Partiel FTML

Ivan LIN
SCIA
2021

exercice 1

a) Y_EST_1 :

6	1
8	2

$$\tilde{D}(\hat{f}, \tau) = \frac{1}{17}(1 \times 8 + x \times 1) = \frac{x+8}{17}$$

Y_EST_2 :

4	3
1	9

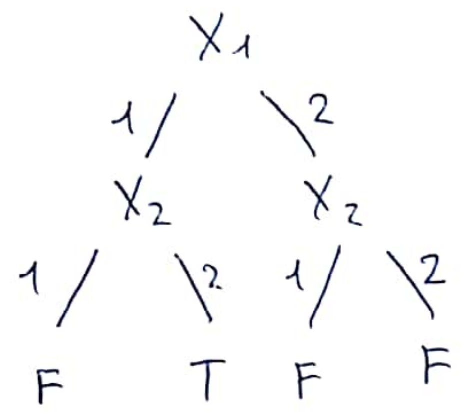
$$\tilde{D}(\hat{f}, \tau) = \frac{1}{17}(1 \times 1 + x \times 3) = \frac{3x+1}{17}$$

$$\frac{x}{17} + \frac{8}{17} \geqslant \frac{1}{17} + \frac{3x}{17} \iff \frac{7}{17} \geqslant \frac{2x}{17}$$

$$\iff \frac{7}{2} \geqslant x$$

Ainsi : Pour $x \in]-\infty ; \frac{7}{2}[$: Y_EST_1 > Y_EST_2

et pour $x \in [\frac{7}{2} ; +\infty[$: Y_EST_1 < Y_EST_2 b)



$\longrightarrow$ risque empirique :

$\frac{1}{17}(3\times 2 + 1\times 1) = \frac{7}{17}$

c)

d) $y = \text{true}$

	1	2
X_1	5	2
X_2	2	5

$y = \text{false}$

	1	2
X_1	3	7
X_2	6	4

e) Pour $y = \text{True}$

(X_1, X_2)

$(1,1)$. $\frac{7}{17}\times\frac{5}{7}\times\frac{2}{7} = \frac{10}{119}$

$(1,2)$ $\frac{7}{17}\times\frac{5}{7}\times\frac{5}{7} = \frac{25}{119}$

$(2,1)$ $\frac{7}{17}\times\frac{2}{7}\times\frac{2}{7} = \frac{4}{119}$

$(2,2)$. $\frac{7}{17}\times\frac{2}{7}\times\frac{5}{7} = \frac{10}{119}$

Pour $y = \text{fals}$

(X_1, X_2)

$(1,1)$ $\frac{10}{17}\times\frac{3}{10}\times\frac{6}{10} = \frac{9}{85}$

$(1,2)$ $\frac{10}{17}\times\frac{3}{10}\times\frac{4}{10} = \frac{6}{85}$

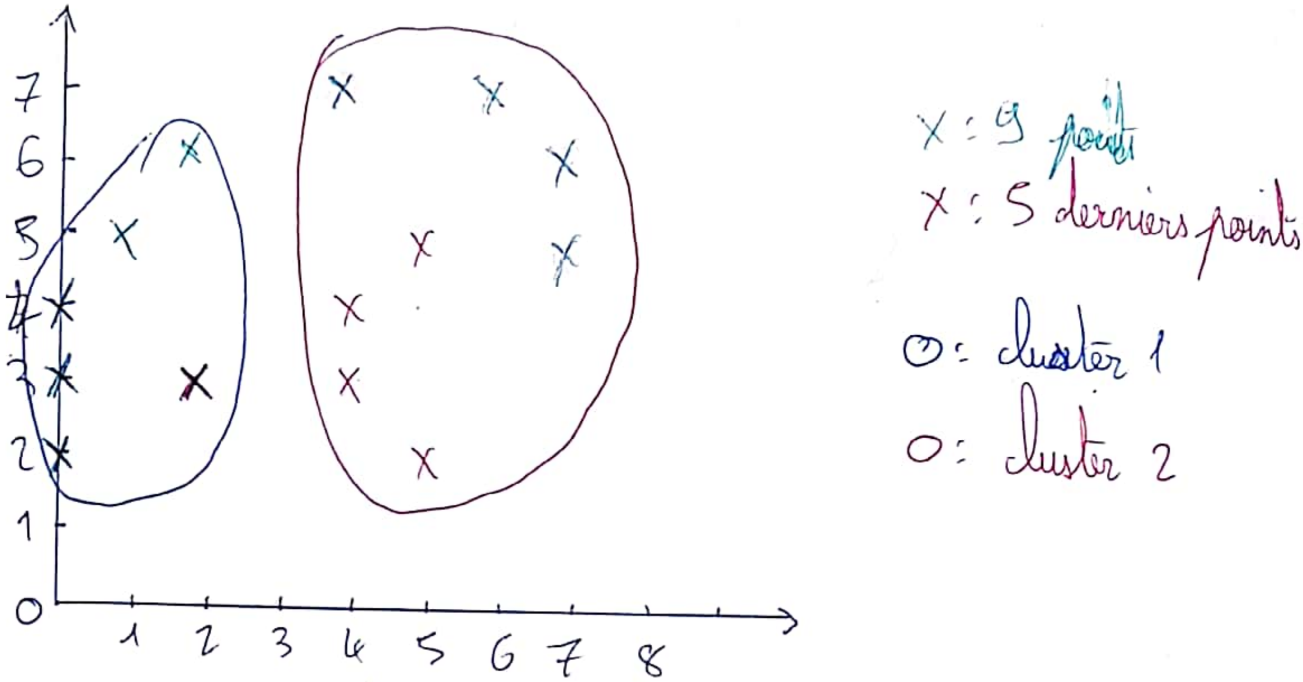
$(2,1)$ $\frac{10}{17}\times\frac{7}{10}\times\frac{6}{10} = \frac{21}{85}$

