CSE 250: Machine Learning Theory	Spring 2017
Homework 2	
Instructor: Raef Bassily	**Due on:** Thu May 11

Instructions and Notes

- For your proofs, you may use any result covered in class, but please cite the result that you use.
- The assignment will be graded on clarity and correctness. If your arguments are not clear and have holes in them, it will be assumed incorrect.
- **Notation:** For any positive integer k, we will use the notation $[k]$ to denote $\{1, \ldots, k\}$.

Problem 1: Some Properties of VC-dimension

- Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes over a domain $\mathcal{X}$. Suppose that $VC(\mathcal{H}_1) = k_1$ and $VC(\mathcal{H}_2) = k_2$ for some positive integers $k_1 \geq k_2$. Prove that $VC(\mathcal{H}_1 \cup \mathcal{H}_2) = O(k_1)$. [**Hint:** *Can $\mathcal{H}_1 \cup \mathcal{H}_2$ shatter a set of $2k_1+2$ points? You may find it useful to prove and use the following claim: Let $\mathcal{G}_1, \mathcal{G}_2 \subseteq \{0,1\}^N$ s.t. $\mathcal{G}_1 \cup \mathcal{G}_2 = \{0,1\}^N$. Suppose, w.l.o.g., that $|\mathcal{G}_1| \geq |\mathcal{G}_2|$. Then, there must exist a set of $N/2$ distinct indices $\{i_1, \ldots, i_{N/2}\} \subset [N]$ s.t. $|\{(b_{i_1}, \ldots, b_{i_{N/2}}) : \forall j \in [N/2],\ b_{i_j}$ is the i_j-th bit of $\mathbf{b}, \mathbf{b} \in \mathcal{G}_1\}| = 2^{N/2}$.*]
- Let $h_1 : \mathcal{X}_1 \to \{0,1\}$ and $h_2 : \mathcal{X}_2 \to \{0,1\}$ be hypotheses defined over domains $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. We say that $h_1 \odot h_2 : \mathcal{X}_1 \times \mathcal{X}_2 \to \{0,1\}$ is pointwise binary operation on (h_1, h_2) if for every $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, $h_1 \odot h_2\ (x_1, x_2)$ refers to the logical binary operation $h_1(x_1)$ AND $h_2(x_2)$.

  Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two hypothesis classes over $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Let $\mathcal{H}$ be the hypothesis class over the product domain $\mathcal{X}_1 \times \mathcal{X}_2$ defined as $\mathcal{H} = \{h_1 \odot h_2 :\ h_1 \in \mathcal{H}_1,\ h_2 \in \mathcal{H}_2\}$. Suppose $VC(\mathcal{H}_1) = k_1$ and $VC(\mathcal{H}_2) = k_2$ where $k_1 \geq k_2 \geq 1$. Show that $VC(\mathcal{H}) \geq k_1$.

Problem 2: VC Dimension of subsets of a finite domain

Let $\mathcal{X} = [d]$ be a finite domain. Let $0 \leq t \leq d$ be an integer. Obtain the VC-dimension of each of the following classes? (Prove your claims about the VC-dimension in each case.) Hence, for each class, give an upper bound on the sample complexity of agnostic-PAC learning.

- $\mathcal{H}_{=t} = \{h : \mathcal{X} \to \{0,1\} \text{ s.t. } |\{x : h(x) = 1\}| = t\}$. That is, the set of all functions that assign the value 1 to exactly t elements of $\mathcal{X}$.
- $\mathcal{H}_{\leq t} = \{h : \mathcal{X} \to \{0,1\} \text{ s.t. } |\{x : h(x) = 1\}| \leq t\}$. That is, the set of all functions that assign the value 1 to most at t elements of $\mathcal{X}$.

Problem 3: VC-dimension of Boolean Conjunctions

Let $\mathcal{H}_{\mathsf{Conj}}$ be the class of Boolean conjunctions over variables $x_1, \ldots, x_d$ for some fixed $d \geq 2$. Show that $d \leq VC(\mathcal{H}_{\mathsf{Conj}}) \leq d \log_2(3)$. [*Clarification:* You may assume that each variable may appear *at most once* in a conjunction; therefore, formulas such as $x_1 \wedge \overline{x}_1$ are excluded.]

Problem 4: VC-dimension of Convex sets in $\mathbb{R}^2$

Let $\mathcal{X} = \mathbb{R}^2$, and let $\mathcal{H}$ be the hypothesis class of all convex sets in the real plane. That is, each $h_\kappa \in \mathcal{H}$ corresponds to a convex set $\kappa \subseteq \mathbb{R}^2$, such that $h_\kappa(x) = 1$ if $x \in \kappa$ and 0 otherwise. What is $VC(\mathcal{H})$? (Prove your claims.) [**Hint:** When you think of sets that can be shattered by $\mathcal{H}$, think of sets of points located on the perimeter of a circle.]