MCS 548 – Mathematical Theory of Artificial Intelligence
Fall 2020
Problem Set 2

Lev Reyzin

Due: 10/19/20 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. [10 pts.] Show that for `AdaBoost` $\prod_t Z_t$ is a monotonically decreasing function of t.

2. [10 pts.] 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. [10 pts.] 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 $(1 - 4\gamma^2)^{T/2}$.

4. [10 pts.] Show that conjunctions are efficiently learnable with statistical queries.

5. [10 pts.] Consider modifying the statistical query model by letting the learner request a polynomial number of unlabeled examples from the target distribution, in addition to having access to the usual statistical query oracle. We would also then include a failure parameter δ in the learning criterion. Call this new model "SQUE" (for $\underline{\text{SQ}}$ with $\underline{\text{U}}$nlabeled $\underline{\text{E}}$xamples). For efficient learnability, we know that

$$\text{SQ} \subseteq \eta\text{−PAC} \subseteq \text{PAC}.$$

Where would SQUE fit in this hierarchy for efficient learning?

6. [10 pts.] Let $X = \{0,1\}^n$, $c \in \{0,1\}^n$, and U be the uniform distribution over X. Define parity functions $\chi_c : X \to \{-1, 1\}$ as

$$\chi_c(x) = (-1)^{c \cdot x}.$$

Prove the following important fact used in the computation of SQ-DIM:

$$\forall c \neq c',\ \mathrm{E}_{x \sim U}[\chi_c(x)\chi_{c'}(x)] = 0.$$