FIELD 2 FOR THE (54) INPUTS AND FIELD NUMBER 57 FOR THE (5) OUTPUTS]; a hidden layer of 8 elements, and an output layer of 5 elements. Such nets have been successfully trained (using the Backpropagation algorithm) to over 90% correct responses on both the train and generalize data sets, with the nets being trained from 6,000 to 30,000 pattern presentations (N = 6,000; N = 30,000). See READ.ME file that is with the data.

Take "snapshots" of the runs at appropriate intervals (use NWorks *checkpoints* in the Run menu), and after the run is completed, go through and do a recall process on each of the snapshots. Use whatever tools are at your disposal (e.g., within NWorks or outside of NWorks: EXCEL, MATLAB, etc.) to determine the performance of the NN at various stages during the learning process (each of the partially trained NNs was saved as a sequence of .nnd files by the snapshots/checkpoints process-- make sure you understand the naming convention). Create plots of the *learning dynamics* for each experiment. Show comparative results for different experimental conditions. NOTE: when you do the recalls, if you instruct NeuralWorks to APPEND the results in the .nnr file, all the results are sequentially put into the same results file. The alternative is an OVERWRITE option, and if you weren't aware of this, there would be some surprises.

Additional Assignment: familiarize yourself [via lecture material, Texts 1 & 2, and any other source you wish to consult] with the following paradigms sufficiently to allow yourself to bring a list of 3 choices with you to class on Wednesday, Nov. 17th, and I will make assignments accordingly: LVQ, SOM, RBF, PNN, GRNN, Counterprop, Fuzzy Artmap -- part of your PROJECT will be to learn enough about the selected paradigm to carry out meaningful experiments.