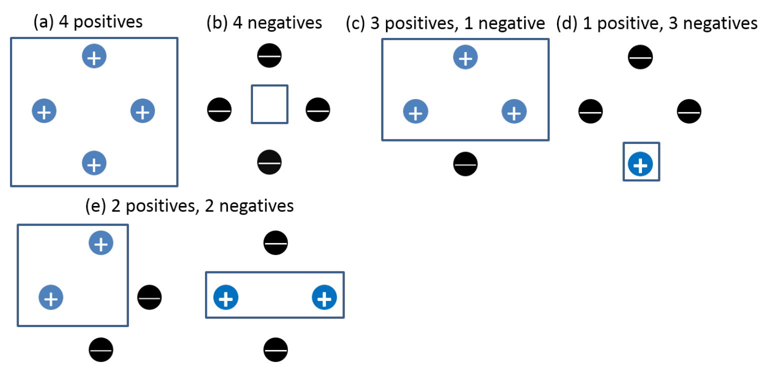
Solution of Assignment 2

October 30, 2014

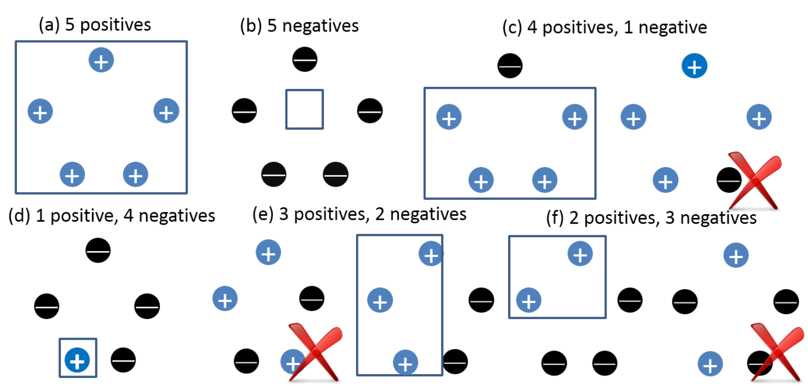
1. Suppose the input dimension is 2. Show that the VC dimension of the hypothesis class of axis-aligned rectangles that include positive instances inside and exclude negative instances outside is 4.

   Answer :

   - There exists a placement that positive rectangle can shatter four instances.

   - As we know, the five points of pentagon are the best placement to be shattered; however, there still exist some cases that the five points can't be shattered.

2. Read the proof of Hoeffding's inequality, and show that given a random variable Z and a real-valued function f with values $f(Z) \in [a, b]$, for any real number η, we have

$$E[e^{\eta f(Z)}] \leq \frac{b - E[f(Z)]}{b-a} e^{\eta a} + \frac{E[f(Z)] - a}{b-a} e^{\eta b}.$$

(Hint: $\exp(\cdot)$ is a convex function)

Answer :

Since $f(z) \in [a,\ b]$, $\theta = \frac{f(z)-a}{b-a} \in [0, 1]$ and $(1-\theta) = \frac{b-f(z)}{b-a} \in [0, 1]$.

Let $g(x) = e^{\eta x}$.

Since $g(x)$ is convex function, $g((1-\theta)a + \theta b) \leq (1-\theta)g(a) + \theta g(b)$

$g(f(z)) = g(\frac{b-f(z)}{b-a} \cdot a + \frac{f(z)-a}{b-a} \cdot b) \leq \frac{b-f(z)}{b-a} \cdot g(a) + \frac{f(z)-a}{b-a} \cdot g(b)$

$\Rightarrow e^{\eta f(z)} \leq \frac{b-f(z)}{b-a} \cdot e^{\eta a} + \frac{f(z)-a}{b-a} \cdot e^{\eta b}$

$\Rightarrow E[e^{\eta f(z)}] \leq E[\frac{b-f(z)}{b-a} \cdot e^{\eta a}] + E[\frac{f(z)-a}{b-a} \cdot e^{\eta b}] = \frac{b-E[f(z)]}{b-a} \cdot e^{\eta a} + \frac{E[f(z)]-a}{b-a} \cdot e^{\eta b}$.

3. Let $\mathcal{H}$ be a function class of VC dimension v. Based on Sauer's lemma, show that for $N > v$,

$$\mathcal{S}_{\mathcal{H}}(N) \leq \left(\frac{eN}{v}\right)^v.$$

(Hint: recall that $e^v = \lim_{n\to\infty}(1+\frac{v}{n})^n \geq (1+\frac{v}{N})^N$ for any $0 \leq \frac{v}{N} < 1$)

Answer :

Since $N > v$, we have $0 \leq \frac{v}{N} \leq 1$, and

$(\frac{v}{N})^v S_H(N) \leq (\frac{v}{N})^v \sum_{i=0}^{v} \binom{N}{i} \leq \sum_{i=0}^{v} (\frac{v}{N})^i \binom{N}{i} \leq \sum_{i=0}^{N} (\frac{v}{N})^i \binom{N}{i} = (1+\frac{v}{N})^N \leq e^v$.

So $S_H(N) \leq (\frac{Ne}{v})^v$.

4. Show that $Var[1(g^*(\boldsymbol{x}) \neq r)] \leq \frac{1}{4}$.
(Hint: $1(g^*(\boldsymbol{x}) \neq r) \in \{0, 1\}$ so if $P(1(g^*(\boldsymbol{x}) \neq r)) = p$, then $P(1(g^*(\boldsymbol{x}) = r)) = 1-p$)

Answer :

Let $X = 1(g^*(\boldsymbol{x}) \neq r) \Rightarrow X' = 1(g^*(\boldsymbol{x}) = r)$ and $P(X) = p \Rightarrow P(X') = 1-p$

$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$

$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1-p) = p$

$Var[X] = E[X^2] - E[X]^2 = p - p^2 = p(1-p)$

According to inequality of arithmetic-geometric mean, $Var[x] = p(1-p) \leq (\frac{p+(1-p)}{2})^2 = \frac{1}{4}$.

5. Given three datasets and three hypothesis classes, plot their generalization errors and VC bounds (see coding assignment for details). Which dataset and hypothesis class combination results in the tightest VC bound?

Answer :

see coding solution.