MAT 5314

Assignment 3

Soufiane Fadel

September 23, 2019

Exercise 1

Let $b \leq c$ and $\epsilon > 0$ and assume $\mathcal{P}_D((b,c))$, where D is a distribution on $\mathbb{R}$ according to which points are drawn.

- **a)** Show that the probability that m points are drawn i.i.d. according to D without any of them falling in the interval (b, c) is at most $e^{m\epsilon}$.
- **b)** What is the VC dimension of the concept class C consisting of concepts that are the union of two closed intervals, e.g. $[a,b] \cup [c,d]$?
- **c)** Using only theorems we stated in class show that if a concept class has finite VC dimension then it is PAC-learnable
- **d)** Is the concept class C in (b) PAC learnable? Justify

♣ **solution for Q1:**

Since the probability of each point falling outside the interval is $1-\epsilon$:

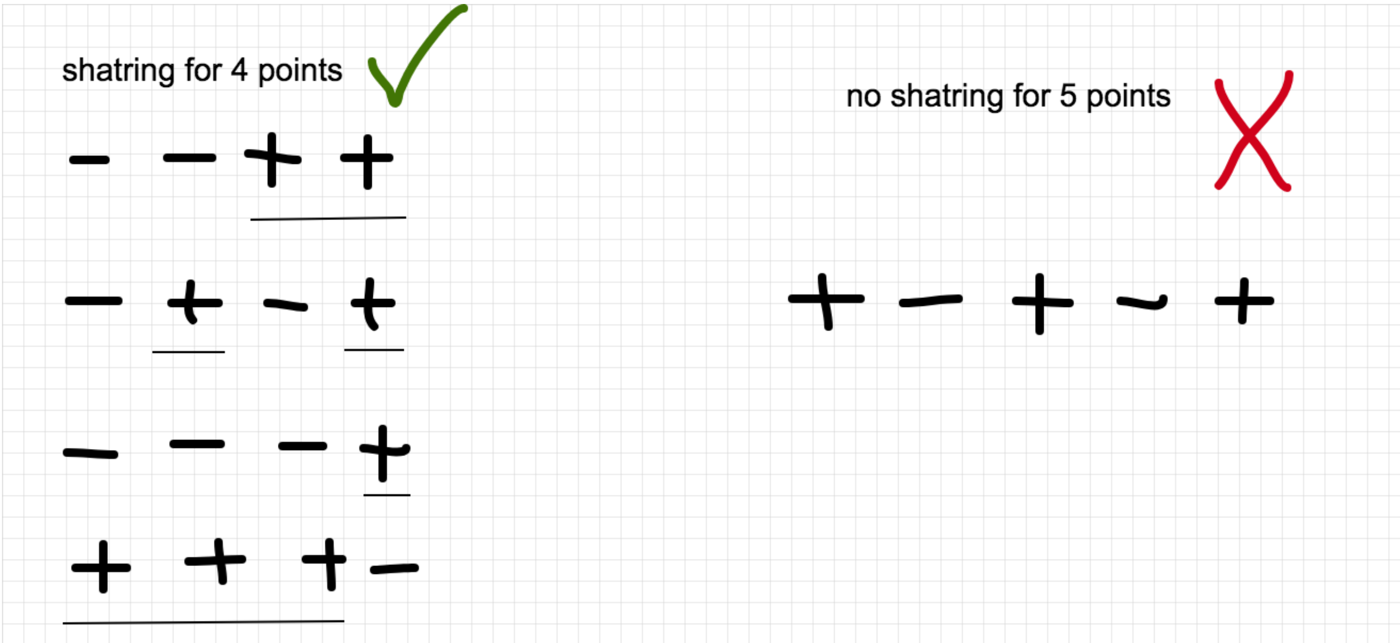
$$\begin{aligned}\mathcal{P}_x((b,c)) &= 1-\epsilon \\ &\leq e^{\epsilon}\end{aligned}$$

then we have for the whole data set :

$$\begin{aligned}\mathcal{P}_D((b,c)) &= (1-\epsilon)^m \\ &\leq e^{m\epsilon}\end{aligned}$$

♣ **solution for Q2:**

the VC Dimension for thoses concepts is : 4



♣ **solution for Q3:**

♣ **solution for Q4:**

using the claim on the question 1.c) and since the vc dimendion is $4 < \infty$ then the concept class C consisting of concepts that are the union of two closed intervals is PAC learnable.

