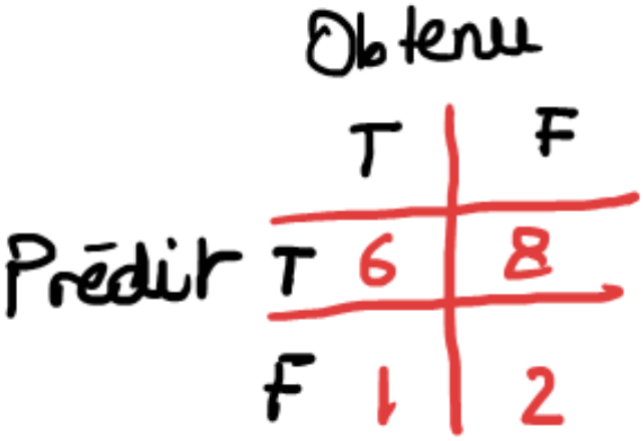
FTML

Exercice 1

a)

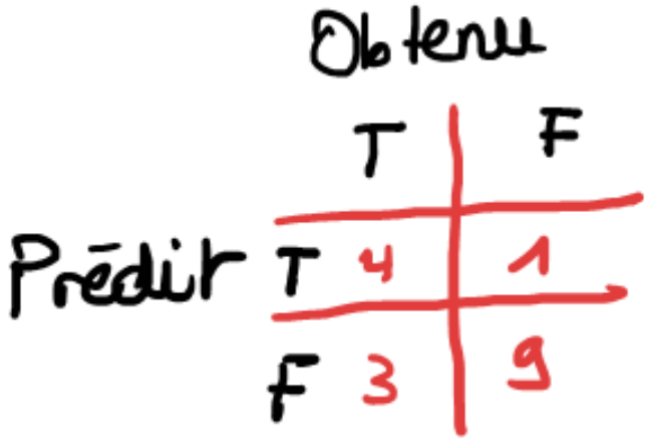
Matrice de confusion pour Est_1:



Risque empirique de Est_1:

$LF(\tilde{f}, f)$= 8*1 + x*1

Matrice de confusion pour Est_2



Risque empirique de Est_2:

$LF(\tilde{f}, f)$= 1*1 + x*3

On peut changer la valeur x pour donner l'avantage à un estimateur ou à l'autre:

Si on veut avantager l'estimateur 1 on a:

- $x + 8 < 3x + 1 \iff$ x > $\frac{7}{2}$

Et donc si on veut avantager l'estimateur 2 on a:

- $x + 8 > 3x + 1 \iff$ x < $\frac{7}{2}$

b)

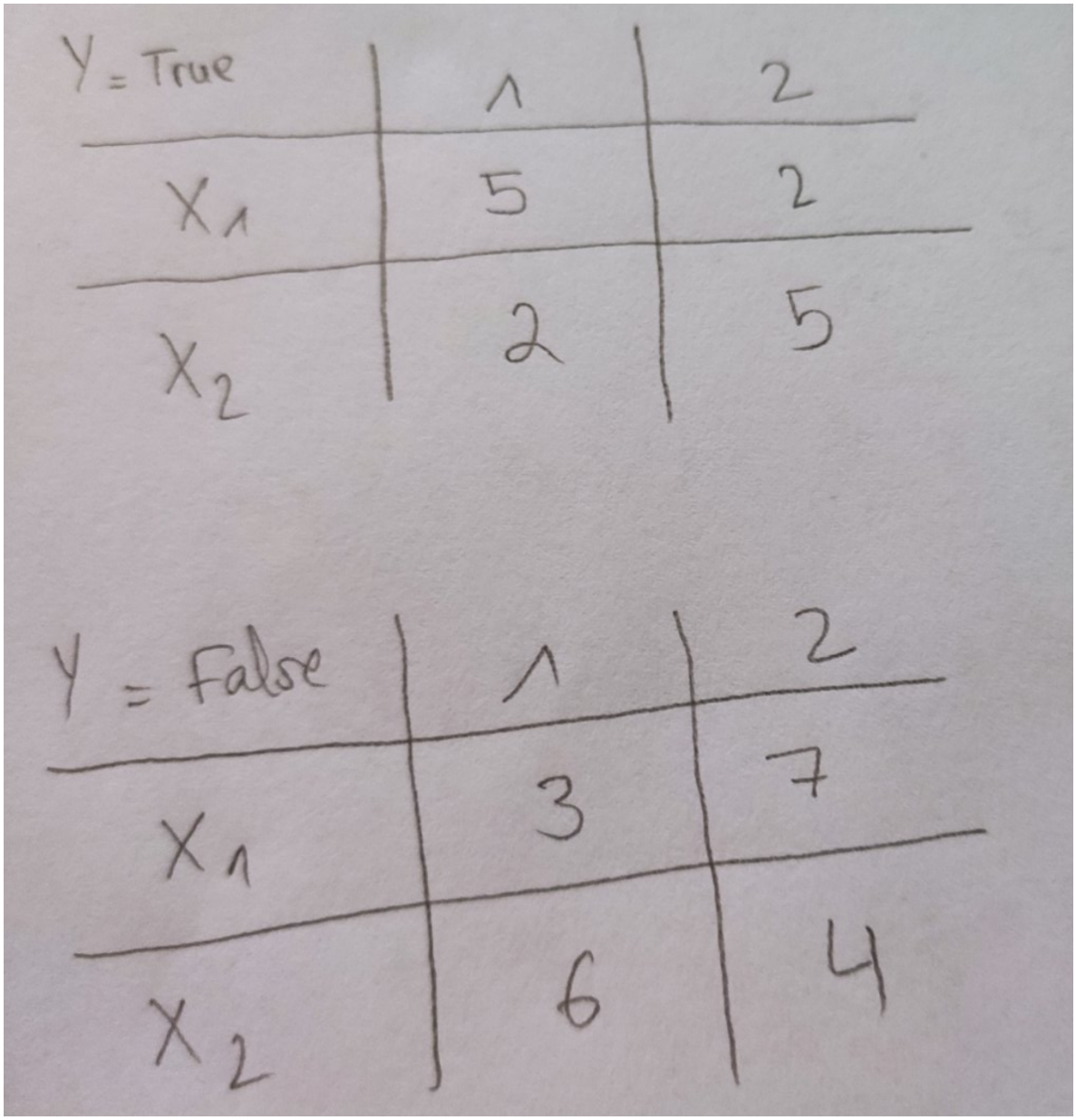
Un estimateur qui minimise le risque empirique est un estimateur pour lequel l'erreur est nulle. C'est le cas si notre estimateur correspond à la bonne réponse (et donc la colonne y)

