Blackboard 4.1 : Bagging

mismatch of pattern μ in model (copy) k

(1) $\delta_k^\mu = t^\mu - \hat{y}_k^\mu$

μ ← pattern, k ← model

slides (example)

$\delta_k^\mu = \{0.1; -0.1; 0; -0.9; +0.9; \ldots\}$

average mismatch for model k

(2) $\frac{1}{P} \sum_{\mu=1}^{P} \delta_k^\mu = d$

assumption: same for all models!

define:

(3) $\varepsilon_k^\mu = \delta_k^\mu - d \overset{(2)}{\Longrightarrow} \frac{1}{P} \sum_{\mu=1}^{P} \varepsilon_k^\mu = 0$

(4) $V = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{P} \sum_{\mu=1}^{P} (\varepsilon_k^\mu)^2$

all models

---

quadratic error for model k

$$E_k = \frac{1}{P} \sum_{\mu=1}^{P} (\delta_k^\mu)^2 \overset{(3)}{=} \frac{1}{P} \sum_\mu (\varepsilon_k^\mu + d)^2$$

$$= \frac{1}{P} \sum_\mu (\varepsilon_k^\mu)^2 + d^2 + 2d \underbrace{\frac{1}{P} \sum_\mu \varepsilon_k^\mu}_{= 0 \; (3)}$$

$\Longrightarrow$ expected error ("typical" model)

$$\langle E_k \rangle_k = \frac{1}{K} \sum_{k=1}^{K} E_k \overset{(4)}{=} d^2 + V$$

4.1 Bagging (continued)

bagged output

$$\hat{y}_{bag} = \frac{1}{K} \sum_k \hat{y}_k$$

Claim: (Slides)

$$E_{bag} \leq \langle E_k \rangle_k$$

analogous

(1') $\delta^\mu_{bag} = t^\mu - \hat{y}_{bag} = \frac{1}{K} \sum_k \left[t^\mu - \hat{y}^\mu \right] = \frac{1}{K} \sum_k \delta^\mu_k$

(2') $\frac{1}{P} \sum_{\mu=1}^{P} \delta^\mu_{bag} = d$

(3') $\varepsilon^\mu_{bag} = \delta^\mu_{bag} - d \overset{(1')+(3)}{=} \frac{1}{K} \sum_k \varepsilon^\mu_k \Rightarrow \sum_\mu \varepsilon^\mu_{bag} = 0$

---

quadratic error for bagged output

$$E_{bag} = \frac{1}{P} \sum_{\mu=1}^{P} \left(\delta^\mu_{bag} \right)^2 = \frac{1}{P} \sum_{\mu=1}^{P} \left(d + \varepsilon^\mu_{bag} \right)^2$$

$$\overset{(3')}{=} d^2 + 2d \cdot \underbrace{\cancel{\frac{1}{P} \sum_{\mu=1}^{P} \varepsilon^\mu_{bag}}}_{=0} + \frac{1}{P} \sum_{\mu=1}^{P} \frac{1}{K} \sum_{k=1}^{} \varepsilon^\mu_k \cdot \frac{1}{K} \sum_{n=1}^{P} \varepsilon^\mu_n$$

$k=n$ $\quad$ $k \neq n$

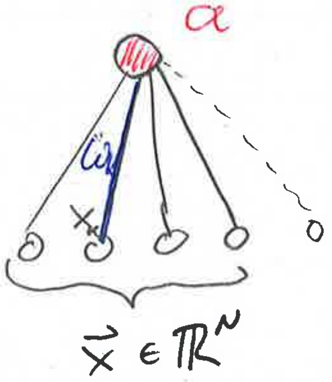
$$= d^2 + \frac{1}{K} \frac{1}{P} \frac{1}{K} \sum_\mu \sum_k \left(\varepsilon^\mu_k \right)^2 + \underbrace{\frac{1}{P} \frac{1}{K^2} \sum_\mu \sum_n \sum_{\substack{k \\ k \neq n}} \varepsilon^\mu_k \cdot \varepsilon^\mu_n}$$

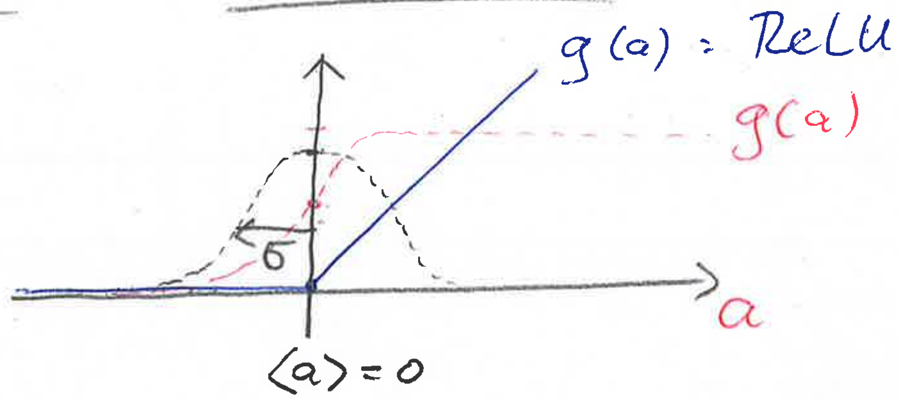
$$= d^2 + \frac{1}{K} \cdot V + C \leq \langle E_k \rangle_k \quad \blacksquare$$

scales with $1/K$ "number of copies"

$C = 0$ if independent: different data for different k

Blackboard 4.2 : Initialisation





$$a = \sum_k w_k \cdot x_k$$

chosen independently

$$\langle a \rangle = \sum_k \langle w_k \rangle \langle x_k \rangle = 0$$

= 0 Eq. (2) slides

= 0; Eq (1) slides

$$\sigma^2 = \langle a^2 \rangle = \left\langle \sum_{k=1}^{N} w_k \cdot x_k \cdot \sum_{n=1}^{N} w_n \cdot x_n \right\rangle$$

k = n

$$= \sum_{k=1}^{N} \langle w_k^2 \cdot x_k^2 \rangle + \sum_{k \neq n} \sum_{n} \underbrace{w_k \cdot w_n \cdot x_n \cdot x_n}_{} \quad = 0$$

independent   indendent

$$= N \cdot \langle w_k^2 \rangle \cdot \langle x_n^2 \rangle$$

preprocessing: $\langle x_k^2 \rangle = 1$

I want $\sigma = 2$

$$\frac{\sigma}{\sqrt{N}} = \sqrt{\langle w_k^2 \rangle} = \frac{2}{\sqrt{N}}$$