Instructor: Rocco Servedio
Class Manager: Andrew Wan

Computer Science 4252: Introduction to Computational Learning Theory
Problem Set #2 Spring 2006

Due 5:00pm Monday, Feb 27, 2006

Problem 1 An online mistake-bound algorithm is said to be *conservative* if it changes its hypothesis only when a mistake is made.

Let C be a concept class and let A be an online learning algorithm (not necessarily conservative) which has a finite mistake bound M for C. Prove that there is a conservative learning algorithm A' for C which also has mistake bound M.

Problem 2 Determine the exact value of the VC dimension of the class C of conjunctions of literals (so negations are allowed) over $x_1, \ldots, x_n$. You must define a function $f(n)$; give a set of $f(n)$ examples and show that it is shattered by C; and show that no set of size $f(n)+1$ is shattered by C.

Problem 3

(i) Let C denote the class of monotone disjunctions of length at most k over $x_1, \ldots, x_n$. Show that the VC dimension of C is $\Theta(k \log n)$. (You can assume that k is small compared with n, e.g. $k \leq \sqrt{n}$). What does this say about the performance of the Winnow algorithm relative to the best possible mistake-bound algorithm for learning monotone disjunctions of length k?

(ii) Now let C denote the set of all decision lists over $x_1, \ldots, x_n$ of length at most k. Show that the VC dimension of C is $\Theta(k \log n)$. You can use the result from part (i) in your solution here even if you can't prove part (i).

Hint: The following fact may be useful: for any finite concept class C, we have $VC - DIM(C) \leq \log_2 |C|$. You should prove this fact if you use it.

Problem 4 Recall that in the setup for the basic Weighted Majority algorithm, there is a pool of N experts each of whom makes a binary prediction at each trial.

(i) Suppose that there is a subpool of k experts each of whom makes at most m mistakes over a sequence of trials. What bound can you prove for the Weighted Majority algorithm on this sequence of trials?

(ii) Now suppose that there is a subpool of k experts which together make a total of at most m mistakes. What bound can you prove for the Weighted Majority algorithm? (The following fact may be useful: for any convex function $f : [0, t] \to R$ and any values $x_1, \ldots, x_n \in [0, t]$, we have $\sum_{i=1}^{n} f(x_i) \geq n f(\frac{x_1 + \cdots + x_n}{n})$.)

Problem 5 You are going to the racetrack with a group of N friends, where you will each bet on the same set of T races. In each bet, each of your friends will either incur no loss or will incur a loss of 1 dollar.

Show that by using an appropriate strategy, the expected loss you incur per race will be at most $O(\sqrt{(\ln N)/T} + (\ln N)/T)$ more than the average loss incurred per race by the most successful of your friends. **Hint:** You may use the fact that $\ln \frac{1}{\beta} \leq (1 - \beta^2)/(2\beta)$ for $\beta \in (0, 1]$.