Assignment 6

Pegah Khazaei

14 April 2020

Chapter 6

Problem 2

Part a

In order for the relation $VCdim(H_{=K}) = \min\{k, |x| - k\}$ to be true, we should demonstrate that: $VCdim(H_{=K}) \leq \min\{k, |X| - k\}$ and $VCdim(H_{=K}) \geq \min\{k, |X| - k\}$. For the first case, suppose that $C \subseteq X$ is a set of size $k + 1$. Then, $\forall x \in C \nexists h \in H_{=k}$ s.t. $h(x) = 1$ and $\forall x \in X \setminus C h(x) = \mathbb{1}_{[E]}$. We conclude that C is shattered by $H_{=k}$. Then $VCdim(H_{=K}) \geq \min\{k, |X| - k\}$.

Part b

We prove that $VCdim(H_{=K}) = k$. Suppose that $C \subseteq X$ be a subset of size $k + 1$. Therefore, there is no $h \in H_{\leq k}$ that satisfies $h(x) = 1$ for all $x \in C$.

Suppose that $C = \{x_1, x_2, ..., x_m\}$ is a set of size m such that $m \leq k$. Let $(y_1, y_2, ..., y_m) \in \{0, 1\}^m$ be a vector of labels. This labeling can be obtained by some hypothesis $h \in H_{\leq k}$ which satisfies $h(x_i) = y_i$ for every $x_i \in C$ and $h(x) = 0$.

Similarly if $C \subseteq X$ has a size of $|X| - k + 1$, for all $x \in C \nexists h \in H_{=k}$ s.t. $h(x) = 0$. Hence $VCdim(H_{=K}) \leq \min\{k, |X| - k\}$. For the second case we assume that $C = \{x_1, x_2, ..., x_m\} \subseteq X$ be a set of size m such that $m \leq \min\{k, |X| - k\}$. Let $(y_1, y_2, ..., y_m) \in \{0, 1\}^m$ be a vector of labels. We prove that $\sum_{i=1}^{m} y_i$ by s. Pick an arbitrary subset $E \subseteq X \setminus C$ of $k - s$ elements. Also suppose that $h \in H_{=k}$ be a hypothesis such that $\forall x_i \in C, h(x_i) = y_i$, for every $x \in X \setminus C$, we conclude that C is shattered by $H \leq k$ and this completes the proof.

Problem 9

We want to show that $VCdim(H_{a,b,c}) = 3$. We consider four arbitrary points $\{x_1, x_2, x_3, x_4\}$ and without the loss of generality we suppose that $x_1 < x_2 < x_3 < x_4$. If we want to consider a labeling over this set that the label of points are different, this will not be possible with each of the elements of $H_{a,b,c}$. Thus $VCdim(H_{a,b,c}) \leq 3$.

We consider 3 arbitrary points $\{1, 2, 3\}$. These 3 points have 8 different labeling. It suffices that the values of a, b and c to be put as indicated in the following table to generate all the possible cases:

1	2	3	a	b	s
+	+	+	0.5	3.5	1
-	+	+	1.5	3.5	1
+	-	+	1.5	2.5	-1
+	+	-	0.5	2.5	1
-	-	+	2.5	3.5	1
-	+	-	1.5	2.5	1
+	-	-	0.5	1.5	1
-	-	-	0.5	3.5	-1

Thus $VCdim(H_{a,b,c}) \geq 3$ and the proof is complete.

Problem 10

Part a

Based on the question, we have $VCdim(H) \geq d$. Thus there exists a set C such that C is shattered by H. Without loss of generality, we assume that $X = C$. Since H contains all functions from C to $\{0, 1\}$, for every algorithm, there is a distribution $\mathcal{D}$ such that $\min_{h \in H} L_{\mathcal{D}}(h) = 0$ in which $|X| = |C| = d = km$ thus $k = \frac{d}{m}$. As a result we have:

$$\mathbb{E}[L_{\mathcal{D}}(A(S))] \geq \frac{\frac{d}{m} - 1}{\frac{2d}{m}} = \frac{d - m}{2d}$$

Part b

Assume that $VCdim(H) = \infty$. We want to show that in this case H cannot be PAC learnable. Since $VCdim(H) = \infty$, we can consider the size of the set as large as possible. If we consider $d = 2m$, based on the previous part we could say that there is a distribution $\mathcal{D}$ that $\min_{h \in H} L_{\mathcal{D}}(h) = 0$ but $\mathbb{E}[L_{\mathcal{D}}(A(S))] \geq \frac{1}{4}$. Thus with of probability of at least $\frac{1}{7}$ we have:

$$L_{\mathcal{D}}(A(S)) - \min_{h \in H} L_{\mathcal{D}}(h) = L_{\mathcal{D}}(A(S)) \geq \frac{1}{8}$$

Now considering $\epsilon = \frac{1}{9}$ and $\delta = \frac{1}{8}$ it is obvious that H is not PAC learnable.

Problem 11

Part a

Consider H as follows: $H = \bigcup_{i=1}^{r} H_i$. Let $k \in [d]$ such that $\mathcal{T}_H(k) = 2^k$. Based on the definition of the growth function we have:

$$\mathcal{T}_{\bigcup_{i=1}^{r} H_i}(k) \leq \sum_{i=1}^{r} \mathcal{T}_{H_i}(k)$$

Now using Sauer's Lemma on each term of $\mathcal{T}_{H_i}$, we get:

$$\mathcal{T}_{\bigcup_{i=1}^{r} H_i}(k) \leq rm^d$$

Taking the logarithm of both sides yields:

$$k < d \log m + \log r$$

Now using the Lemma A.2 we have:

$$k < 4d \log(2d) + 2 \log r$$

Part b

Without loss of generality, we assume that $VCdim(H_1) = VCdim(H_2) = d$. Let k be such that $k \geq 2d + 2$. Using Sauer's Lemma we have:

$$\begin{aligned}
&\mathcal{T}_{H_1 \bigcup H_2}(k) \leq \mathcal{T}_{H_1}(k) + \mathcal{T}_{H_2}(k) \leq \\
&\qquad \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{i} = \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{k-i} \\
&\qquad = \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=k-d}^{k} \binom{k}{i} \leq \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=d+2}^{k} \binom{k}{i} \\
&\qquad\qquad < \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=d+1}^{k} \binom{k}{i} = \sum_{i=0}^{k} \binom{k}{i} = 2^k
\end{aligned}$$

Thus, $VCdim(H_1 \bigcup H_2) \leq k \leq 2d + 1$.