9 Sept 2021

Chap 2 ... Partial Answers to Some Exercises.

Ex 2.1

(1) Two oracle variant of PAC.

$(\mathcal{C}, \mathcal{H})$ is two-oracle PAC learnable if $\exists$ an algo. $\mathcal{A}$ (randomized) and a polynomial $\text{poly}(\cdot,\cdot,\cdot)$ s.t. $\forall D_+, D_- \in \Pr(\mathcal{X})$ $\forall \varepsilon, \delta > 0$ $\forall m > 0$, if $m \geq \text{poly}(\frac{1}{\varepsilon}, \frac{1}{\delta}, n)$ (size of the input) and $\exists c \in \mathcal{C}$ s.t. $\forall x\, (D_+(x) > 0 \Rightarrow c(x) = 1)$ $\wedge\, (D_-(x) > 0 \Rightarrow c(x) = 0)$,

$$\mathbb{P}\Big[\ \mathbb{P}_{x \sim D_+}[\,h_{S_+,S_-}(x) = 0\,] \leq \varepsilon \ \wedge\ \mathbb{P}_{x \sim D_-}[\,h_{S_+,S_-}(x) = 0\,] \leq \varepsilon\ \Big] \geq 1 - \delta.$$

where h_{S_+,S_-} is the result of the algo. $\mathcal{A}$ given samples $S_+ \sim D_+^m$ and $S_- \sim D_-^m$.

(2) $h_0 = x \mapsto 0$, $h_1 = x \mapsto 1$. We want to show:

$(\mathcal{C}, \mathcal{H} \cup \{h_0, h_1\})$ is two-oracle PAC learnable

iff. $(\mathcal{C}, \mathcal{H})$ is PAC learnable. (randomized variant).

(3). Proof sketch.

$\Leftarrow$). Assume that $(\mathcal{C}, \mathcal{H})$ is PAC-learnable.

$\mathcal{A}$ ... algo., $\text{poly}(\cdot,\cdot,\cdot)$ ... polynomial from the def'n of PAC.

We should build an algo. $\mathcal{A}'$ and a polynomial poly' required in 2-oracle PAC.

- $\mathcal{A}'((x_1^+, x_2^+, \ldots, x_m^+), (x_1^-, x_2^-, \ldots, x_m^-)) =$

  for $i = 1$ to m

  $(x_i', y_i') = (x_i^+, 1)$ with prob $\frac{1}{2}$ and $(x_i^-, 0)$ with prob $\frac{1}{2}$.

  return $\mathcal{A}((x_1', y_1'), \ldots, (x_m', y_m'))$.

$- \mathrm{poly}'\left(\frac{1}{\varepsilon}, \frac{1}{\delta}, n\right) = \mathrm{poly}\left(\frac{2}{\varepsilon}, \frac{1}{\delta}, n\right).$

Claim: A' and poly' satisfy the condition in 2-oracle PAC.

Pick $D_+, D_-, \varepsilon, \delta, m$ s.t. $m \geq \mathrm{poly}'\left(\frac{1}{\varepsilon}, \frac{1}{\delta}, n\right)$.

Let $D = \frac{1}{2}D_+ + \frac{1}{2}D_-$, i.e., the $\left(\frac{1}{2}, \frac{1}{2}\right)$ mixture of D_+ and D_-. Then, since A and poly satisfy the usual PAC condition,

$$\mathop{\mathbb{P}}_{S\sim D}\left[\mathop{\mathbb{P}}_{x\sim D}[h_S(x) \neq c(x)] \leq \frac{\varepsilon}{2}\right] \geq 1-\delta.$$

where

c is some fn in $\mathcal{C}$ s.t. $\forall x. \left(D_+(x) > 0 \Rightarrow c(x) = 1\right)$

and $\left(D_-(x) > 0 \Rightarrow c(x) = 0\right)$.

Then,

$$\mathop{\mathbb{P}}_{\substack{S_+\sim D_+ \\ S_-\sim D_- \\ S\sim \mathrm{mix}(S_+,S_-)}}\left[\frac{1}{2}\mathop{\mathbb{P}}_{x\sim D_+}[h_S(x) = 1] + \frac{1}{2}\mathop{\mathbb{P}}_{x\sim D_-}[h_S(x)=0] \leq \frac{\varepsilon}{2}\right] \geq 1-\delta.$$

$$\wedge|$$

$$\mathbb{P}\left[\mathop{\mathbb{P}}_{x\sim D_+}[h_S(x)=1] \leq \varepsilon \wedge \mathop{\mathbb{P}}_{x\sim D_-}[h_S(x)=0] \leq \varepsilon\right]$$

$$\|$$

$$\mathbb{P}\left[\mathop{\mathbb{P}}_{x\sim D_+}[h_{S_+,S_-}(x)=1] \leq \varepsilon \wedge \mathop{\mathbb{P}}_{x\sim D_-}[h_{S_+,S_-}(x)=0] \leq \varepsilon\right].$$

$\Rightarrow$) Given: A' and poly' from 2-oracle PAC.

To construct: A and poly from PAC.

$$A.\ (\underbrace{(x_{1,1}, y_{1,1}), \ldots, (x_{1,r}, y_{1,r}), \ldots (x_{m,1}, y_{m,1}), \ldots (x_{m,r}, y_{m,r})}_{S})$$

m rounds and r examples in each round.

if S doesn't contain at least m positive samples, return h_0.

if S doesn't contain at least m negative samples, return h_1

Otherwise, let S_+ be the x values of the first positive samples and let S_- be those of the first negative samples.

return $A'(S_+, S_-)$.

$\mathbb{P}[R(h_S) > \varepsilon] \leq \mathbb{P}[R(h_0) > \varepsilon \wedge \text{return case 1}]$

$\vee\ \mathbb{P}[R(h_1) > \varepsilon \wedge \text{return case 2}]$

$\vee\ \mathbb{P}[R(h_{S_+,S_-}) > \varepsilon \wedge \text{return case 3}]$

So, bounding each by $\delta/3$ gives what we want.

" $\leq (1-\varepsilon)^r \leq e^{-\varepsilon r} \leq \frac{\delta}{3}$.

$\hookrightarrow \quad r \geq \frac{1}{\varepsilon} \log \frac{\delta}{3}$.

The second leads to the same requirement on r.

The third leads to the condition that

$m \geq \text{poly}'(\frac{1}{\varepsilon}, \frac{3}{\delta}, n)$.

$\therefore \text{poly}(\frac{1}{\varepsilon}, \frac{1}{\delta}, n) = \frac{1}{\varepsilon} \log \frac{\delta}{3} \times \text{poly}'(\frac{1}{\varepsilon}, \frac{3}{\delta}, n)$.

$\square$

Ex 2.10.

$(x_1, y_1), \ldots, (x_m, y_m)$

Consider a uniform dist. D over $\{x_1, \ldots, x_m\}$.

and a concept c def. by the above pairs.

Pick $m' \geq \text{poly}(\frac{1}{(1/2m)}, \frac{1}{\delta}, n)$ ind. examples.

Run the given PAC algo. A on those examples.