Exercice 1

1)

$$E = -\sum_{i=1}^{k} \overbrace{\underbrace{\hat{p}_\ell}_{\substack{\geq 0 \\ \text{car } \in [0,1]}} \underbrace{\log(\hat{p}_\ell)}_{\leq 0 \text{ car } \hat{p}_\ell \leq 1}}^{\leq 0}$$

Ainsi E est la somme des $-\hat{p}_\ell \log(\hat{p}_\ell)$ qui sont des termes ≥ 0 donc E est ≥ 0

Si $v_1 = \dots = v_N$ alors si $v_1 = \dots = v_N = l_0$

$$\hat{p}_\ell = \begin{cases} 1 & \text{si } \ell = l_0 \\ 0 & \text{sinon} \end{cases}$$

et $\hat{p}_\ell \log(\hat{p}_\ell) = 0$ car $0 \log(0) = 1 \log(1) = 0$

donc $E = 0$

Si les $v_1, \dots, v_N$ ne sont pas tous égaux alors il y a $\ell_1 \neq \ell_2$ tels que $\hat{p}_{\ell_1} > 0$ et $\hat{p}_{\ell_2} > 0$ donc

$\hat{p}_{\ell_1} \in]0,1[$ et $\hat{p}_{\ell_2} \in]0,1[$ et donc

$\hat{p}_{\ell_1} \log(\hat{p}_{\ell_1}) \neq 0$

et on a vu $-\hat{p}_{\ell_1}\log(\hat{p}_{\ell_1}) \geq 0$ donc $-\hat{p}_{\ell_1}\log(\hat{p}_{\ell_1}) > 0$

donc $$E = -\sum_{\ell=1}^{k} \hat{p}_\ell \log(\hat{p}_\ell) = \underbrace{-\hat{p}_{\ell_1}\log(\hat{p}_{\ell_1})}_{>0} + \underbrace{\sum_{\substack{\ell=1 \\ \ell\neq\ell_1}}^{k} -\hat{p}_\ell \log(\hat{p}_\ell)}_{\geq 0}$$

$$> 0$$

Donc $E = 0 \iff v_1 = \cdots = v_N$

2) $N = 10$ $\quad \hat{p}_1 = \frac{2}{10}$ $\quad \hat{p}_2 = \frac{6}{10}$ $\quad \hat{p}_3 = \frac{2}{10}$

$$E = -\frac{2}{10}\log\left(\frac{2}{10}\right) - \frac{6}{10}\log\left(\frac{6}{10}\right) - \frac{2}{10}\log\left(\frac{2}{10}\right)$$