Toutefois sur les deux estimateurs fourni, puisque x = 2, on sait que c'est l'estimateur 2 qui a le risque empirique le plus faible.

c)

On pourrait se placer dans le cadre bayésien naif, toutefois cela n'est pas nécessaire puisque

d)



e)

c)

$P(Y=T / X_1=1, X_2=1) = P(Y=T) \times P(X_1=1 / Y=T) \times P(X_2=1 / Y=T)$

$= \frac{7}{17} \times \frac{5}{7} \times \frac{2}{7} \simeq 0,084$

$P(Y=T / X_1=1, X_2=2) = \frac{7}{17} \times \frac{5}{7} \times \frac{5}{7} \simeq 0,210$

$P(Y=T / X_1=2, X_2=1) = \frac{7}{17} \times \frac{2}{7} \times \frac{2}{7} \simeq 0,034$

$P(Y=T / X_1=2, X_2=2) = \frac{7}{17} \times \frac{2}{7} \times \frac{5}{7} \simeq 0,084$

$P(Y=F / X_1=1, X_2=1) = \frac{10}{17} \times \frac{3}{10} \times \frac{6}{10} \simeq 0,106$

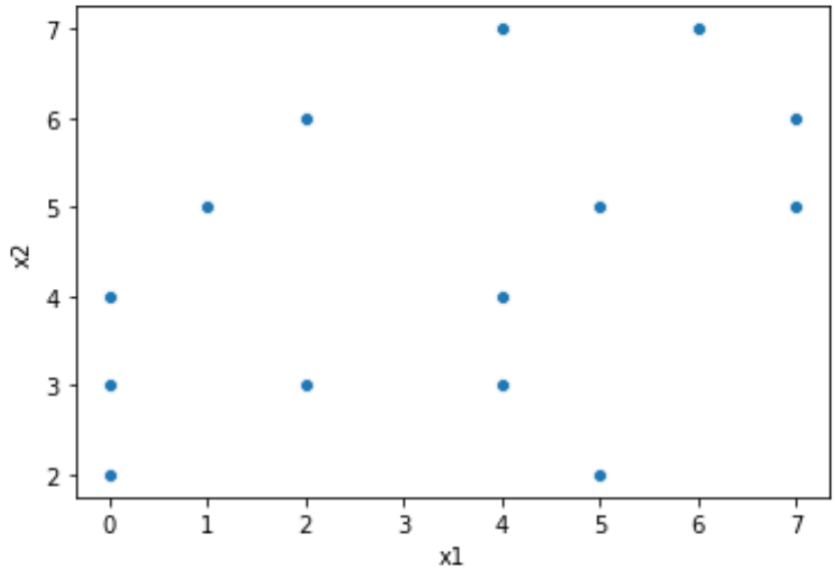
$P(Y=F / X_1=1, X_2=2) = \frac{10}{17} \times \frac{3}{10} \times \frac{4}{10} \simeq 0,071$

$P(Y=F / X_1=2, X_2=1) = \frac{10}{17} \times \frac{7}{10} \times \frac{6}{10} \simeq 0,247$

