Homework #3B

Spring 2020, CSE 446/546: Machine Learning
Prof. Kevin Jamieson and Prof. Jamie Morgenstern
Due: 5/28/20 11:59 PM

B.1

a. Since $R(f) > \epsilon$:
$\mathbb{P}(\hat{R}_n(f) = 0) = \mathbb{P}(\frac{1}{n}\sum_{i=1}^{n}(f(x_i) \neq y_i)) = \prod_{i=1}^{n}\mathbb{P}(f(x_i) = y_i)$
$= (1 - \mathbb{P}(f(x_1) \neq y_1))^n = (1 - \mathbb{E}_{XY}[(f(X) \neq Y)])^n = (1 - R(f))^n \leq (1-\epsilon)^n \leq e^{-n\epsilon}$

b. For every $f \in \mathcal{F}$, define $A_f = \{R(f) > \epsilon \text{ and } \hat{R}_n(f) = 0\}$. Note that $\mathbb{P}(\{R(f) > \epsilon \text{ and } \hat{R}_n(f) = 0\}) = \mathbb{P}(\hat{R}_n(f) = 0 | R(f) > \epsilon)\mathbb{P}(R(f) > \epsilon) \leq \mathbb{P}(\hat{R}_n(f) = 0 | R(f) > \epsilon)$ Then:

$$\mathbb{P}(\exists f \in \mathcal{F} : R(f) > \epsilon \text{ and } \hat{R}_n(f) = 0) \leq \mathbb{P}(\cup_{f\in\mathcal{F}} A_f) \leq \sum_{f\in\mathcal{F}} \mathbb{P}(A_f) \leq |\mathcal{F}|e^{-\epsilon n} \qquad \square$$

c. $|\mathcal{F}|e^{-\epsilon n} \leq \delta$ so $\epsilon \geq \frac{1}{n}\log\frac{|\mathcal{F}|}{\delta}$
so $\epsilon^* = \frac{1}{n}\log\frac{|\mathcal{F}|}{\delta}$

d. $1 - \mathbb{P}(\hat{R}_n(f) = 0 \text{ and } R(f) - R(f^*) > \epsilon^*)$
$\geq 1 - \mathbb{P}(\hat{R}_n(f) = 0 \text{ and } R(f) > \epsilon^* \geq 1 - \mathbb{P}(\exists f \in \mathcal{F} : R(f) > \epsilon^*)$ and
$\hat{R}_n(f) = 0) \geq 1 - |\mathcal{F}|e^{-\epsilon^* n} \geq 1 - \delta$