CMSC5724: Exercise List 4

Problem 1. A *rectangular classifier* h in $\mathbb{R}^2$ is described by an axis-parallel rectangle $r = [x_1, x_2] \times [y_1, y_2]$. Function h maps all the points covered by r to label 1, and all the points outside r to label -1. Give a set of 4 points $\mathbb{R}^2$ that can be shattered by the class of rectangular classifiers.

Problem 2. Prove: there does not exist any set of 5 points in $\mathbb{R}^2$ that can be shattered by the class of rectangular classifiers.

Problem 3. Let $\mathcal{P}$ be a set of points in $\mathbb{R}^d$ for some integer $d > 0$. Let $\mathcal{H}$ be a set of classifiers each of which maps $\mathbb{R}^d$ to $\{-1, 1\}$. Prove: for any $\mathcal{H}' \subseteq \mathcal{H}$, it holds that VC-dim$(\mathcal{P}, \mathcal{H}') \leq$ VC-dim$(\mathcal{P}, \mathcal{H})$.

Problem 4. Denote by $\mathcal{X} = \mathbb{R}^d$ (where d is an integer) the instance space and by $\mathcal{Y} = \{-1, 1\}$ the label space. Recall that a classifier is a function $h : \mathcal{X} \to \mathcal{Y}$. Given a classifier h, define its *complement* as the function $\bar{h} : \mathcal{X} \to \mathcal{Y}$ which, given an instance $x \in \mathcal{X}$, outputs 1 if $h(x) = -1$, or -1 otherwise. Let $\mathcal{H}$ be a set of classifiers. Define another set of classifiers as follows: $\bar{\mathcal{H}} = \{\bar{h} \mid h \in \mathcal{H}\}$. Prove: $(\mathcal{X}, \mathcal{H})$ and $(\mathcal{X}, \bar{\mathcal{H}})$ have the same VC dimension.

Problem 5*. In this problem, we will see that deciding *whether* a set of points is linearly separable can be cast as an instance of linear programming.

In the *linear programming* (LP) problem, we are given n constraints of the form:

$$\boldsymbol{\alpha}_i \cdot \boldsymbol{x} \geq 0$$

where $i \in [1, n]$, $\boldsymbol{\alpha}_i$ is a constant d-dimensional vector (i.e., $\boldsymbol{\alpha}_i$ is explicitly given), and $\boldsymbol{x}$ is a d-dimensional vector we search for. Let $\boldsymbol{\beta}$ be another constant d-dimensional vector. Denote by S the set of vectors $\boldsymbol{x}$ satisfying all the n constraints. The objective is to

- either find the best $\boldsymbol{x} \in S$ that maximizes the *objective function* $\boldsymbol{\beta} \cdot \boldsymbol{x}$ — in this case we say that the LP instance is *feasible*;
- or declare that S is empty — in this case we say that the instance is *infeasible*.

Suppose that we have an algorithm $\mathcal{A}$ for solving LP in at most $f(n, d)$ time. Let P be a set of n points in $\mathbb{R}^d$, each given a label that is either 1 or -1. Explain how to use $\mathcal{A}$ to decide in $O(nd) + f(n, d + 1)$ time whether P is linearly separable, i.e., whether there exists a vector $\boldsymbol{w}$ such that:

- $\boldsymbol{w} \cdot \boldsymbol{p} > 0$ for each $\boldsymbol{p} \in P$ of label 1;
- $\boldsymbol{w} \cdot \boldsymbol{p} < 0$ for each $\boldsymbol{p} \in P$ of label -1.

Note that the inequalities in the above two bullets are strict, while the inequality in LP involves equality.