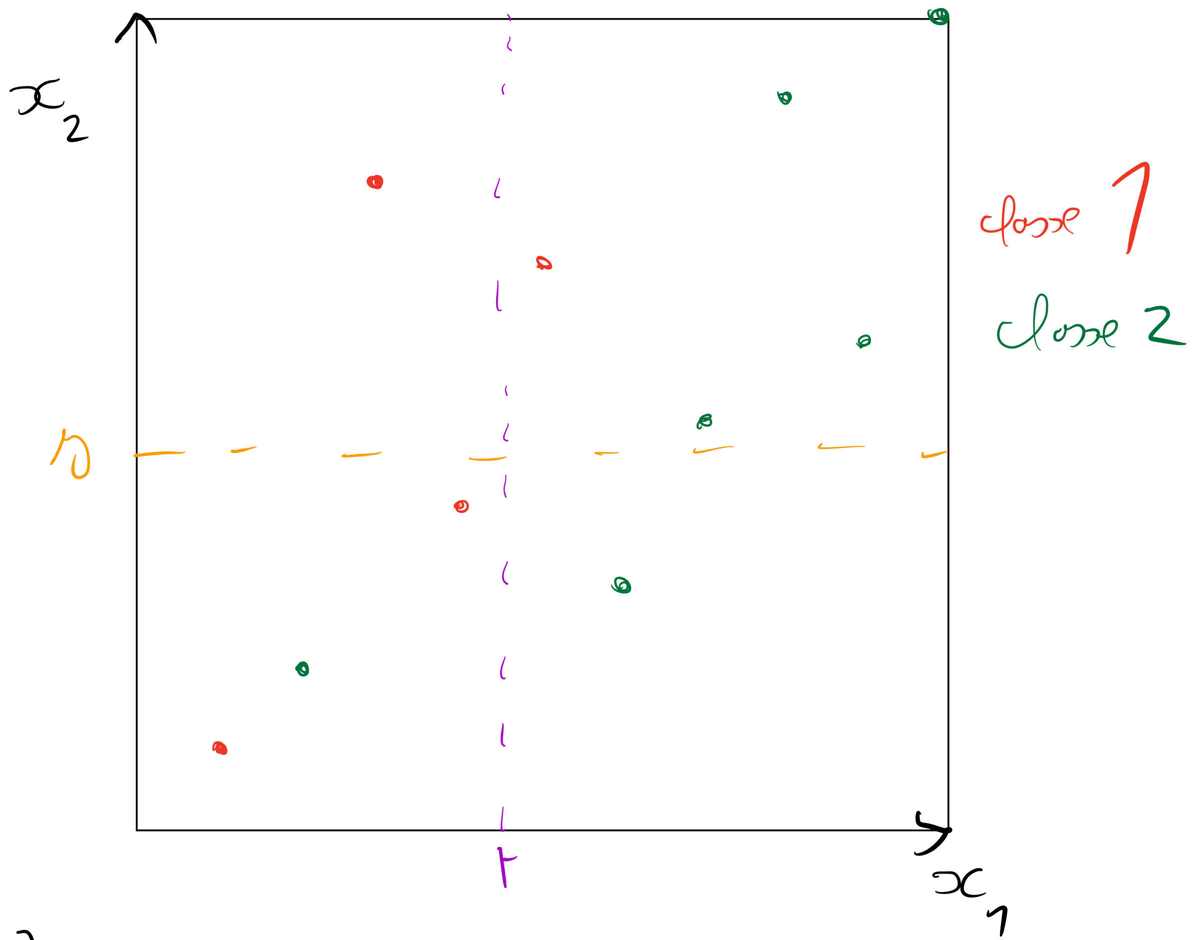
$$\simeq 0.95$$

3) $N = 12$ $\quad \hat{p}_1 = \frac{5}{12}$ $\quad \hat{p}_2 = \frac{7}{12}$

$$E = -\frac{5}{12}\log\left(\frac{5}{12}\right) - \frac{7}{12}\log\left(\frac{7}{12}\right)$$

$\simeq 0.68$

Exercice 2



1) $E = E(y_1, \dots, y_{10})$

$\hat{p}_1 = \frac{4}{10} \qquad \hat{p}_2 = \frac{6}{10}$

$$E = -\frac{4}{10}\log\left(\frac{4}{10}\right) - \frac{6}{10}\log\left(\frac{6}{10}\right)$$

$\simeq 0.67$

2)

$$C_1 = \{ i ; x_1^{(i)} \leq t \}$$

$$= \{ 1, 2, 3, 4 \}$$

$$C_2 = \{ i ; x_1^{(i)} > t \}$$

$$= \{ 5, 6, 7, 8, 9, 10 \}$$

$$E_t = \frac{4}{10} E(y_1, \dots, y_4) + \frac{6}{10} E(y_5, \dots, y_{10})$$

$$= \frac{4}{10} \left[-\frac{3}{4} \log\left(\frac{3}{4}\right) - \frac{1}{4} \log\left(\frac{1}{4}\right) \right] + \frac{6}{10} \left[-\frac{1}{6} \log\left(\frac{2}{6}\right) - \frac{5}{6} \log\left(\frac{5}{6}\right) \right]$$

$$\simeq 0.49$$

3)

$$C_1 = \{ i ; x_2^{(i)} \leq s \}$$

$$= \{ 1, 2, 4, 6 \}$$

$$C_2 = \{ i ; x_2^{(i)} > s \}$$

$$= \{ 3, 5, 7, 8, 9, 10 \}$$

$$E_t = \frac{4}{10} E(y_1, y_2, y_4, y_6) + \frac{6}{10} E(y_3, y_5, y_7, y_8, y_9, y_{10})$$

$$= \frac{4}{10} \left[-\frac{2}{4} \log\left(\frac{2}{4}\right) - \frac{2}{4} \log\left(\frac{2}{4}\right) \right] + \frac{6}{10} \left[-\frac{2}{6} \log\left(\frac{2}{6}\right) - \frac{4}{6} \log\left(\frac{4}{6}\right) \right]$$

