

# Homework 4 – Classification

Machine Learning / Greg Hamerly

March 26, 2007

This homework is due in class on Wednesday, April 4, 2007.

For this homework, you will use Matlab to implement LDA, and also use Support Vector Machines for the task of classification of images of handwritten digits.

**Dataset** Download the small digits dataset at

`http://cs.baylor.edu/~hamerly/courses/5v93\_07s/small.digits.tar.gz`

and read the description in the zip.info file. The training data is in 10 separate files (train.0 - train.9), and the test data is in zip.test. The training data does not have labels (examples are grouped by class in separate files), but the first column in the test data is the class.

**Linear discriminant analysis** Implement an LDA classifier for the model with the training set for all ten classes, and also just for classes 3 and 8.

**Support vector machine learning** Download the SVMLight software from `http://svmlight.joachims.org/`.

Read the documentation on how to do classification, including the data representation required. Try this software just on the digits 3 and 8, and try several kernels – linear, polynomial, and radial basis functions (aka Gaussian). For each kernel, try several kernel parameters, and try setting the -c value (which we discussed as  $\gamma$ , the multiplier for  $\sum \xi_i$  in the SVM optimal criterion). Try also letting the software choose -c by itself. Finally, you should allow the software to estimate the leave-one-out estimates with '-x 1'.

**Analyzing results** Describe the training and test error for each method. Which works the best? Which is the easiest to interpret? For SVMs, consider the number of support vectors chosen. Go beyond these basic questions, and show what you learned from using these models.

Your report should be written in L<sup>A</sup>T<sub>E</sub>X. You should describe how you set up your experiments, make any relevant graphs to illustrate your findings, and analyze your results. Please also comment on what you learned. Finally, please provide the code you wrote as an appendix.