Final Exam

CSE 254 / Winter 2011

Name: ______________________________

ID: ______________________________

Be clear and concise. Write your answers in the space provided. Use the backs of pages for scratchwork.

1	
2	
3	
4	
5	

1. Let F be the set of functions defining disks in R^2, i.e. functions of the form

$$f(\vec{x}) = \begin{cases} 1 & \text{if } \|\vec{x} - \vec{c}\|_2 \leq r \\ 0 & \text{otherwise} \end{cases}$$

where $\vec{c} \in R^2$ is the center of the disk and $r \in R$ is it's radius.

Find the VC dimension of F, prove your answer.

2. Let F be the set of indicator function for second degree polynomials on defines over $(x, y) \in R^2$, i.e. functions of the form

$$f(x,y) = \begin{cases} 1 & \text{if } \sum_{0 \leq i+j \leq 2} a_{i,j} x^i y^j > 0 \\ 0 & \text{otherwise} \end{cases}$$

give the tightest (smallest) upper bound you can on the number of labelings for a set of m instances.

3. Consider the problem of two people crossing an intersection from two different directions. The main source of conflict is that if both people cross at the same time, there is likely to be an accident, which will hurt both people.

   We can describe this interaction using the following game matrix between a row player and a column player (each entry defines row loss/column loss).

	cross	stop
cross	100/100	0/5
stop	5/0	5/5

   (a) Find all of the Nash equilibria for this game and the expected loss of each player under each of the equilibria.

   (b) We say that an equilibrium is "fair" if the expected losses of all the players are equal. One of the Nash equilibria for this game is fair. Explain in words what behaviour of the two people is represented by this equilibrium.

   (c) Find a *correlated* equilibrium that is fair and that has a lower expected loss for both players than the fair Nash equilibrium.

4. **Question 4.6 in Cesa-Bianchi and Lugosi**
   Suppose your have N experts whose loss at each iteration is a number in the range $[0, 1]$. Show that the regret of "follow the leader" algorithm is bounded by the number of times the "leader" (i.e. the action with minimal cumulative loss) changes during the sequence of plays.

5. **Question 3.6 in Schapire and Freund**
Let $\mathcal{X}_n = \{0,1\}^n$, and let $\mathcal{G}_n$ be a class of boolean functions $g : \mathcal{X}_n \to \{-1,+1\}$. Let $\mathcal{M}_{n,k}$ be the class of all boolean functions that can be written as a simple majority vote of k functions in $\mathcal{G}_n$:

$$\mathcal{M}_{n,k} \doteq \left\{ f : x \mapsto \operatorname{sign}\left(\sum_{j=1}^{k} g_j(x) \right) \mid g_1, \ldots, g_k \in \mathcal{G}_n \right\}.$$

In this problem, we will see, roughly speaking, that if f can be written as a majority vote of polynomially many functions in $\mathcal{G}_n$, then under any distribution, f can be approximated by some function in $\mathcal{G}_n$. But if f cannot be so written as a majority vote, then there exists some "hard" distribution under which f cannot be approximated by *any* function in $\mathcal{G}_n$.

(a) Show that if $f \in \mathcal{M}_{n,k}$ then for all distributions D on $\mathcal{X}_n$, there exists a function $g \in \mathcal{G}_n$ for which

$$\mathrm{Pr}_{x\sim D}\left[f(x) \neq g(x)\right] \leq \frac{1}{2} - \frac{1}{2k}.$$

(b) Show that if $f \notin \mathcal{M}_{n,k}$ then there exists a distribution D on $\mathcal{X}_n$ such that

$$\mathrm{Pr}_{x\sim D}\left[f(x) \neq g(x)\right] > \frac{1}{2} - \sqrt{\frac{n \ln 2}{2k}}$$

for every $g \in \mathcal{G}_n$.