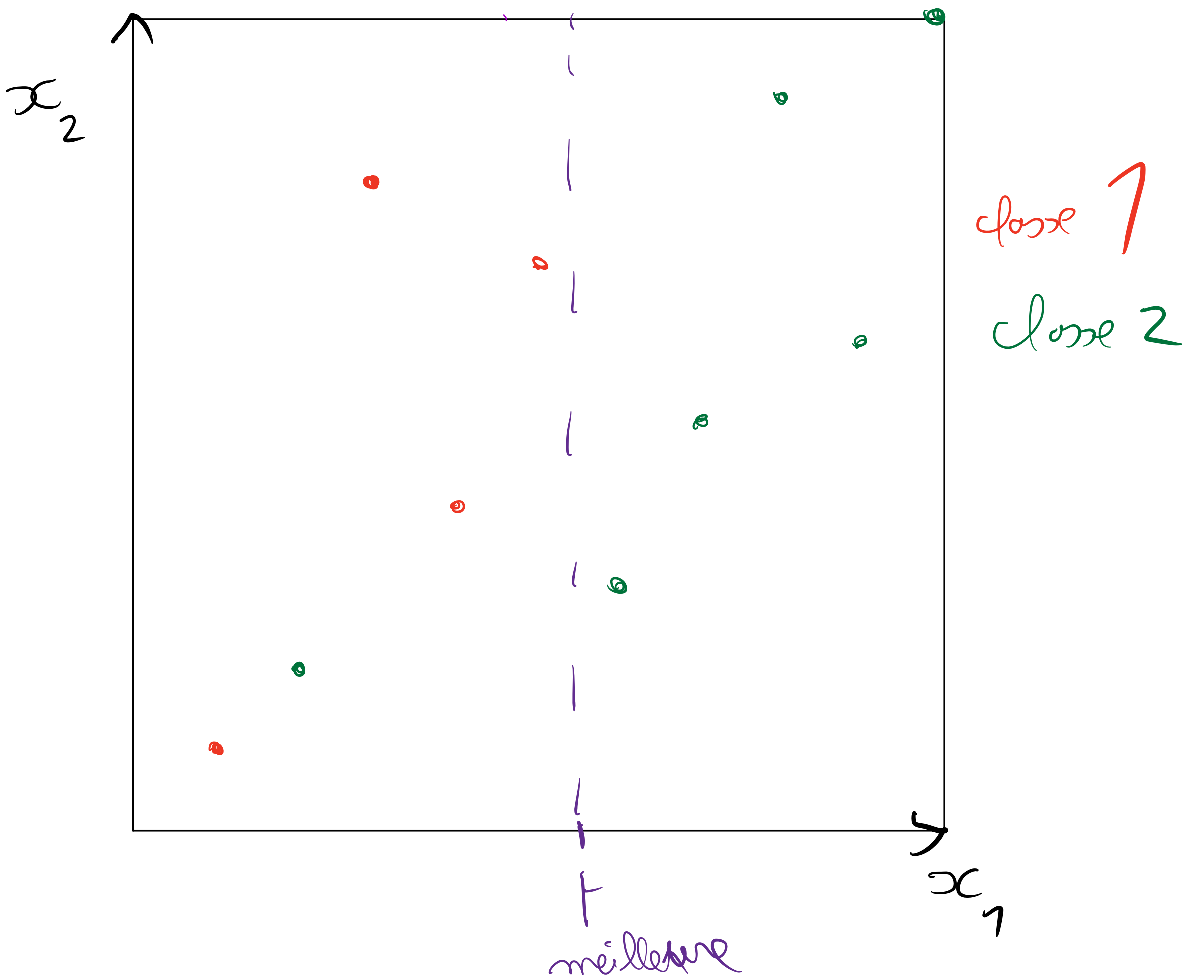
$$\simeq 0.66$$

4) On préfère le premier classifieur

y classe 1 ssi $x_1 \leq t$

car il sépare mieux les classes 1 et 2
En effet "$x_1 \leq t$" correspond empiriquement à presque que des classes 1 dans la base d'apprentissage et "$x_2 > t$" presque seulement des classes 2.

