MCS 548 – Mathematical Theory of Artificial Intelligence
Fall 2018
Problem Set 2

Lev Reyzin

Due: 10/31/18 at the beginning of class

Instructions: Atop your problem set, please write your name and list your collaborators. You may consult outside references, but cite all the resources used (e.g. which resources on the internet you consulted). You should not, however, search for answers to these questions. All problems in this assignment require proof.

Problems

1. Show that conjunctions are efficiently learnable with statistical queries.

2. Prove that statistical query dimension is lower-bounded by VC-dimension. Namely show that for every concept class C,

$$\text{VC-DIM}(C) \leq \text{SQ-DIM}(C).$$

Extra credit: give a bound in the other direction, namely give a function f such that for every concept class C, $\text{SQ-DIM}(C) \leq f\left(\text{VC-DIM}(C)\right).$

3. 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.

4. Suppose that in boosting, the weak learning condition is guaranteed to hold so that $\epsilon_t \leq \frac{1}{2} - \gamma$ for some $\gamma > 0$ that is known before boosting begins. Describe a modified version of AdaBoost whose final classifier is an *unweighted* majority vote and whose training error is at most $(1 - 4\gamma^2)^{T/2}$.

5. 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 $p > 1$. Derive the dual formulation of the optimization problem in this general case.