Pen and Paper 4

(2) We need a dataset, where for one point

$$\arg\min_{\substack{x_i \in D \\ x_i \neq x}} \|x - x_i\|_2^2 \neq \arg\min_{\substack{x_i \in D \\ x_i \neq x}} k(x, x_i)$$

$D \subset [0,1]^2$

$\|x - y\|_2^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2$

$k(x, y) = \min(x_1, y_1) + \min(x_2, y_2)$

$x = (1,0)$ $\quad \|x - x_1\|_2^2 = 1$

$x_1 = (0,0)$ $\quad \|x - x_2\|_2^2 = \frac{1}{2}$ $\quad \Rightarrow x_2$ is the n.n.

$x_2 = (\frac{1}{2}, \frac{1}{2})$

$k(x, x_1) = \min(1, 0) + \min(0,0) = 0$

$k(x, x_2) = \min(1, \frac{1}{2}) + \min(0, \frac{1}{2}) = \frac{1}{2}$

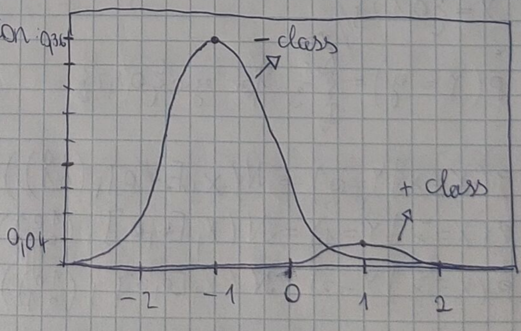
$\Rightarrow x_1$ is the n.n.

(3) a) data distribution

$P(N(\mu,1) = \mu) \approx 0{,}4$

$0{,}4 \cdot 0{,}9 = 0{,}36$

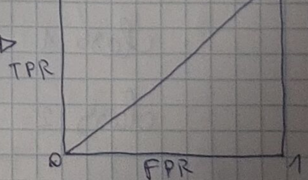
$0{,}4 \cdot 0{,}1 = 0{,}04$



$$TPR = \frac{TP}{TP + FN} = \frac{0{,}1 \cdot \tau}{0{,}1 \cdot \tau + 0{,}1(1-\tau)} = \tau$$

$$FPR = \frac{FP}{FP + TN} = \frac{0{,}9\tau}{0{,}9\tau + 0{,}9(1-\tau)} = \tau$$

$\rightarrow$



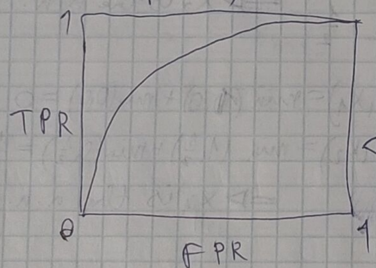
b) $TPR = \frac{TP}{TP+FN} = \frac{P(\text{pred}=+ \text{ and real}=+)}{P(\text{real}=+)} = \cancel{\frac{P(Y=+)\,\hat{Y}=1}{P(Y=+)}}$

$= \frac{P(Y=+)\, P(\hat{Y}=+ \mid Y=+)}{P(Y=+)} = P(X > \tau \mid Y=+) = 1 - P(X<\tau \mid Y=+) =$

$= 1 - \Phi(\tau - 1)$

$FPR = \frac{FP}{FP+TN} = \frac{P(Y=-)\cdot P(\hat{Y}=+ \mid Y=-)}{P(Y=-)} = P(\hat{Y}=+ \mid Y=-) =$

$= P(X \geq \tau \mid Y=-) = 1 - P(X<\tau \mid Y=-) = 1 - \Phi(\tau+1)$



$\Phi(\tau+1) > \Phi(\tau-1)$

$\Leftarrow \; 1-\Phi(\tau+1) < 1-\Phi(\tau-1)$

4 $\mathcal{Y} = \{0,1,2\}$

$P(Y=y) = \frac{1}{3}$ for $y \in \{0,1,2\}$

$P(X=x \mid Y=0) = N(x, [-3,0], \begin{pmatrix}1&0\\0&1\end{pmatrix})$

$P(X=x \mid Y=1) = N(x, [3,0], \begin{pmatrix}1&0\\0&1\end{pmatrix})$

$P(X=x \mid Y=2) = N(x, [0,3], \begin{pmatrix}1&0\\0&1\end{pmatrix})$

a) class 0: $w_0 = [-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]$, $b_0 = 0 \Rightarrow y = -x$

class 1: $w_1 = [\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]$, $b_1 = 0 \Rightarrow y = x$

