PAC Learning

$\mathcal{H}$ is PAC learnable if $\exists\, m_{\mathcal{H}} : (0,1)^2 \to \mathbb{N}$ and algorithm $\mathcal{A}$ s.t.: $\forall\, \epsilon, \delta \in (0,1)$, $\forall\, \mathcal{D} \sim \mathcal{X} \times \mathcal{Y}$, if realizability assumption holds, then $\mathcal{A}$ with $m \geq m_{\mathcal{H}}(\epsilon, \delta)$ iid sampes output $h = \mathcal{A}(S)$ s.t

$$\underset{S \sim \mathcal{D}^m}{P}\left[\mathcal{L}_{\mathcal{D}}(h) \leq \epsilon \right] \geq 1 - \delta$$

Def^n: S is $\underline{\epsilon\text{-representative}}$ wrt $(\mathcal{X}, \mathcal{Y}, \mathcal{H}, \mathcal{D}, \ell)$ if

$$\sup_{h \in \mathcal{H}} \left| L_{\mathcal{D}}(h) - L_S(h) \right| \leq \epsilon$$

$$\mathcal{F} := \ell \circ \mathcal{H} := \{ z \mapsto \ell(h, z) : h \in \mathcal{H} \}$$

Then,

$$\mathcal{L}_{\mathcal{D}}(f) = \underset{z \sim \mathcal{D}}{E}[f(z)] \qquad \mathcal{L}_S(f) = \frac{1}{m} \sum_{i=1}^{m} f(z_i)$$

Def^n: representativeness of S wrt $\mathcal{F}$

$$\mathrm{Rep}_{\mathcal{D}}(\mathcal{F}, S) := \sup_{f \in \mathcal{F}} \left(L_{\mathcal{D}}(f) - L_S(f) \right)$$

Goal: Estimate $\mathrm{Rep}_{\mathcal{D}}$ with only S

Idea: Split S

$$S = S_1 \cup S_2 \quad \text{where } S_1 \cap S_2 = \emptyset$$

Then,

$$\sup_{f \in \mathcal{F}} \left(L_{S_1}(f) - L_{S_2}(f) \right)$$

Let $|S_1| = |S_2|$

$$= \sup_{f \in \mathcal{F}} \frac{2}{m}\left(\sum_{S_1} f(z_i) - \sum_{S_2} f(z_i)\right)$$

Def'n: $\sigma = (\sigma_1, \dots, \sigma_m) \in \{\pm 1\}^m$

$$\sigma_i = \begin{cases} 1 & \text{if } z_i \in S_1 \\ -1 & z_i \in S_2 \end{cases}$$

$$\triangleright \sup_{f \in \mathcal{F}} \frac{2}{m} \sum \sigma_i f(z_i)$$

Radamacher Complexity

$$\mathcal{F} \circ S = \{(f(z_1), \dots, f(z_m)) : f \in \mathcal{F}\}$$

Let σ_i iid w/ $P[\sigma_i = 1] = P[\sigma_i = -1] = \frac{1}{2}$

$$R(\mathcal{F} \circ S) := \frac{1}{m} \mathop{E}_{\sigma \sim \{\pm 1\}^m}\left[\sup_{f \in \mathcal{F}} \sum_{i=1}^{m} \sigma_i f(z_i)\right]$$

Thm:

$$\mathop{E}_{S \sim D^m}\left[\mathrm{Rep}_D(\mathcal{F}, S)\right] \leq 2 \mathop{E}_{S \sim D^m} R(\mathcal{F} \circ S)$$

Pf: Let $S' = \{z_1', \dots, z_m'\}$

$$\forall f \in \mathcal{F} \quad L_D(f) = \mathop{E}_{S'}\left[L_{S'}(f)\right]$$

$$\mathop{E}_{S \sim D^m}\left[\sup_f \left(L_D(f) - L_S(f)\right)\right]$$

$$= \mathop{E}_{S'}\left[L_{S'}(f)\right] - L_S[f] = \mathop{E}_{S'}\left[L_{S'}(f) - L_S(f)\right]$$

$$\leq \mathop{E}_{S'}\left[\sup_{f \in F}\left(L_{S'}(f) - L_S(f)\right)\right]$$

$$E_S\left[\sup_{f\in\mathcal{F}}\left(L_D(f)-L_S(f)\right)\right]\le \underset{S,S'}{E}\left[\sup_{\mathcal{F}}\left(L_{S'}(f)-L_S(f)\right)\right]$$

$$=\frac{1}{m}\underset{S,S'}{E}\left[\sup_{f\in\mathcal{F}}\sum_{i=1}^{m}\left(f(z_i')-f(z_i)\right)\right]$$

Fix $j\in[m]$

$$\frac{1}{m}\underset{S,S'}{E}\left[\sup\left(\left(f(z_j')-f(z_j)\right)+\sum_{i\ne j}\left(f(z_i')-f(z_i)\right)\right)\right]$$

$$=\frac{1}{m}\underset{S,S'}{E}\left[\sup\left(\left(f(z_j)-f(z_j')\right)+\sum_{i\ne j}\left(f(z_i')-f(z_i)\right)\right)\right]$$

Let σ_j s.t. $P[\sigma_j=1]=P[\sigma_j=-1]=\frac{1}{2}$

$$\frac{1}{m}\underset{S,S',\sigma_j}{E}\left[\sup_{f\in\mathcal{F}}\left(\sigma_j\left(f(z_j')-f(z_j)\right)+\sum_{i\ne j}\left(f(z_i')-f(z_i)\right)\right)\right]$$

$$=\frac{1}{2}\bigstar+\frac{1}{2}\varphi$$

$$=\frac{1}{m}\underset{S,S'}{E}\left[\sup_{f\in\mathcal{F}}\left(\left(f(z_j')-f(z_j)\right)+\sum_{i\ne j}\left(f(z_i')-f(z_i)\right)\right)\right]$$

Repeat all j

$$\frac{1}{m}\underset{S,S'}{E}\left[\sup_{f\in\mathcal{F}}\sum\left(f(z_i')-f(z_i)\right)\right]=\frac{1}{m}\underset{S,S',\sigma}{E}\underbrace{\left[\sup_{f\in\mathcal{F}}\sum\sigma_i\left(f(z_i')-f(z_i)\right)\right]}$$

$$\le\sup_{\mathcal{F}}\sum\sigma_i f(z_i')+\sup_{\mathcal{F}}\sum-\sigma_i f(z_i)$$

$P[\sigma]=P[-\sigma]$

$$\frac{1}{m} \underset{S,S',\sigma}{E} \left[\sup_{f \in \mathcal{F}} \sum \sigma_i f(z_i') + \sup_{f \in \mathcal{F}} \sum \sigma_i f(z_i) \right]$$

$$\underset{S'}{E} \left[R(\mathcal{F} \circ S) \right] + \underset{S}{E} \left[R(\mathcal{F} \circ S) \right] = 2 \underset{S}{E} \left[R(\mathcal{F} \circ S) \right]$$