CSE252C – Object Recognition – Assignment #4

Instructor: Prof. Serge Belongie.

`http://www-cse.ucsd.edu/classes/fa07/cse252c`

Due Date: Friday Dec. 7, 2007.

1. Experimenting with AdaBoost.

   - Use `generate_data_1.m` with $n = 1000$ to create a synthetic training and testing set.
   - Implement Discrete AdaBoost using a weak learner of your choice, and demonstrate it on the above synthetic data using 60 boosting iterations.
   - As a function of the number of iterations, plot the training error, the testing error, and the upper bound on the empirical error.

2. Suppose you and your homework partner obtain a dataset $\boldsymbol{x}^i \in \mathbb{R}^d$, $i = 1, \ldots, n$, from which you assemble a data matrix $X = [\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n] \in \mathbb{R}^{d \times n}$ and compute the inner product matrix $Q = X^\top X \in \mathbb{R}^{n \times n}$. After you compute Q, you realize you wanted to center the data before computing the inner products, i.e., to use $\boldsymbol{x}^i - \boldsymbol{\mu}$ in place of $\boldsymbol{x}^i$, where $\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{x}^i$. Unfortunately, you deleted the dataset.

   Show your partner that hope is not lost, since Q', the inner product matrix for the centered data, can be obtained from Q via the expression $Q' = HQH$ with $H = I - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top$, where $\mathbf{1}_n$ denotes a column vector of n ones.

3. Kernel PCA Experiment on Toy Data.

   (a) Implement Kernel PCA using a Gaussian kernel.

   (b) Reproduce the result in Fig. 4 of Schölkopf et al. (1999).