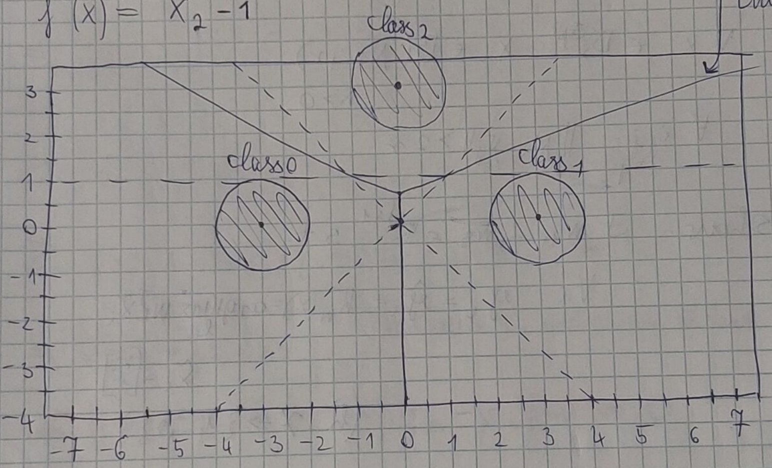
class 2: $w_2 = [0,1]$, $b_2 = -1 \Rightarrow y = 1$

$f^0(x) = -\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2 \quad \Rightarrow \quad x_2 - 1 = \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2$

$f^1(x) = \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2 \qquad \sqrt{2}x_2 - \sqrt{2} = x_1 - x_2$

$f^2(x) = x_2 - 1 \qquad 0 = x_1 - (1+\sqrt{2})x_2 + \sqrt{2}$

this is this



decision boundaries: - - -

one - vs - all classifiers: ——

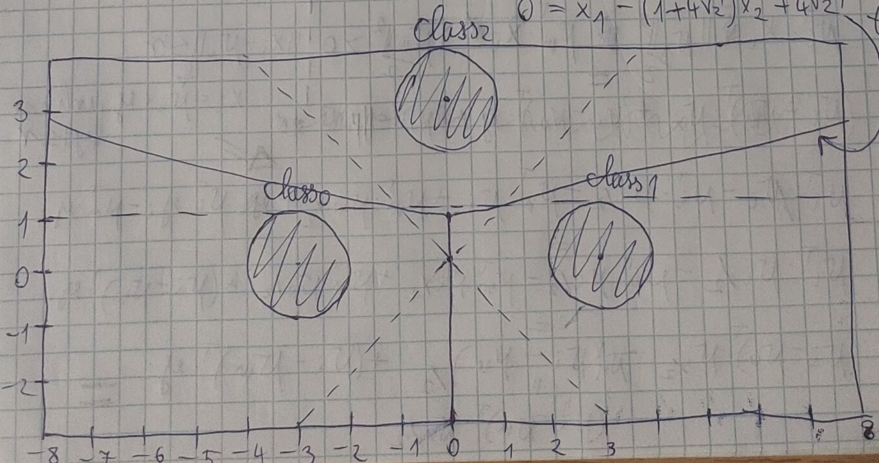
b) almost everything is the same, just we need:

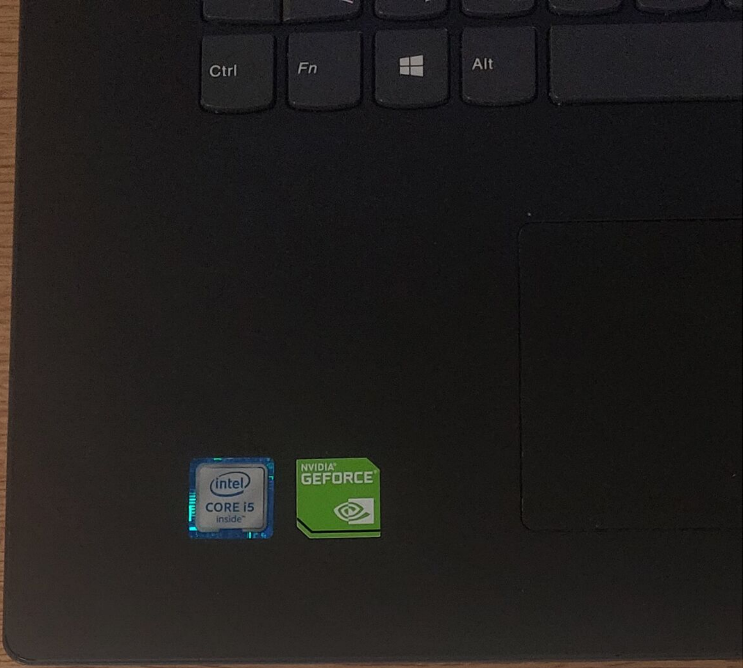
$$4x_2 - 4 = \frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_2$$

