CN550 CLASS PROJECT
Comparative Analysis of Supervised Learning Systems

The CN550 class project develops a systematic analysis of supervised learning systems, neural networks and others. You will investigate a variety of systems that explore the following problem.

During training, a set of input vectors **a** learns to predict a set of output vectors **b** :
Training $\mathbf{a} \rightarrow \mathbf{b}$
During testing, a new set of input vectors **a** try to predict the correct output vectors **b** :
Testing $\mathbf{a} \rightarrow$?? [compare with correct output **b**]

Supervised learning systems include multi–layer perceptrons (MLP), support vector machines (SVM), ARTMAP, logistic regression, and K nearest neighbors (KNN).

Class groups will analyze many different algorithms, with common benchmark problems and system evaluation criteria. Toward the end of the semester you will integrate findings, draw conclusions, and present results. You will post your results on the CN550 wiki:
http://cns.bu.edu/cn550/pmwiki/pmwiki.php?n=PmWiki.SystemComparisons ,
and write a final essay summarizing your contributions and experiences.

BENCHMARK PROBLEMS

Your group will test performance of various classifiers on the following problems.
Circle–in–the–square (CIS)
TBA
The problems are described in the document:
http://cns.bu.edu/cn550/550Benchmarks2009.pdf

For additional benchmark data and descriptions, see the the UCI machine learning repository:
http://www.ics.uci.edu/~mlearn/MLRepository.html/

and the CN710 wikis:
Fall 2008 http://cns.bu.edu/cn710/Fall2008/
Fall 2007 http://cns.bu.edu/cn710/Fall2007/
Fall 2006 http://cns.bu.edu/cn710/Fall2006/pmwiki.php?n=Main.HomePage/
Spring 2006 http://cns.bu.edu/cn710/Spring2006/pmwiki.php?n=Main.HomePage/

Each group will have a designated coordinator, whose responsibilities include scheduling meetings, monitoring progress, and organizing presentations. You might also wish to delegate another student to keep track of your records, files, reference articles, etc. and/or a web master.

CN550 CLASS PROJECT ASSIGNMENTS

Due Monday, February 2: Group organization and system evaluation criteria.
Discuss how your group will operate. You may decide to assign primary responsibility for each system to one group member. As you progress, keep in mind that your group as a whole is responsible for producing an accurate set of results. You should think about procedures for accomplishing this, including testing one another's algorithms and reading system specifications.
As a group, start to develop a list of system evaluation criteria.

Due Friday, February 6:
Each group coordinator should post on the wiki a list of the systems that the group has chosen to study.
Select a total of (at least) two systems per group member, according to the guidelines below.
You should choose systems that you have not implemented previously.
Take risks: try to choose at least one system you don't know much about. If you change your mind, you can make a new choice later on.

Due Monday, May 11 (at the final exam): A well written essay (approximately 2-3 pages) on the class project.
Include a summary of your own contributions to your group's effort and an assessment of the experience for you personally. Also include a discussion of your results vis-à-vis others posted on the wiki. E.g., if you implemented an SVM model, discuss possible reasons for different results posted for other SVMs.
Demonstrate the writing skills you have been practicing all semester.
Attach to this essay a printout of all your contributions to the CN550 wiki.

K nearest neighbors (KNN)
[*Each group should implement this baseline system.*]

Multi–layer perceptrons (MLPs)
[*Each group choose one or more.*]
Backprop with momentum
Cascade correlation
Elliptical basis functions
Levenberg-Marquardt backprop
Madaline
Quickprop
Radial basis function network (RBF)
Vanilla backprop
Other MLPs (provide references)

ARTMAP systems
[*Each group choose one or more.*]
ART–EMAP
ARTMAP–IC
Default ARTMAP
Distributed ARTMAP
Fuzzy ARTMAP
Gaussian ARTMAP
PROBART (Mariott & Harrison)
Other ARTMAP systems (provide references)

Support vector machines (SVMs)
[*Each group choose one or more.*]
Burges: Simplified SVM rules
Osuna's algorithm
Platt: Fast training using sequential minimal optimization
Other SVMs (provide references)

Additional supervised learning systems
Bayesian estimation
Decision trees
EM (expectation maximum) algorithm
Fuzzy classifiers – various types
Fuzzy KNN
Genetic algorithms (GA)
Learning vector quantization (LVQ)
Maximum likelihood
Memory–based systems (other than KNN)
Mixtures of experts models (e.g., Jordan & Jacobs)
Probabilistic Neural Network (PNN)
Sparse distributed memory
Other neural or non-neural systems (provide references)