

MCS 548 – Mathematical Theory of Artificial Intelligence
Fall 2014
Problem Set 2

Lev Reyzin

Due: 10/28/14 at the beginning of class

Instructions: Atop your problem set, please write your name and list your collaborators.

Problems

1. Assume that the weak learning assumption of AdaBoost holds. Let h_t be the base learner selected at round t . Show that the base learner h_{t+1} selected by AdaBoost at round $t + 1$ must be different from h_t .
2. Let the training sample be $S = ((x_1, y_1), \dots, (x_m, y_m))$. Suppose we wish to penalize the errors on x_i and x_j differently. To do that, we associate a non-negative weight w_i to each point x_i and we define the objective function $F(\alpha) = \sum_i^m w_i e^{-y_i g(x_i)}$, where $g = \sum_{t=1}^T \alpha_t h_t$. Use this function to derive a boosting algorithm.
3. The function of the slack variables we used in the optimization for soft margin hyperplanes has the form $\sum_i^m \xi_i$. Instead we could use $\sum_i^m \xi_i^p$ with $p > 1$. Give the dual formulation of the optimization problem in this general case.