CSE 546 Homework #3 B

Lu Yu

Intro to sample complexity

B1.

a. *[2 points]*

$$1 - \epsilon \le e^{-\epsilon}$$
$$1 - R(f) < 1 - \epsilon \le e^{-\epsilon}$$
$$Pr(R(f) = 0) \le e^{-\epsilon} \ since\ data\ drawn\ i.i.d$$
$$Pr(\hat{R}_n(f) = 0) = (Pr(R(f) = 0))^n \le e^{-n\epsilon}$$

b. *[2 points]*

$$Pr(\exists f \in \mathbb{F}\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0)$$
$$= Pr(\exists f_1\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0) \cdot Pr(\exists f_2\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0) \ldots Pr(\exists f_n\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0)$$
$$\le \sum_{i=1}^{|\mathbb{F}|} Pr(\exists f_i\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0) \qquad union\ bound$$
$$\le |\mathbb{F}| e^{-n\epsilon} \qquad according\ to\ part\ a$$

c. *[2 points]*

$$|\mathbb{F}| e^{-n\epsilon} \le \delta$$
$$\epsilon \ge -\frac{log(\frac{\delta}{|\mathbb{F}|})}{n} = \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}$$

So, the minimum $\epsilon = \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}$

d. *[4 points]* Since $\hat{R}_n(f) \ge 0$ and $\hat{R}_n(\hat{f}) = 0$, then $\hat{f} \in arg\ min_{f \in \mathbb{F}} \hat{R}_n(f)$
From part a, we have

$$Pr(\exists f \in \mathbb{F}\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(f) = 0) = Pr(\exists \hat{f} \in arg\ min_{f \in \mathbb{F}} \hat{R}_n(f)\ s.t. R(f) > \epsilon\ and\ \hat{R}_n(\hat{f}) = 0)$$
$$\ge Pr(\hat{R}_n(\hat{f}) = 0 \to R(\hat{f}) > \epsilon)$$

From part b, we have

$$Pr(\exists \hat{f} \in arg\ min_{f \in \mathbb{F}} \hat{R}_n(\hat{f})\ s.t. R(\hat{f}) > \epsilon\ and\ \hat{R}_n(\hat{f}) = 0) \le |\mathbb{F}| e^{-n\epsilon}$$

From part c, we have $\epsilon = \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}$

$$Pr(\exists \hat{f} \in arg\ min_{f \in \mathbb{F}} \hat{R}_n(f)\ s.t. R(\hat{f}) > \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}\ and\ \hat{R}_n(\hat{f}) = 0) \le \delta$$

Thus, we have

$$Pr(\hat{R}_n(\hat{f}) = 0 \to R(\hat{f}) > \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}) \le Pr(\exists \hat{f} \in arg\ min_{f \in \mathbb{F}} \hat{R}_n(f)\ s.t. R(\hat{f}) > \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}\ and\ \hat{R}_n(\hat{f}) = 0) \le \delta$$

Using inversion,

$$Pr(\hat{R}_n(\hat{f}) = 0 \rightarrow R(\hat{f}) \leq \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}) \geq 1 - \delta$$

Since $f^* \in arg\ min_{f \in \mathbb{F}} R(f), f^* > 0\ and\ R(\hat{f}) \leq \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}$

$$R(\hat{f}) - R(f^*) \leq \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}$$

$$Pr(\hat{R}_n(\hat{f}) = 0 \rightarrow R(\hat{f}) - R(f^*) \leq \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}) \geq Pr(\hat{R}_n(\hat{f}) = 0 \rightarrow R(\hat{f}) \leq \frac{log(\frac{|\mathbb{F}|}{\delta})}{n}) \geq 1 - \delta$$

Thus, we get that with probability at least $1 - \delta$:

$$R(\hat{f}) - R(f^*) \leq \frac{log(|\mathbb{F}|/\delta|)}{n}$$