$$4\sqrt{2}x_2 - 4\sqrt{2} = x_1 - x_2$$

$$0 = x_1 - (1+4\sqrt{2})x_2 + 4\sqrt{2}$$

this is this





(1) $D = \{(x_i, y_i)\}_{i=0}^{n} \quad x_i \in \mathbb{R}^d, \; y_i \in \{1, \dots, k\}$

k centroids: $\mu_l \in \mathbb{R}^d$
$\qquad l \in \{1, \dots, k\}$

$\forall x_i \in D_x \; \exists \mu_l : \; x_i \in B(\mu_l, r)$
$\qquad r > 0$

$\forall \underset{i \neq j}{i, j} : \; \|\mu_i - \mu_j\| > 2r$

Show: $\exists \; w_1, \dots, w_k \in \mathbb{R}^{d+1}$ s.t.

$$\forall i: \; y_i = \hat{y}_i = h_w(x) = \underset{l}{\operatorname{argmax}} \, w_l^T \tilde{x}_i$$

$$\tilde{x}_i = \begin{bmatrix} x_i \\ 1 \end{bmatrix}$$

$$(h : X \to \{1, \dots, k\})$$

$w_i = \left[\mu_i, -\frac{\|\mu_i\|^2}{2}\right]$ will work:

We need that if $y_i = l$, then

$$w_l^T \tilde{x}_i > w_m^T \hat{x}_i \quad \forall \; \underset{m \neq l}{m = 1, \dots, k}$$

$$\mu_i^T \cdot x_i - \frac{\|\mu_i\|^2}{2} - \mu_m^T \cdot x_i + \frac{\|\mu_m\|^2}{2} > 0$$

$\|x_i - \mu_i\| \le r$
$\Downarrow$
$x_i = \mu_i + y$, where $\|y\| \le r$

$=$

$\cancel{(\mu_i - \mu_m)^T \cdot \mu_i + (\mu_i - \mu_m)^T y + \|\mu_i\|^2 - \|\mu_m\|^2}$

$$\mu_i^T \cdot \mu_i - \mu_i^T \cdot \mu_i/2 - \mu_m^T \cdot \mu_i + \mu_m^T \cdot \mu_m/2 + \mu_i^T \cdot y_i - \mu_m^T \cdot y_i$$

$$=$$

$$\mu_i^T \cdot \mu_i/2 - \mu_m^T \cdot \mu_i/2 - \mu_m^T \cdot \mu_i/2 + \mu_m^T \cdot \mu_m/2 + (\mu_i - \mu_m)^T \cdot y_i$$

$$=$$

$$(\mu_i - \mu_m)^T \cdot \mu_i/2 \; - \underbrace{\mu_m^T(\mu_i - \mu_m)/2}_{= (\mu_i - \mu_m)^T \cdot \mu_m/2} + (\mu_i - \mu_m)^T \cdot y_i \;=

$$= (\mu_i - \mu_m)^T \cdot (\mu_i - \mu_m)/2 + (\mu_i - \mu_m)^T \cdot y_i =$$

$$= \|\mu_i - \mu_m\|_2^2/2 + (\mu_i - \mu_m)^T \cdot y_i =$$

$$= \|\mu_i - \mu_m\|^2/2 - (\mu_m - \mu_i)^T \cdot y_i \geq$$

$$\geq \|\mu_i - \mu_m\|^2/2 - \|\mu_m - \mu_i\| \cdot \|y_i\| \geq$$

$$\geq \|\mu_i - \mu_m\|^2/2 - \|\mu_m - \mu_i\| \cdot r \geq$$

because μ_i and μ_m are at least $2r$ apart

$$\geq \|\mu_i - \mu_m\| \cdot r - \|\mu_m - \mu_i\| \cdot r = 0 \qquad \square$$