MCS 548 – Mathematical Theory of Artificial Intelligence
Fall 2016
Problem Set 2

Lev Reyzin

Due: 10/21/16 at the beginning of class

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

Problems

1. Give as strong a lower bound as you can on the statistical query dimension of DNF formulae of length n, where n is to total number of (not necessarily unique) literals occurring in the formula. What, if anything, does your lower bound imply about their learnability in the SQ model of learning? In particular, can we hope for an SQ algorithm for the class of DNF formulae that uses a polynomial number of queries with polynomially large tolerances.

2. 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.

3. Let the training sample be $S = ((x_1, y_1), \ldots, (x_m, y_m))$. 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 $g = \sum_{t=i}^T \alpha_t h_t$. Use this function to derive a boosting algorithm.

4. 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.