Exercise 2

Let $X = \mathbb{R}^{\nvDash}$ and consider the set of concepts of the form $C = \{(x, y) \in \mathbb{R}^{\nvDash} | x^2 + y^2 \leq r\}$ for some real number r. Show that this class can be (ϵ, δ)-PAC-learned from training data of size $m \geq (\frac{1}{\epsilon})log(\frac{1}{\delta})$.

solution:

Exercise 3

If two concept classes both have VC dimension d what can you say about the VC dimension of their union? Try to make the strongest statement you can. Justify

solution:

we will proof that:

$$VCdim(H \cup H') \leq 2d + 1$$

in the Mohri textbook the VC-dimension of a hypothesis set is defined by :

$$m = \textbf{max}\, \{m : \Gamma_H(m) = 2^m\}$$

where $\Gamma_H(m)$ is the growth function which is the maximum number of ways points can be classified using H.

So since The number of ways m particular points can be classified using $H \cup H'$ is at most the number of classifications using H plus the number of classifications using H', This gives immediately the following inequality for growth functions for any $m \geq 0$:

$$\Gamma_{H \cup H'}(m) \leq \Gamma_H(m) + \Gamma_{H'}(m)$$

therefore by using saurez lemma we have :

$$\Gamma_{H \cup H'}(m) \leq \sum_{i=1}^{d} \binom{m}{i} + \sum_{i=1}^{d} \binom{m}{i}$$

and since we know that : $\binom{m}{i} = \binom{m}{m-i}$ then :

$$\begin{aligned} \Gamma_{H \cup H'}(m) &\leq \sum_{i=1}^{d} \binom{m}{i} + \sum_{i=1}^{d} \binom{m}{i} \\ &\leq \sum_{i=1}^{d} \binom{m}{i} + \sum_{i=1}^{d} \binom{m}{m-i} \\ &\leq \sum_{i=1}^{d} \binom{m}{i} + \sum_{i=m-d}^{m} \binom{m}{i} \end{aligned}$$

then w, if $m - d0 > d + 1$, that is $m \geq 2d + 2$

$$\begin{aligned} \Gamma_{H \cup H'}(m) &\leq \sum_{i=0}^{m} \binom{m}{i} - \binom{m}{d+1} \\ &= 2^m - \binom{m}{d+1} \\ &< 2^m \end{aligned}$$

and we we have :

$$\mathbf{max}\{m : \Gamma_H(m) = 2^m\} < 2d + 2$$

and finally:

$$VCdim(H \cup H') \leq 2d + 1$$

Exercise 4

Let F be a finite-dimensional vector space of real functions on $\mathbb{R}^n$ and $dim(F) = r < \infty$ Consider the concept class $C = \{c_f : f \in F\}$ where c_f is the concept corresponding to the set $\{x : f(x) \geq 0\}$ Show that the VC-dimension of C is at most r.

solution:

Asume $VCdim(C) \geq r + 1$ then : $\exists\{x_1 \cdots x_{r+1}\}$ of r+1 points in $\mathcal{R}^n$ such that C shatters $\{x_1 \cdots x_{r+1}\}$ which inplies that for each $i \in \{1, \cdots m\}$ (m=r+1) there is some $f_i \in \mathcal{F}$ such that:

$$f_i(x_1) < 0, f_i(x_2) < 0, \cdots f_i(x_i) \geq 0 \cdots f_i(x_m) < 0$$

. now lets consider ($\forall x \in \mathbb{R}^n$):

$$g_i(x) = \frac{1}{m-1} \sum_{j \neq i} \frac{f_i(x) - f_i(x_j)}{f_i(x_i) - f_i(x_j)}$$

then we have $g_i(x_j) = 0 \ \forall j \neq i$ and $g_i(x_i) = 1$.
moreover we have $g_i \in \mathcal{F}$. thus for any $i \in [|1, m|]$ there is some $g_i \in \mathcal{F}$ such that: $u(g_i) = e_i$
where

$$\begin{array}{ccccc} u & : & \mathcal{F} & \to & \mathbb{R}^m \\ & & f & \mapsto & (f(x_1), \cdots, f(x_m)) \end{array}$$

and

$$e_i = (0, \cdots, \underset{i}{1}, \cdots 0)$$

Therefore:

$$\begin{aligned} u \text{ is subjective } &\Longrightarrow dim(F) \geq dim(\mathbb{R}^n) \\ &\Longrightarrow r \geq m = r + 1 \\ &\Longrightarrow r \geq r + 1 \end{aligned}$$

which is impossible!