Assignment 3

Fateme Rahimi

March 15, 2020

Question 1. We should show that if set

$$\{(x_i, y_i)\}_{i=1}^{m} \subseteq (\mathbb{R}^d, \{0,1\})^m$$

as training set, and

$$h_s(x) = \begin{cases} 1 \text{ if } \exists \text{i} \in [\text{m}]\text{s.t x}_i = x \\ 0 \text{ o.w} \end{cases}$$

there exists a polynomial P_s such that:

$$h_s(x) = 1 \Leftrightarrow P_s(x) \geq 0$$

Proof. We can define P_s as:

$$P_s(x) = -\prod_{i=1}^{m} (||x - x_i||^2)$$

(Right to Left)if h$_s$(x)=1 $\Rightarrow$ $\exists$i $\in$ [m]s.t x$_i = x \Rightarrow P_s(x) = 0$
so if h$_s$(x)=1 $\Rightarrow P_s(x) = 0$
(Left to Right)if h$_s$(x) $\neq 1 \Rightarrow$ $\forall$i $\in$ [m]x$_i \neq x \Rightarrow P_s(x) < 0$
so if , h$_s$(x) $\neq 1 \Rightarrow P_s(x) < 0$
□

Question 2. We should proof that:

$$\underset{S|x\sim D^m}{E}[L_S(h)] = L_{(D,f)}(h)$$

Proof. by definition we konw:

$$L_S(h) = \frac{|i \in [m] : h(x_i) \neq f(x_i)|}{m}$$

we rewrite that as:

$$L_S(h) = \frac{\sum\limits_{i\in[m]} 1_{\{h(x_i)\neq f(x_i)\}}}{m}$$

now we calculate expectection value:

$$\underset{S|x\sim D^m}{E}[L_S(h)] = \underset{S|x\sim D^m}{E}(\frac{\sum\limits_{i\in[m]} 1_{\{h(x_i)\neq f(x_i)\}}}{m})$$

Expectection is linear function so we have:

$$\underset{S|x\sim D^m}{E}\left(\frac{\sum_{i\in[m]} 1_{\{h(x_i)\neq f(x_i)\}}}{m}\right) = \sum_{i\in[m]} \underset{S|x\sim D^m}{E}\left(1_{\{h(x_i)\neq f(x_i)\}}\right)$$

We know expectection of indicator is probability function:

$$\sum_{i\in[m]} \underset{S|x\sim D^m}{E}\left(1_{\{h(x_i)\neq f(x_i)\}}\right) = \frac{1}{m}\sum_{i\in[m]} \underset{S|x\sim D^m}{P}\left(h(x_i)\neq f(x_i)\right) =$$

$$\frac{1}{m}\times m\times \underbrace{\underset{S|x\sim D^m}{P}\left(h(x_i)\neq f(x_i)\right)}_{L_{(D,f)}(h)} = L_{(D,f)}(h)$$

□

Question 3.1. we want show that:

$$A \in \underset{h\in H_{rec}}{\arg}\ \min L_s(h)$$

Proof. It is clear that, A returns all the positive instances in the training set correctly.The tightest rectangle enclosing all positive examples beside realizability assumption guarantee all the negative instance labels correctly by A, so A is an ERM. □

Question 3.2. Show that if A receives a training set of size $\geq \frac{4\log(4/\delta)}{\varepsilon}$ then, with probability of at least $1-\delta$ it returns a hypothesis with error of at most ϵ

Proof. First we define R^* as in the hint. if $R(S)$ be the rectangle returned by A, from definition of A it is clear that:

$$R(S)\subseteq R^*$$

If rec be the corresponding hypothesis with R(S), since errors occur outside of $R(S)$ we have:

$$L_{(D,f)}(rec) = D(R^*\backslash R(S))$$

so

$$\{S : L_{(D,f)}(rec_s) > \varepsilon\} \subseteq \{S : S\cap(\bigcup_{i=1}^{4} R_i) = \varnothing\}$$

$$D^m(\{S : L_{(D,f)}(rec_s) > \varepsilon\}) \leq D^m(\{S : S\cap(\bigcup_{i=1}^{4} R_i) = \varnothing\})$$

If we write right side as union of $\{S : S\cap R_i = \varnothing\}$:

$$D^m(\{S : L_{(D,f)}(rec_s) > \varepsilon\}) \leq \sum_{i=1,2,3,4} D^m(\{S : S\cap R_i = \varnothing\})$$

and wen konw:

$$D^m(\{S : S \cap R_i = \varnothing\}) = (1 - \frac{\varepsilon}{4})^m$$

so:

$$\sum_{i=1,2,3,4} D^m(\{S : S \cap R_i = \varnothing\}) \leq 4 \times (1 - \frac{\varepsilon}{4})^m \leq 4 \exp(\frac{-m\varepsilon}{4})$$

from assumption $m \geq \frac{4 \log(4/\delta)}{\varepsilon}$ we conclude our proof.

$\square$