5)



Selon l'axe des x_1

le meilleur t

t	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5
entropie	0.67	0.57	0.66	0.67	0.67	0.25	0.38	0.47	0.55
t	9.5	10.5							
	0.67	0.67							

Exemple de calcul pour $t = 2.5$

groupe 1 : $\{1, 2\}$

groupe 2 : $\{1, 1, 1, 2, 2, 2, 2, 2\}$

$$\frac{k}{10} E(u_1, \dots, u_k) + \frac{l}{10} E(v_1, \dots, v_l)$$

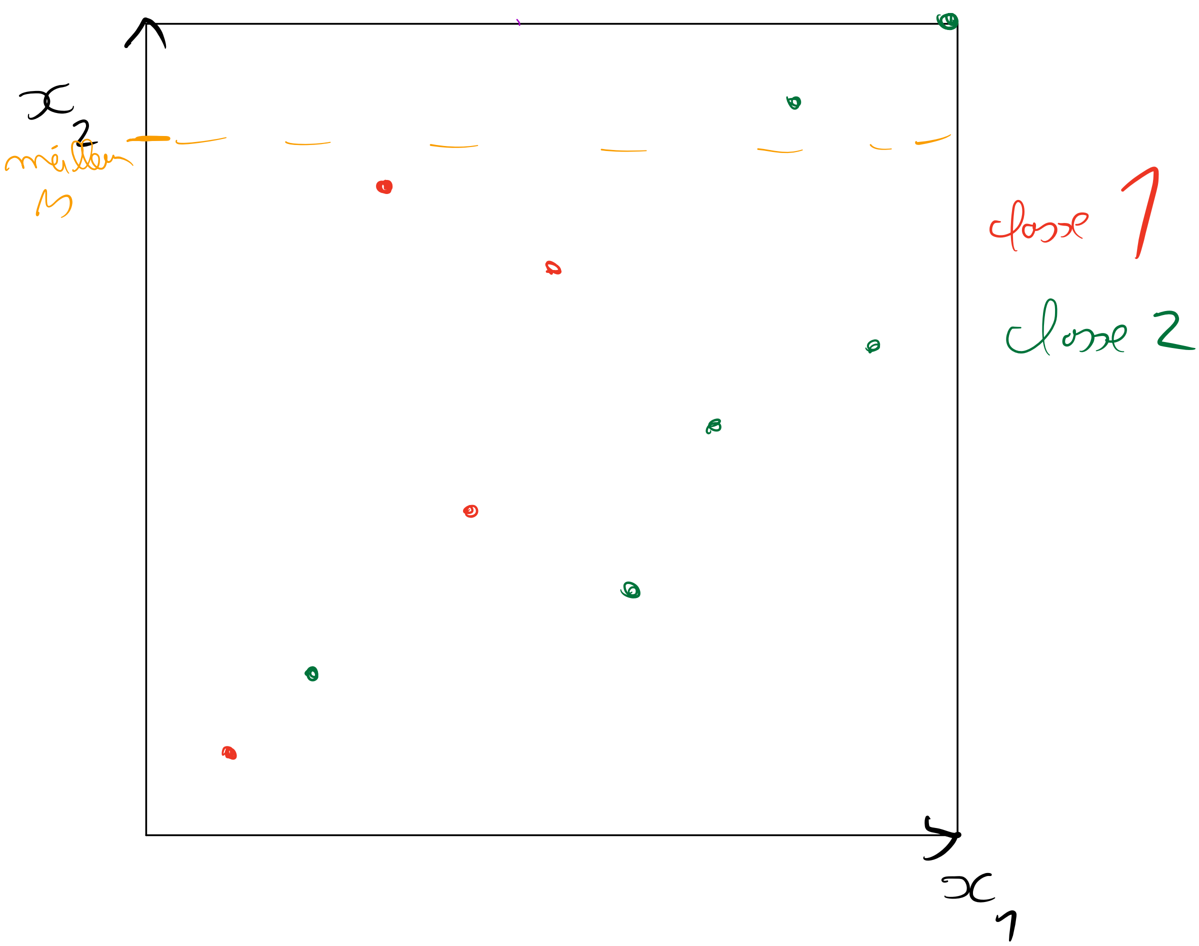
$$= \frac{2}{10}\left[\frac{1}{2}\log\left(\frac{1}{2}\right) + \frac{1}{2}\log\left(\frac{1}{2}\right)\right] + \frac{8}{10}\left[\frac{3}{8}\log\left(\frac{3}{8}\right) + \frac{5}{8}\log\left(\frac{5}{8}\right)\right]$$

