Neural Networks I, Winter, 2010
G.G.Lendaris

PATTERN RECOGNITION PROJECT for rest of term.

Course Project part b: Wed., February 24, 2010 DUE: Wed., March 17, 2010, by 5:00pm.

As I indicated in "PRELIMINARY INFORMATION re. COURSE PROJECT" assignment sheet, each of you is to carry out a set of experiments with the SKYLAB Phoenix data using one of the various paradigms available in the NeuralWorks simulator, except the Backpropagation paradigm. It was suggested that you bring a list of 3 choices with you to class on **Wednesday, February 24th,** and I will make assignments accordingly.

Your first important "task" in this assignment is to learn a sufficient amount about your selected/ assigned NN paradigm, so you can formulate some "intelligent" experiments. Your next set of tasks are to define the problem appropriately for the given paradigm, and do training runs to (hopefully) achieve successful classification (5 land-use classes; the *desired-output* part of the .nna file is annotated) -- in the high 90%'s

Following this, assume the perspective of a researcher who has good reason to believe that the photographic images can be classified reasonably well via the **52 feature set**, and you want to ***discover*** a minimal set that can do a good (high 90's percent) classification. Also, you want to develop as much personal insight as you can about relationships among the features. The expected approach is to start with the (successful) network you developed for the 52 features, and study this NN for hints about which of the 52 features are used the most, and/or in ways you can conjecture seem important, etc... Then, create a new NN configuration which uses the selected features... Get down to at least a "best 15" set of features. It is known that a "best 9" is capable of yielding over 90%a generalization performance. Perhaps you can get it down to even fewer?

ASSIGNMENT: Design and execute experiments to accomplish all of the above.

Keep a good log book in which you document your ***thought process***, including speculations on good experiments to run (by definition, this means you DO THIS DOCUMENTATION AS YOU GO ALONG; waiting until the end to enter this information is the wrong approach!). Choose experiments and run them. Print out appropriate documentation for at least the starting and final networks. Test your nets with the train and generalize data sets, and "grade" the results for both data sets. Note that you can get NeuralWorks to give you evaluation results in its "error rate" and "confusion" and/or "classification" tables. Make sure you learn the details of what each of these tables does, and use as you desire. Alternatively, you may wish to export the .nnr results file to some other environment (e.g., spreadsheet program, MATLAB, MathWorks, a program you write, etc.) and develop evaluation and presentation format(s) of your choosing. Develop a way to show learning dynamics, as appropriate. Include print outs of the .nnr files for your *final* design, and be sure they are well annotated [you can use an editor to add information to the header prior to printing it out, or you may (neatly) add in information by hand] -- depending on the form your report takes, these may be exhibited in an Appendix.

SUBMIT a "RESEARCH REPORT", in which you describe the project, including your objectives, plan of attack, results, insights, etc., as well as a good conclusion section. Submit your **Log Book** as a *separate addendum* to the Research Report.

NOTE TO GRADUATE STUDENTS:

Your submission will be evaluated from a higher level of expectation (re. your methodology, the experiments you run, the analysis you perform, and the quality of the report itself) than for the undergrad students.

NOTE TO THE UNDERGRADUATE STUDENTS:

The above is a statement of higher expectancy for the graduate students. It is NOT to serve as a license for you to do less than what you are capable of.

Evaluation categories: Written Presentation: 20%; Academic content: 50%; Log Book: 30%.