FUNDAÇÃO GETULIO VARGAS

ESCOLA DE MATEMÁTICA APLICADA

MESTRADO 2015.3

APRENDIZAGEM POR MÁQUINAS

Prof Eduardo Mendes

Resolução do dever de casa #2

KIZZY TERRA

RIO DE JANEIRO

OUTUBRO/2015

1 Problem 2.1

Sabemos que:

$$E_{out}(g) \leq E_{in}(g) + \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$$

$$\epsilon(M, N, \delta) = \sqrt{\frac{1}{2N} ln \frac{2M}{\delta}}$$

a) Para $M = 1$, $\delta = 0,03$ e $\epsilon \leq 0,05$

$$\sqrt{\frac{1}{2N} ln \frac{2M}{\delta}} \leq (0,05)^2 \Rightarrow \frac{1}{2N} ln \frac{2M}{\delta} \leq (0,05)^2$$

$$N \geq \frac{1}{2} ln \frac{2}{0,03} \frac{1}{(0,05)^2} = 839,941$$

b) Para $M = 100$, $\delta = 0,03$ e $\epsilon \leq 0,05$

$$\sqrt{\frac{1}{2N} ln \frac{2M}{\delta}} \leq (0,05)^2 \Rightarrow \frac{1}{2N} ln \frac{2M}{\delta} \leq (0,05)^2$$

$$N \geq \frac{1}{2} ln \frac{200}{0,03} \frac{1}{(0,05)^2} = 1760,97$$

c) Para $M = 10000$, $\delta = 0,03$ e $\epsilon \leq 0,05$

$$\sqrt{\frac{1}{2N} ln \frac{2M}{\delta}} \leq (0,05)^2 \Rightarrow \frac{1}{2N} ln \frac{2M}{\delta} \leq (0,05)^2$$

$$N \geq \frac{1}{2} ln \frac{20000}{0,03} \frac{1}{(0,05)^2} = 2682,009$$

2 Problem 2.22

De antemão, pelo enunciado sabemos que $\mathbb{E}[\epsilon] = 0$ e $var(\epsilon) = \sigma^2$ e $E_{out}(g) = E_{x,y}[(g(x) - y(x))^2]$.

Portanto, podemos fazer:

$$\mathbb{E}_{\mathcal{D}}[E_{\text{out}}(g^{(\mathcal{D})})] = \mathbb{E}_{\mathcal{D}}[\mathbb{E}_{\mathbf{x}}[(g^{(\mathcal{D})}(\mathbf{x}) - y(\mathbf{x}))^2]] = \mathbb{E}_{\mathbf{x}}[\mathbb{E}_{\mathcal{D}}[g^{(\mathcal{D})}(\mathbf{x})^2] - \overline{g}(\mathbf{x})^2 + \overline{g}(\mathbf{x})^2 - 2\overline{g}(\mathbf{x})y(\mathbf{x}) + y(\mathbf{x})^2]$$

Usando o fato $\mathbb{E}_{\mathcal{D}}[g^{(\mathcal{D})}(\mathbf{x})^2] - \overline{g}(\mathbf{x})^2 = var(x)$ e substituindo $y(x)$ por $f(x) + \epsilon$, obtemos:

$$\mathbb{E}_{\mathcal{D}}[E_{\text{out}}(g^{(\mathcal{D})})] = \mathbb{E}_{\mathbf{x}}[var(\mathbf{x}) + \overline{g}(\mathbf{x})^2 - 2\overline{g}(\mathbf{x})f(\mathbf{x}) + f(\mathbf{x})^2 - 2\overline{g}(\mathbf{x})\epsilon + 2f(\mathbf{x})\epsilon + \epsilon^2]$$

Sabemos que: $bias(x) = \overline{g}(\mathbf{x})^2 - 2\overline{g}(\mathbf{x})f(\mathbf{x}) + f(\mathbf{x})^2$. Logo:

$$\mathbb{E}_{\mathcal{D}}[E_{\text{out}}(g^{(\mathcal{D})})] = var + bias - 2 \cdot \cancelto{0}{\mathbb{E}_{\mathbf{x},y}[\overline{g}(\mathbf{x})\epsilon]} + 2\cancelto{0}{\mathbb{E}_{\mathbf{x},y}[f(\mathbf{x})\epsilon]} + \mathbb{E}_{\mathbf{x},y}[\epsilon^2]$$

Como $\mathbb{E}[\omega^2] = (\mathbb{E}[\omega])^2 + var(\omega)$ temos:

$$\mathbb{E}[\epsilon^2] = \sigma^2$$

Assim:

$$\mathbb{E}_{\mathcal{D}}[E_{\text{out}}(g^{(\mathcal{D})})] = var + bias + \sigma^2$$