Chap 2 – PAC learning framework

1. Reminder

(1) Risk or generalisation error

(2) Empirical risk or empirical error

(3) PAC-learnability.

$(\mathcal{C}, \mathcal{H})$ is PAC-learnable

$\subseteq X \to Y$ $\subseteq X \to Y$

$\{0,1\}$

$\|$

$X \to Y$

$\downarrow$

$R(h) = \mathbb{P}_{x \sim D}[h(x) \neq c(x)] = \mathbb{E}_{x \sim D}[\mathbb{1}\{h(x) \neq c(x)\}]$.

$\widehat{R}_S(h) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}\{h(x) \neq c(x)\} = \mathbb{E}_{x \sim \widehat{D}}[\mathbb{1}\{h(x) \neq c(x)\}]$

$S \sim D^m$

$S, y_1, \dots, y_m \mapsto h$.

if $\exists$ an algo. $\mathcal{A}$ and a poly $\mathrm{poly}(\dots)$ s.t.

$\forall c \in \mathcal{C}\ \forall D \in \mathcal{P}r(X).\ \forall \varepsilon > 0\ \forall \delta > 0.$

$\forall m \geq \mathrm{poly}\left(\frac{1}{\varepsilon}, \frac{1}{\delta}, n\right)$.

(size of the input)

$\mathbb{P}_{S \sim D^m}[R(h_S) \leq \varepsilon] \geq 1 - \delta$.

$\uparrow$