Exc 1

a) $\dfrac{(FP \times 1 + FN \times x)}{\text{nb element}}$ = risque empirique

~~Est 1:~~

Est 1 : $\frac{1}{17} \times (8 \times 1 + 1 \times x) = \frac{8+x}{17}$

Est 2 : $\frac{1}{17} \times (1 \times 1 + 3 \times x) = \frac{1+3x}{17}$

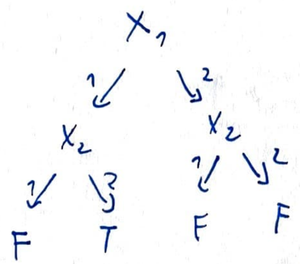
Si $x < \frac{7}{2}$ alors Est 2 est plus performant

$x = \frac{7}{2}$ alors les deux estimateurs sont aussi performants

$x > \frac{7}{2}$ alors Est 1 est plus performant

b) $\frac{1}{nb}(FP \times 1 + FN \times 2)$ = risque empirique

Est custom:



Pour (1, 1), F et T équivalent

$\frac{1}{17}(2 + 2 + 1 + 1) = \frac{6}{17}$ risque empirique

c)

- On ne sais pas ~~comment~~ si les variables sont ~~colléssé~~ corélé
- les données sont souvent contradictoire

d)

$Y=T$	1	2
x_1	5	2
x_2	2	5

$Y=F$	1	2
x_1	3	7
x_2	6	9

e) ~~[illegible]~~

$$P(Y=T \mid x_1=1 \ \& \ x_2=1) = \frac{P(X_1=1 \mid Y=T) \times P(X_2=1 \mid Y=T) \times P(Y=T)}{P(X_1=1 \ \& \ X_2=1)}$$

$$= \frac{\frac{5}{7} \times \frac{2}{7} \times \frac{7}{17}}{\frac{3}{17}}$$

$$= \frac{5}{7} \times \frac{2}{7} \times \frac{7}{17} \times \frac{17}{3} = \frac{10}{21}$$

$$P(Y=F \mid x_1=1 \ \& \ x_2=1) = \frac{P(X_1=1 \mid Y=F) \times P(X_2=1 \mid Y=F) \times P(Y=F)}{P(X_1=1 \ \& \ X_2=1)}$$

$$= \frac{3}{10} \times \frac{7}{10} \times \frac{10}{17} \times \frac{17}{3} = \frac{7}{10} = 0{,}7$$

Donc, comme $0{,}7 > \frac{10}{21}$; $x_1=1$ & $x_2=1$ renvoie F ~~positif~~

~~De même~~

~~$P(Y=T \mid X_1=1 \ \& \ x_2=2) = \frac{10}{[illegible]}$~~

~~$P(Y=F \mid X_1=1 \ \& \ x_2=2) = [illegible]$~~ T

~~$P(Y=T \mid x_1=2 \ \& \ x_2=1) = [illegible]$~~

~~$P(Y=F \mid x_1=2 \ \& \ x_2=1) =$~~ F

De même

$P(Y = T \mid X_1 = 1 \;\&\; X_2 = 2) > P(Y = F \mid X_1 = 1 \;\&\; X_2 = 2)$

$X_1 = 1 \;\&\; X_2 = 2$ renvoie T

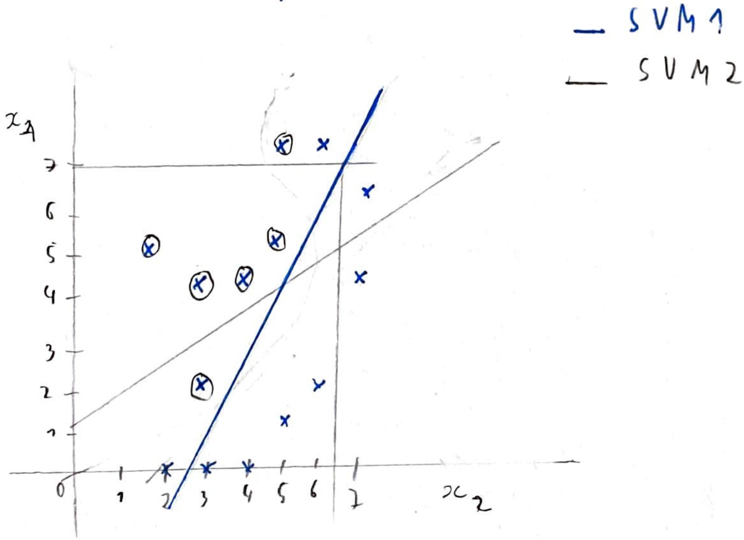
$X_1 = 2 \;\&\; X_2 = 1$ renvoie F

$X_1 = 2 \;\&\; X_2 = 2$ renvoie F

Donc l'estimateur Bayesien naïf renvoie les même résultat que l'estimateur optimal

Ex 2

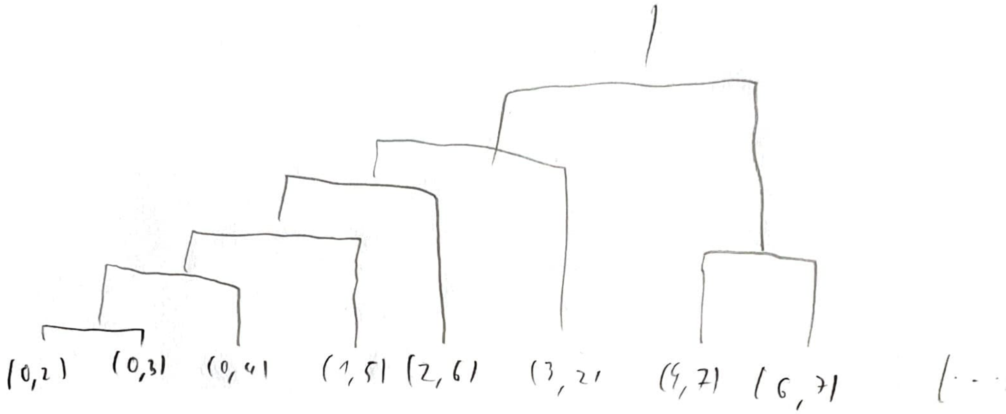
a)



D'après la doc Scikit learn, - l'Agglomerative Clustering
- le DBscan
- l'OPTICS
- Spectral Clustering

sont des méthode adapté

b) En prenant la distance de Manhattan



Pas le temp de finir

c) Ce n'est pas possible de parfaitement séparer par un SVM linéaire

d) Soit le SVM 1 avec un fort coef de pénalisation (ie : 800)
Soit le SVM 2 " " faible " " " (ie : 1)

e) Un kernel polynomial pourrait séparer les deux classes

Exc 3.

a) Test des moindres carrés

b) Je testerai toute les pulvérisation sur ~~un~~ des polyèdres

c) Ambiguïté : Données confuse
Risque : Danger quantifié