

Instructor: Rocco Servedio
Class Manager: Andrew Wan

Computer Science 4252: Introduction to Computational Learning Theory
Problem Set #5 Spring 2006

Due 5:00pm Monday, Apr 10, 2006

Problem 1 In class we worked through a simple boosting procedure which runs a weak learning algorithm 3 times (on 3 different distributions) and combines the resulting hypotheses h_1, h_2, h_3 by taking majority vote. We saw that if each weak hypothesis has error rate exactly 40% under its distribution, then the resulting final hypothesis has error rate 35.2%.

Now generalize what we did by assuming that each weak hypothesis has error rate exactly β (for some value $\beta < \frac{1}{2}$) under its distribution. Prove that the error rate of the resulting final hypothesis is strictly less than β . (Hint: Show that this final error rate is $3\beta^2 - 2\beta^3$.)

Problem 2 Show that in the Adaboost algorithm, the error of hypothesis h_t on distribution \mathcal{D}^{t+1} is exactly $1/2$.