$(2,2)$ $\frac{10}{17}\times\frac{7}{10}\times\frac{4}{10} = \frac{14}{85}$

	1	2
1	False	True
2	False	False

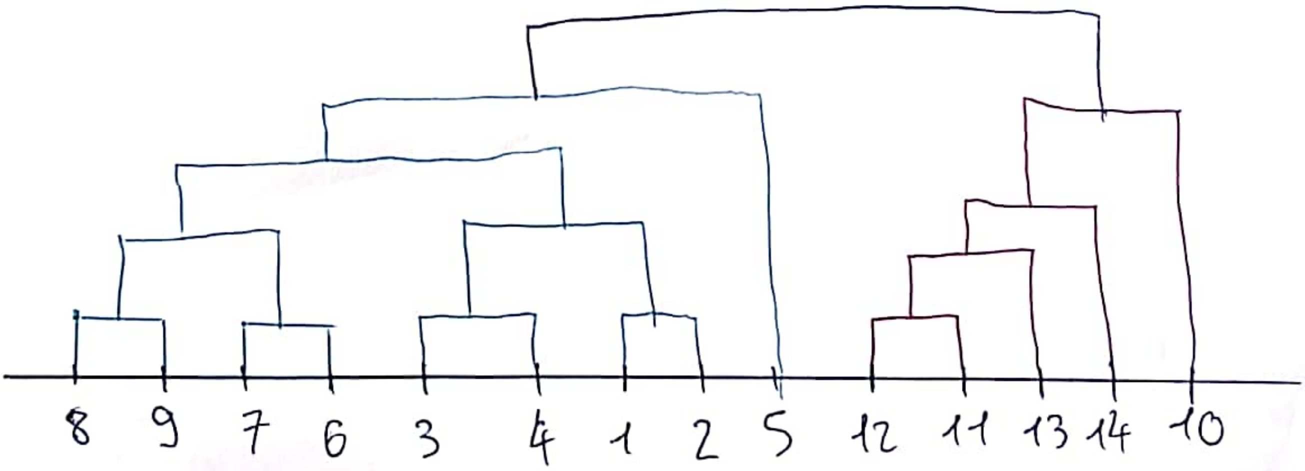
Exercice 2

a)



On utiliserait un kmeans ou une classification hiérarchique

b)



exercice 3

a) Pour savoir si les deux variables concernées par un tableau croisé d'effectifs sont independant on peut utiliser le test du x^2

- calcul des effectifs croisés

$$E_{i,j} = \frac{1}{n} \sum card(\{X = i\}) \, card(\{Y = j\})$$

- calcul des effectifs ~~ce~~ croisé observés

$$O_{i,j} = card(\{X = i \cap Y = j\})$$

- et on peut calculer l'écart relatif en les deux variables

$$T = \sum_{i,j} \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

avec un T faible cela signifie que les varibles sont independants

b) Il est possible de calculer la dimension de VC.
Cette demension corresponds au nombre maximum d'éléments pulvérisable.