1) Boosting the confidence

Let $\varepsilon, \delta \in (0,1)$, pick k "chunks" of size $m_H(\varepsilon/2)$. Apply A on each of chunks, to obtain $\hat{h}_1, \ldots, \hat{h}_k$. Note that the probability that $\min_{i \in [k]} L_D(\hat{h}_i) \leq \min L_D(h) + \varepsilon/2$ is at least $1 - \delta_0^k \geq 1 - \delta/2$.

Now apply an ERM over the class $H := \{h_1, \ldots, h_k\}$ with the training data being the last chunk of size $\left\lceil \frac{4 \log(2k/\delta)}{\varepsilon^2} \right\rceil$

Denote the output hypothesis by $\hat{h}$. Using corollary 4.6 we obtain that with probability at least $1 - \delta/2$

$L_D(\hat{h}) \leq \min_{i \in [k]} L_D(h_i) + \varepsilon/2$. Applying the union bound we obtain that with probability at least $1 - \delta$

$$L_D(h) \leq \min_{i \in [k]} L_D(h_i) + \varepsilon/2 \leq \min_{h \in H} L_D(h) + \varepsilon$$

(4) Let X be a finite set of size n. Let B be a class of all functions from X to $\{0,1\}$. Then $L(B,T) = B$, and both are finite. Hence for any $T \geq 1$,

$$\mathrm{VCdim}(B) = \mathrm{vcdim}(L(B,T)) = \log 2^n = n$$

(b) denote by B the class of decision stumps in $\mathbb{R}^d$. Formally

$B = \{h_{j,b,\theta} : j \in [d], b \in \{-1,1\}, \theta \in \mathbb{R}\}$ where $h_{j,b,\theta}(x) = b.\mathrm{sign}(\theta - x_i)$

For each $j \in [d]$, $b \in \{-1,1\}$, $\theta \in \mathbb{R}\}$ where $h_{b,\theta}(x) = b.\mathrm{sign}(\theta - x_j)$

Note the $\mathrm{vcdim}(B_j) = 2$

clearly $B = \bigcup_{j=1}^{d} B_j$. Applying Exercis. 11 we conclude

$$\mathrm{vcdim}(B) \leq 16 + 2\log d$$

Assume that $d = 2^k$ for some $k \in \mathbb{N}$. Let $A \in \mathbb{R}^{k \times d}$ be the matrix whose column range over the (entire) set $\{0,1\}^k$

For each $i \in [k]$, let $x_i = A_{i,\cdot}$. we claim that the set $C = \{x_1 \ldots x_k\}$ is shatter. Let $I \subseteq [k]$. We show that Let $I \subseteq [k]$. we show that we can label the instances in I positively, while the instances $[k] \setminus I$ are labeled negatively. By our construction, there exists an index J such that $A_{i,J} = x_{i,J} = 1$ if $i \in I$. Then $h_{J,-1,1/2}(x_i) = 1$ if $i \in I$.

Scanned with CamScanner