Homework #3 - B

Spring 2020, CSE 446/546: Machine Learning
Richy Yun
Due: Thursday 5/28/2020 11:59 PM

Intro to sample complexity

B.1

a. *[2 points]* Since $R(f)$ is the expected value that $f(X) \neq Y$:

$$R(f) > \epsilon \;\rightarrow\; \mathbb{P}(f(x_i) \neq y_i) > \epsilon$$
$$\mathbb{P}(f(x_i) = y_i) \leq 1 - \epsilon \leq e^{-\epsilon}$$
$$\mathbb{P}(f(x_i) = y_i)^n = \mathbb{P}(\widehat{R}_n(f) = 0) \leq e^{-n\epsilon}$$

b. *[2 points]* We can rewrite the expression:

$$Pr(\exists f \in \mathcal{F} \text{ s.t. } R(f) > \epsilon \text{ and } \widehat{R}_n(f) = 0) =$$
$$Pr(R(f_1) > \epsilon \text{ and } \widehat{R}_n(f_1) = 0 \cup R(f_2) > \epsilon \text{ and } \widehat{R}_n(f_2) = 0 \cup ... \cup R(f_{|\mathcal{F}|}) > \epsilon \text{ and } \widehat{R}_n(f_{|\mathcal{F}|}) = 0)$$

Since we know from part (a) that $Pr(R(f) > \epsilon \text{ and } \widehat{R}_n(f) = 0) \leq e^{-n\epsilon}$, using the union bound we get:

$$Pr(\exists f \in \mathcal{F} \text{ s.t. } R(f) > \epsilon \text{ and } \widehat{R}_n(f) = 0) \leq \textstyle\sum_{i=1}^{|\mathcal{F}|} e^{-n\epsilon} = |\mathcal{F}|e^{-n\epsilon}$$

c. *[2 points]* We simply solve for ϵ:

$$|\mathcal{F}|e^{-\epsilon n} \leq \delta$$
$$-\epsilon n \leq \ln\left(\frac{\delta}{|\mathcal{F}|}\right)$$
$$\epsilon \geq \frac{1}{n}\ln\left(\frac{|\mathcal{F}|}{\delta}\right)$$

So the minimum value is $\frac{1}{n}\ln\left(\frac{|\mathcal{F}|}{\delta}\right)$

d. *[4 points]* Following from parts (b) and (c) if $\widehat{R}_n(\widehat{f}) = 0$ then:

$$Pr(R(\widehat{f}) > \epsilon) \leq |\mathcal{F}|e^{-\epsilon n}$$
$$Pr(R(\widehat{f}) > \epsilon_{min}) \leq |\mathcal{F}|e^{-\epsilon_{min} n} \leq \delta$$
$$Pr(R(\widehat{f}) \leq \epsilon_{min}) \geq 1 - \delta$$

Since $R(f^*) \in [0,1]$, $R(\widehat{f}) - R(f^*) < R(\widehat{f})$. Thus:

$$Pr(R(\widehat{f}) - R(f^*) \leq \epsilon_{min}) \geq 1 - \delta$$
$$Pr\left(R(\widehat{f}) - R(f^*) \leq \frac{\ln(|\mathcal{F}|/\delta)}{n}\right) \geq 1 - \delta$$