Advanced Machine Learning: Homework Problem Set II
Solutions

Guidelines: You have to submit hardcopy of the solutions (printed or hand-written) by March 7, 2018 beginning of lecture class. Write your name and roll number clearly on top of the solution. Be clear and precise in your solution.

Problem 1:

Show that if two hypothesis classes $\mathcal{H}_1$ and $\mathcal{H}_2$, have VC-dimension d each, their union has VC-dimension at most $2d+1$.

Proof. We show that $VCdim(\mathcal{H}_\infty \cup \mathcal{H}_\in) \leq VCdim(\mathcal{H}_\infty) + VCdim(\mathcal{H}_\in) + 1$ Number of ways a given set of m points can be classified using $\mathcal{H}_\infty \cup \mathcal{H}_\in$ is at most the number of classifications using $\mathcal{H}_\infty$ plus the number of classifications using $\mathcal{H}_\in$. This gives the following inequality for growth functions $\zeta_{\mathcal{H}_\infty \cup \mathcal{H}_\in}(m) \leq \zeta_{\mathcal{H}_\infty}(m) + \zeta_{\mathcal{H}_\in}(m)$. By Sauer Lemma $\zeta_{\mathcal{H}_\infty \cup \mathcal{H}_\in}(m) \leq \sum_{i=0}^{d} \binom{m}{i} + \sum_{i=0}^{d} \binom{m}{i}$. Using the identity $\binom{m}{i} = \binom{m}{m-i}$ and a change of variable, this can be re-written as

$$\begin{aligned}
\zeta_{\mathcal{H}_\infty \cup \mathcal{H}_\in}(m) &\leq \sum_{i=0}^{d} \binom{m}{i} + \sum_{i=0}^{d} \binom{m}{m-i} \\
&\leq \sum_{i=0}^{d} \binom{m}{i} + \sum_{i=m-d}^{d} \binom{m}{i} \quad \text{i replaced by m-i}
\end{aligned}$$

Now, if $m - d > d + 1$,i.e,$m \geq 2d + 2$

$$\begin{aligned}
&\leq \sum_{i=0}^{m} \binom{m}{i} - \binom{m}{d+1} = 2^m - \binom{m}{d+1} \\
&< 2^m
\end{aligned}$$

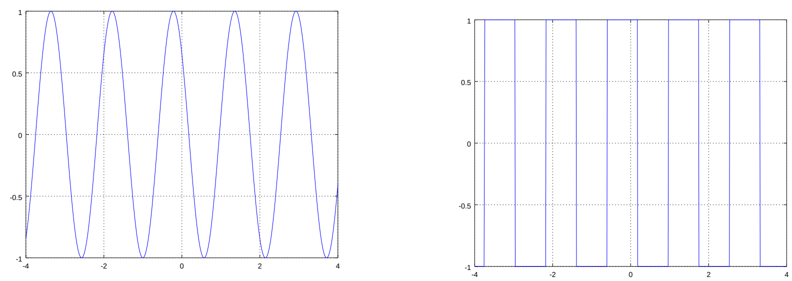
Thus $VCdim(\mathcal{H}_\infty \cup \mathcal{H}_\in)$ cannot be greater than or equal to $2d + 2$. Thus $VCdim(\mathcal{H}_\infty \cup \mathcal{H}_\in) \leq 2d + 1$ □

Problem 2:

Consider the hypothesis class $\mathcal{H}_{\text{sine}} = \{\text{sign}((\sin(ax + b)), a, b \in \mathbb{R}\}$ over the domain $x \in \mathbb{R}$.

(a) Draw a typical function of this class. Show that the points x, $2x$, $3x$, $4x$ cannot be shattered by $\mathcal{H}_{\text{sine}}$.

Proof. The function is a square wave with amplitude 1 and period $2\pi\cos(\tan^{-1}(a))$. Thus changing a,b we can get a function class.



(a) Graph of $sin(ax+b)$

(b) Graph of $\text{sign}(\sin(ax+b))$

□

(b) Show that VC dimension of the hypothesis class $\mathcal{H}_{\text{sine}}$ is infinite. (Note that, there are only two free parameters in this hypothesis class.)

Proof. We show that for any l the set of points $\{x_1, x_2, x_3, ..., x_l\}$ can be shattered where $x_i = 10^{-i}$. Choose any labeling $\{y_1, y_2, ..., y_l\}$. Let $a = \pi\left(1 + \sum_{i=1}^{l} \frac{(1-y_i)10^i}{2}\right)$ and $b = 0$, then

$$\begin{aligned} h(x_j) &= \text{sign}\left(\sin\left(10^{-j}\pi\left(1 + \sum_{i=1} l \frac{(1-y_i)10^i}{2}\right)\right)\right) \\ &= \text{sign}\left(sin\left(10^{-j}\pi + \sum_{i=1}^{l}(1-y_i)10^{i-j}\frac{\pi}{2}\right)\right) \end{aligned}$$

For any $y_i = 1$ the corresponding term in the summation will be zero. Also, any $i > j$, we will be adding an integral number of $\frac{\pi}{2}$ terms to a sine function, which causes no change in value. So those terms can be dropped from summation.

$$
\begin{aligned}
h(x_j) &= \text{sign}\left(sin\left(10^{-j}\pi + \sum_{i:i<j,y_i=-1}(1-y_i)10^{i-j}\frac{\pi}{2}\right)\right) \\
&= \text{sign}\left(sin\left(10^{-j}\pi + (1-y_j)\frac{\pi}{2} + \sum_{i:i<j,y_i=-1} 2*10^{i-j}\frac{\pi}{2}\right)\right) \\
&= \text{sign}\left(sin\left(10^{-j}\pi + (1-y_j)\frac{\pi}{2} + \pi\sum_{i:i<j,y_i=-1} 10^{i-j}\right)\right) \text{ We use } sin(\pi+x) = -sin(x)
\end{aligned}
$$

Summation of the last term(and the second term) is always less than 1. Therefore, $y_j = 1$ argument of the sine function is between 0 and π. Therefore the sine function takes positive values and $h(x_j) = 1 = y_j$. If $y_j = -1$ the first term becomes π and the argument of the sine function is between π and 2π. The sine function takes negative values $h(x_j) = -1 = y_j$. Thus $h(x_j) = y_j \forall j$. So the set $\{10^{-1}, 10^{-2}, ..., 10^{-l}\}$ can be shattered for any value of l. Therefore, the $VCdimension$ is infinite $\square$