$$\simeq 0.66$$

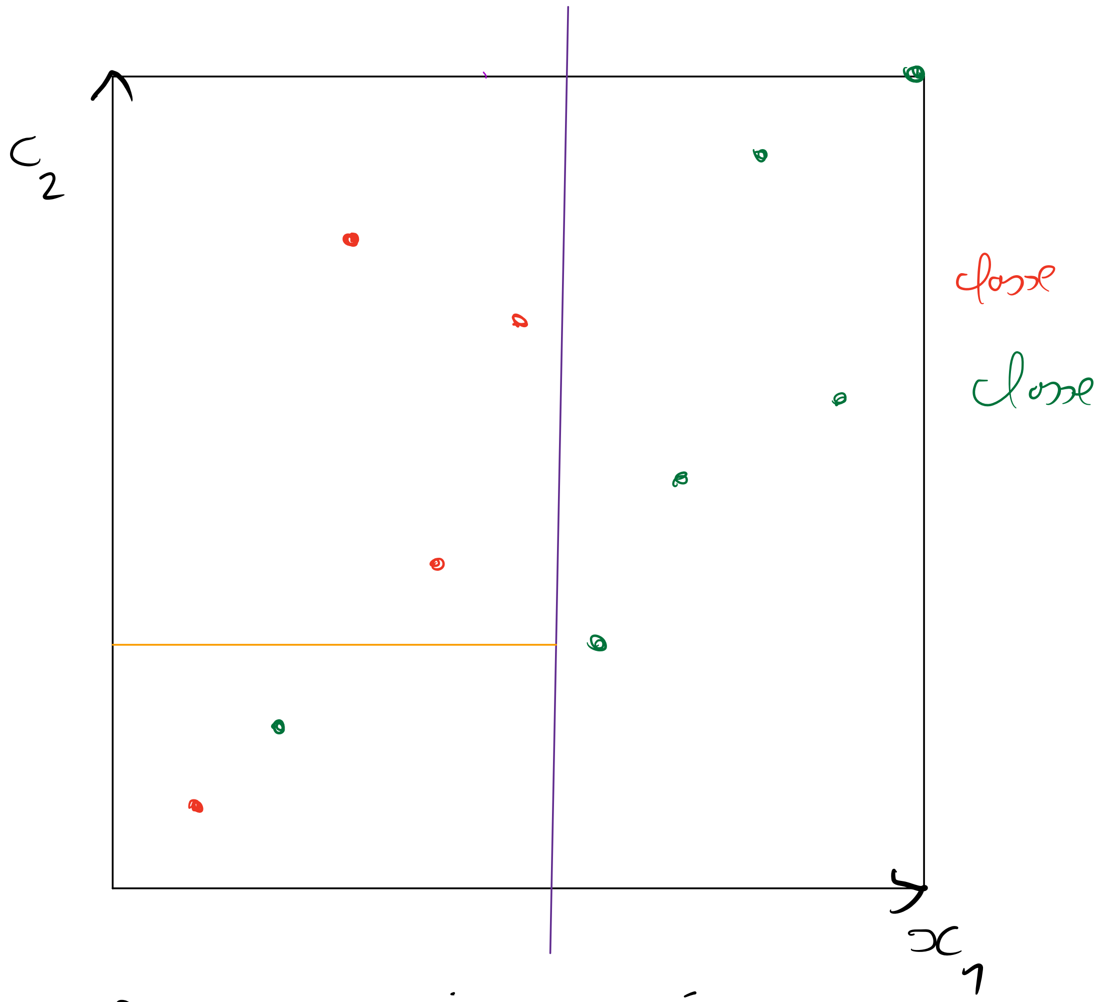
Selon l'axe des x_2

Meilleur s

s	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5
entropie	0.67	0.57	0.66	0.66	0.66	0.67	0.66	0.67	0.55
s	9.5	10.5							
	0.62	0.67							



On choisit de couper selon x_1



Il n'y a rien à faire pour le rectangle de droite.

A gauche on peut calculer que l'optimum sera comme ceci par exemple

A la fin on a 3 rectangles.