T031009

1 a)

- Särdrag troligen beroende ⇒ redundans
- Bättre generalisering i klassificeraren – minskad risk för överanpassning
- enklare tolkning

b) kNN: Långsam respons!

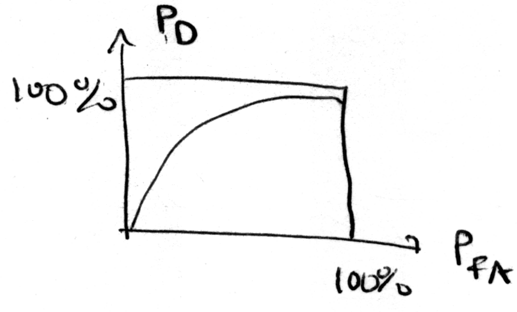
c) Se svaret i a). Kan bli sämre om verkligen alla särdrag är ganska oberoende men informativa

2. a)

$$\frac{e^{-\frac{1}{2}(x-1)^2}}{\sqrt{2\pi}} < \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \Rightarrow \text{klass } 1$$

b) Flytta så att färre exempel klassas till ω_2

c)



3 a) $\tilde{x}_i = \frac{x_i - m_i}{\sigma_i}$

b) $\tilde{y} = \bar{a}^T \tilde{x}$

$\tilde{x} = D(\bar{x} - \bar{m}_x)$

$D_{ii} = 1/\sigma_i$

$\tilde{y} = \frac{y - m_y}{\sigma_y}$

$$\frac{y - m_y}{\sigma_y} = \bar{a}^T \underline{\underline{D}} (\bar{x} - \bar{m}_x)$$

$$y = \underbrace{\sigma_y \bar{a}^T D}_{\bar{w}} \bar{x} + \underbrace{m_y - \sigma_y \bar{a}^T D \bar{m}_x}_{w_0}$$

c) $\lambda^* = 0 \Rightarrow$ OLS ok

4. a) PCR = PCA + OLS

b) (i) Många variable korrelerade $\Rightarrow$ stor varians i OLS-estimat

(ii) Alltför få exempel $\Rightarrow$ OLS lösning existerar ej

c)

$$J(\bar{w}) = \frac{1}{N} \sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x})^2$$

$$\frac{\partial J}{\partial \bar{w}} = -\frac{\cancel{2}}{\cancel{N}} \sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x}_n) \bar{x}_n = 0$$

$$\frac{1}{N} \sum_{n=1}^{N} y_n \bar{x}_n = \left(\frac{1}{N} \sum_{n=1}^{N} \bar{x}_n \bar{x}_n^T \right) \bar{w}$$

$$\hat{w}_{OLS} = \left(\frac{1}{N} \sum_{n=1}^{N} \bar{x}_n \bar{x}_n^T \right)^{-1} \left(\frac{1}{N} \sum_{n=1}^{N} \bar{x}_n y_n \right) \Leftarrow$$

6 a) ~~[illegible]~~ ~~dele~~

Punktskattningen $\frac{2}{20} = 10\%$ säger

mkt lite. ~~Ete~~

Osäkerheten om var e_T ligger kan beskrivas m.h.a. $p(e_T | k, N)$. Denna osäkerhetsfördelning har nästan all massa koncentrerad till intervallet $[2\%, 28\%]$

dvs "troligen" ligger e_T på detta intervall.

b)

⊕ Får info om prestand ~~[illegible]~~ design med $N-1 \approx N$ exempel

⊖ Får ingen info specifikt för den klassificerare som sedan tillämpas.

7 a) $J(\bar{w}) = \dfrac{m_2 - m_1}{\sigma_1^2 + \sigma_2^2}$

projektera $y = \bar{w}^T \bar{x}$ sådan att $J(\bar{w})$ maximeras $\Rightarrow$ "Bästa" seperation av klasserna

7b) $J(\bar{w}) = \| \bar{y} - \mathbf{X}^T \bar{w} \|^2$

7c

$$\bar{y}_i = \begin{cases} \begin{bmatrix} 1 \\ \bar{x}_i \end{bmatrix} & \text{om } \bar{x}_i \in \text{klass 1} \\ \begin{bmatrix} -1 \\ -\bar{x}_i \end{bmatrix} & \text{om } \bar{x}_i \in \text{klass 2} \end{cases}$$

$\bar{w}^T \bar{x}_i > \theta \Rightarrow$ klass 1

$\begin{bmatrix} -\theta & \bar{w}^T \end{bmatrix}^T \begin{bmatrix} 1 \\ \bar{x}_i \end{bmatrix} > 0 \Rightarrow$ klass 1

$\bar{v}^T \tilde{x}_i > 0 \Rightarrow$ klass 1. Ingel fel om $\tilde{x}_i \in$ klass 1

$\Rightarrow$ Fel om $\bar{v}^T \tilde{x}_i > 0$ men $\tilde{x}_i \in$ klass 2

eller

$\bar{v}^T \tilde{x}_i < 0$ men $\tilde{x}_i \in$ klass 1

$\Rightarrow$ Fel om $\bar{v}^T \bar{y}_i < 0$