Partiel de FTML

Exercice 1 :

a) Matrices de confusions :

pour Y_EST_1 :

$$\begin{array}{c|cc} \text{réalités / prédictions} & T & F \\ \hline T & 6 & 8 \\ F & 1 & 2 \end{array}$$

pour Y_EST_2 :

$$\begin{array}{c|cc} \text{réalité / prédictions} & T & F \\ \hline T & 4 & 1 \\ F & 3 & 9 \end{array}$$

Donc les risques empiriques sont :

$$ERM(Y\_EST\_1) = \boxed{\frac{8+x}{17}}$$

$$ERM(Y\_EST\_2) = \boxed{\frac{3x+1}{17}}$$

Donc les estimateurs se valent si $x = 3,5$, si $x > 3,5$ alors l'estimateur 2 aura un risque empirique plus grand sinon c'est l'estimateur 1.

d) Tableaux agregés:

pour Y = True :

X_1 \ X_2	1	2
1	1	4
2	1	1

pour Y = False :

X_1 \ X_2	1	2
1	2	1
2	4	3

e) $P(Y=T \mid x_1=1 \text{ et } x_2=1) = \frac{1}{7} \times \frac{7}{17} = \frac{1}{17}$

$P(Y=F \mid x_1=1 \text{ et } x_2=1) = \frac{10}{17} \times \frac{2}{10} = \frac{2}{17}$

$P(Y=T \mid x_1=2 \text{ et } x_2=1) = \frac{1}{\cancel{17}} \times \frac{7}{17} = \frac{1}{17}$

$P(Y=F \mid x_1=2 \text{ et } x_2=1) = \frac{4}{10} \times \frac{10}{17} = \frac{4}{17}$

$P(Y=T \mid x_1=1 \text{ et } x_2=2) = \frac{4}{7} \times \frac{7}{17} = \frac{4}{17}$

$P(Y=F \mid x_1=1 \text{ et } x_2=2) = \frac{1}{10} \times \frac{10}{17} = \frac{1}{17}$

$P(Y=T / x_1=2 \text{ et } x_2=2) = \frac{1}{7} \times \frac{7}{17} = \frac{1}{17}$

$P(Y=F \mid x_1=2 \text{ et } x_2=2) = \frac{3}{10} \times \frac{10}{17} = \frac{3}{17}$

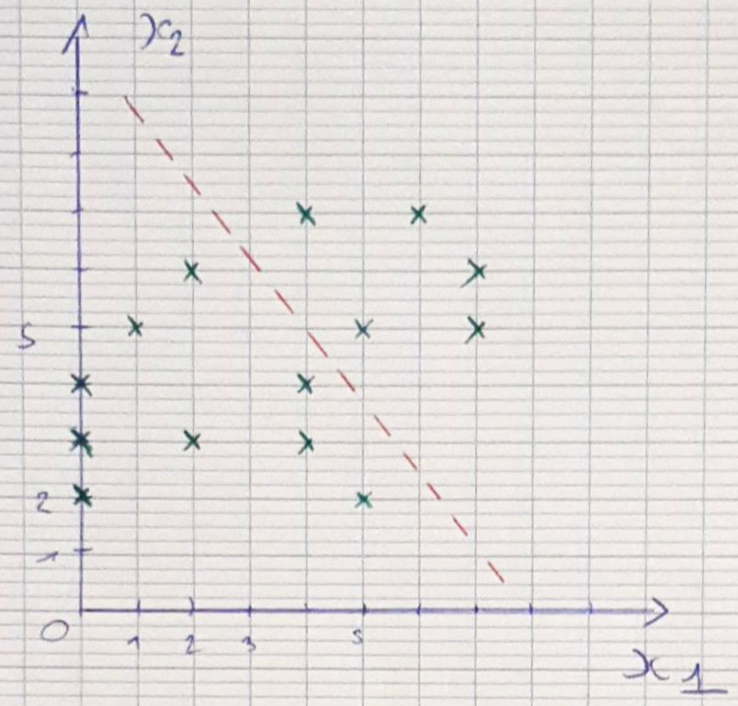
On a donc :

$x_1 \backslash x_2$	1	2
1	False	True
2	False	False

Representant notre estimateur bayesien naif optimal.

Exercice 2 :

a)



Les deux composantes probables observées sont séparées par la droite pointillée rouge.

Exercice 3:

a) Pour tester ~~d~~ l'independance de deux variables on peut utiliser le test du khi2. Si le khi2 calculé est élevé les variables ne sont pas indépendantes.

b) La dimension de Vapnik-Chervonenkis correspond au nombre maximum de points pulverisables on pourrait donc l'utiliser pour determiner le nombre maximum de points pulverisables à la surface d'une sphère.

c)