Mehryar Mohri
Foundations of Machine Learning
Courant Institute of Mathematical Sciences
Homework assignment 2
Due: February 27, 2007

A. Bounds on VC dimension

1. Show that the VC dimension of the set of all closed balls in $\mathbb{R}^n$, that is sets of the form $\left\{x \in \mathbb{R}^n : \|x - x_0\|^2 \leq r\right\}$ for some $x_0 \in \mathbb{R}^n$ and $r \geq 0$ is less than or equal to $n + 2$.

2. Determine the VC dimension of the subsets of the real line formed by the union of k intervals.

B. VC dimension of intersection concepts

1. Let C_1 and C_2 be two concept classes. Show that for any concept class $C = \{c_1 \cap c_2 : c_1 \in C_1, c_2 \in C_2\}$,

$$\Pi_C(m) \leq \Pi_{C_1}(m)\, \Pi_{C_2}(m). \tag{1}$$

2. Let C be a concept class with VC dimension d and let C_s be the concept class formed by all intersections of s concepts from C, $s \geq 1$. Show that the VC dimension of C_s is bounded by $2ds \log_2(3s)$ (*Hint*: show that $\log_2(3x) < 9x/(2e)$ for any $x \geq 2$).

C. Infinite VC dimension

1. Show that if a concept class C has infinite VC dimension, then it is not PAC-learnable.

2. In the standard PAC-learning scenario, the learning algorithm receives all examples first and then computes its hypothesis. Within that setting, PAC-learning of concept classes with infinite VC dimension is not possible as seen in the previous question.

   Imagine now a different scenario where the learning algorithm can alternate between drawing more examples and computation. The objective of this problem is to prove that PAC-learning can then be possible for some concept classes with infinite VC dimension.

To do so, consider for example the special case of the concept class C of all subsets of natural numbers.

(a) Show that the VC dimension of C is infinite.

(b) Professor Vitres has an idea for the first stage of a learning algorithm L PAC-learning C. In the first stage, L draws a sufficient number of points m such that the probability of drawing a point beyond the maximum value M observed be small, with high confidence. Can you complete Professor Vitres' idea by describing the second stage of the algorithm so that it PAC-learns C? The description should be augmented with the proof that L can PAC-learn C.