CAP6615 Project

18th November 2002

The MNIST database (http://yann.lecun.com/exdb/mnist/ and locally at /cise/class/cap6615fa02) consists of a 60,000 example training set and a 10,000 example test set. Since this set is too large for a project, we will focus only on a small subset of examples. For this project, you are only required to use a set of “2”s and “3”s. The algorithms to be used are

1. Linear Fisher discriminant
2. Single hidden layer MLP
3. Support vector machine using a Gaussian radial basis function kernel
4. Kernel Fisher discriminant using a Gaussian radial basis function kernel

Cooperation between groups is encouraged for algorithms i) and ii) and prohibited for iii) and iv). Criteria for good project scores are

- use of a large number of training and test samples
- minimal use of free parameters,
- good training and test set performance and
- good documentation of free parameters used and how they were set

Choices left open to you are

- Number of training and test set samples to use
- Decision criteria in linear and kernel Fisher
- Number of hidden units in MLP

- Choice of stepsize parameter in MLP
- Convergence criteria in MLP
- Choice of nonlinearity (sigmoid etc.) and slope of nonlinearity β in MLP
- Choice of σ in GRBF and C in SVM

Choices not left open

- Please use a standard optimization package for the SVM training. Which one you choose is left up to you.