Machine Learning Theory CSCI599, **Spring 2017**

Problem Set 2

Due: Tuesday, February 28, 3 p.m. by email

Please title your email "CSCI599_PS2".

---

1. **PAC Learning Axis-aligned Rectangles in $\mathbb{R}^n$**

   Do Exercise 1.1 of the Kearns–Vazirani textbook. In other words, give an efficient PAC learning algorithm for the class of axis-aligned rectangles in $\mathbb{R}^n$.

2. **Two-oracle PAC Model**

   Do Exercise 9 in Chapter 3 of the Shalev-Shwartz–Ben-David textbook. In other words, show the equivalence of the standard PAC model and the two-oracle PAC model.

3. **Properties of VC dimension**

   (a) Monotonicity of VC dimension: Do Exercise 1 in Chapter 6 of the Shalev-Shwartz–Ben-David textbook.

   (b) VC dimension versus log of class size: Do Exercise 7 in Chapter 6 of the Shalev-Shwartz–Ben-David textbook.

   (c) VC dimension of union: Do Exercise 11 in Chapter 6 of the Shalev-Shwartz–Ben-David textbook.

4. Recall that the conversion from an online algorithm with mistake bound m to a PAC algorithm given in class works as follows: "Run A on a sequence of examples each drawn independently from $\mathcal{D}$: If hypothesis h ever survives $(1/\epsilon)\log(\frac{m+1}{\delta})$ consecutive examples without making a mistake, stop and output h."

   Now suppose that you have an online algorithm A with some finite mistake bound m, but you don't know what the value of m is. Explain how you can obtain a PAC algorithm from A. What is the best sample complexity (in terms of m, ϵ, and δ) that you can achieve for your PAC algorithm?

5. Let X be the infinite set $\{1, 2, 3, \ldots\}$. Let $P_1, P_2, P_3, \ldots$ be an infinite list of computer programs, each of which takes as input an element $x \in X$ and outputs either 0 or 1. That is, each P_i computes some Boolean function $f_i : X \to \{0, 1\}$. Assume that the list $P_1, P_2, P_3, \ldots$ can be effectively enumerated, meaning that there is some computer program M which, given a value i as input, outputs program P_i.

   Suppose you are learning an unknown function f, which is guaranteed to be one of the f_i's, in the online mistake-bound model. Give a learning algorithm which is guaranteed to make $O(\log t)$ prediction mistakes, where t is the smallest index such that $f = f_t$.