8803 Machine Learning Theory

Homework # 3

Due: November 3rd 2011

This homework is due by the start of class on November 3rd. You can either submit the homework via the course page on T-Square or hand it in at the beginning of the class on November 3rd.

Groundrules:

- Your work will be graded on correctness, clarity, and conciseness.
- You may collaborate with others on this problem set and consult external sources. However, you must write your own solutions and list your collaborators/sources for each problem.

Problems:

1. VC-dimension of linear separators: In this problem you will prove that the VC-dimension of the class $H_n$ of halfspaces (another term for linear threshold functions) in $n$ dimensions is $n + 1$. We will use the following definition: The convex hull of a set of points $S$ is the set of all convex combinations of points in $S$; this is the set of all points that can be written as $\sum_{x_i \in S} \lambda_i x_i$, where each $\lambda_i \geq 0$, and $\sum_{i} \lambda_i = 1$. It is not hard to see that if a halfspace has all points from a set $S$ on one side, then the entire convex hull of $S$ must be on that side as well.

(a) [lower bound] Prove that $\text{VC-dim}(H_n) \geq n + 1$ by presenting a set of $n + 1$ points in $n$-dimensional space such that one can partition that set with halfspaces in all possible ways. (And, show how one can partition the set in any desired way.)

(b) [upper bound part 1] The following is “Radon’s Theorem,” from the 1920’s.

Theorem. Let $S$ be a set of $n + 2$ points in $n$ dimensions. Then $S$ can be partitioned into two (disjoint) subsets $S_1$ and $S_2$ whose convex hulls intersect.

Show that Radon’s Theorem implies that the VC-dimension of halfspaces is at most $n + 1$. Conclude that $\text{VC-dim}(H_n) = n + 1$.

(c) [upper bound part 2] Now we prove Radon’s Theorem. We will need the following standard fact from linear algebra. If $x_1, \ldots, x_{n+1}$ are $n + 1$ points in $n$-dimensional space, then they are linearly dependent. That is, there exist real values $\lambda_1, \ldots, \lambda_{n+1}$ not all zero such that $\lambda_1 x_1 + \ldots + \lambda_{n+1} x_{n+1} = 0$.

You may now prove Radon’s Theorem however you wish. However, as a suggested first step, prove the following. For any set of $n + 2$ points $x_1, \ldots, x_{n+2}$ in $n$-dimensional space, there exist $\lambda_1, \ldots, \lambda_{n+2}$ not all zero such that $\sum_{i} \lambda_i x_i = 0$ and $\sum_{i} \lambda_i = 0$. (This is called affine dependence.)