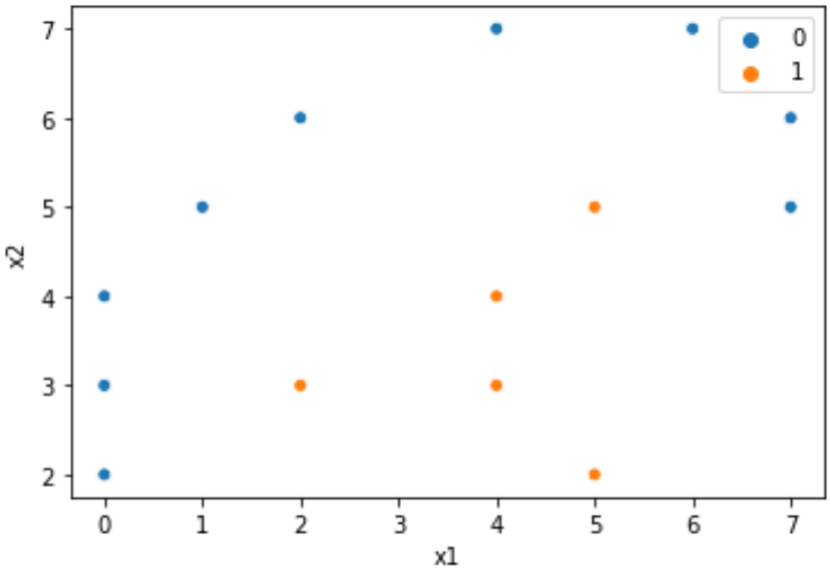
$P(Y=F / X_1=2, X_2=2) = \frac{10}{17} \times \frac{7}{10} \times \frac{4}{10} \simeq 0,165$

Exercice 2

a)



Si on connait le nombre de classe que l'on veut obtenir on peut faire un k-means ou de l'agglomerative clustering. En l'occurrence un spectral clustering serait assez efficace.



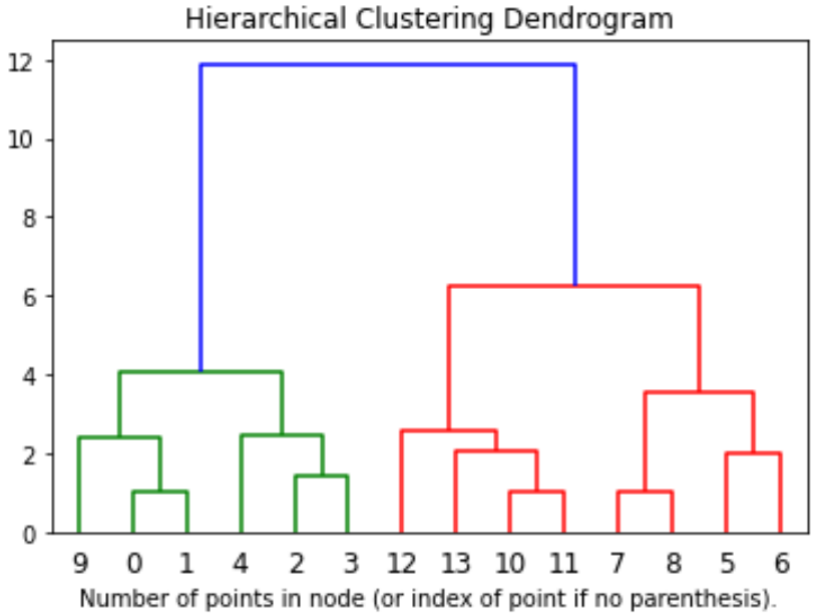
b)

On se munie d'une fonction de distance euclidienne pour faire la distance de Ward.

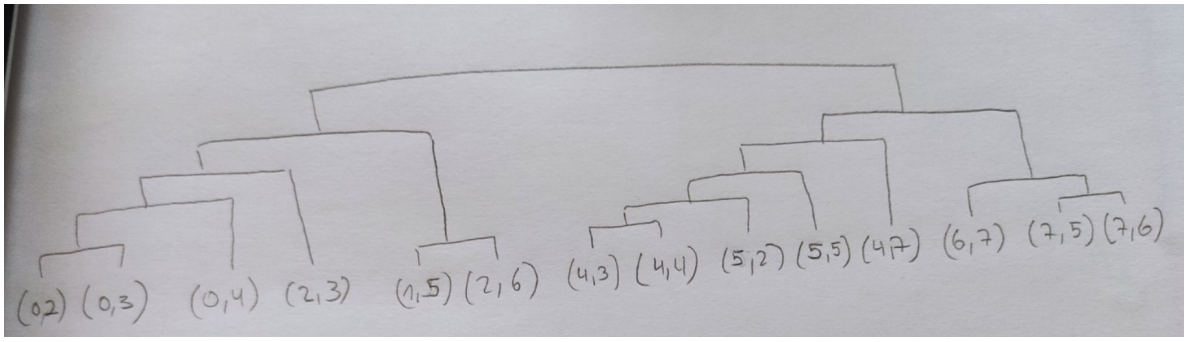
Les points sont les suivants:

```
array([[0, 2],
       [0, 3],
       [0, 4],
       [1, 5],
       [2, 6],
       [4, 7],
       [6, 7],
       [7, 6],
       [7, 5],
       [2, 3],
       [4, 3],
       [4, 4],
       [5, 2],
       [5, 5]])
```

Et le Dendograme associé est donc:



(Ou bien le suivant réaliser à la main)

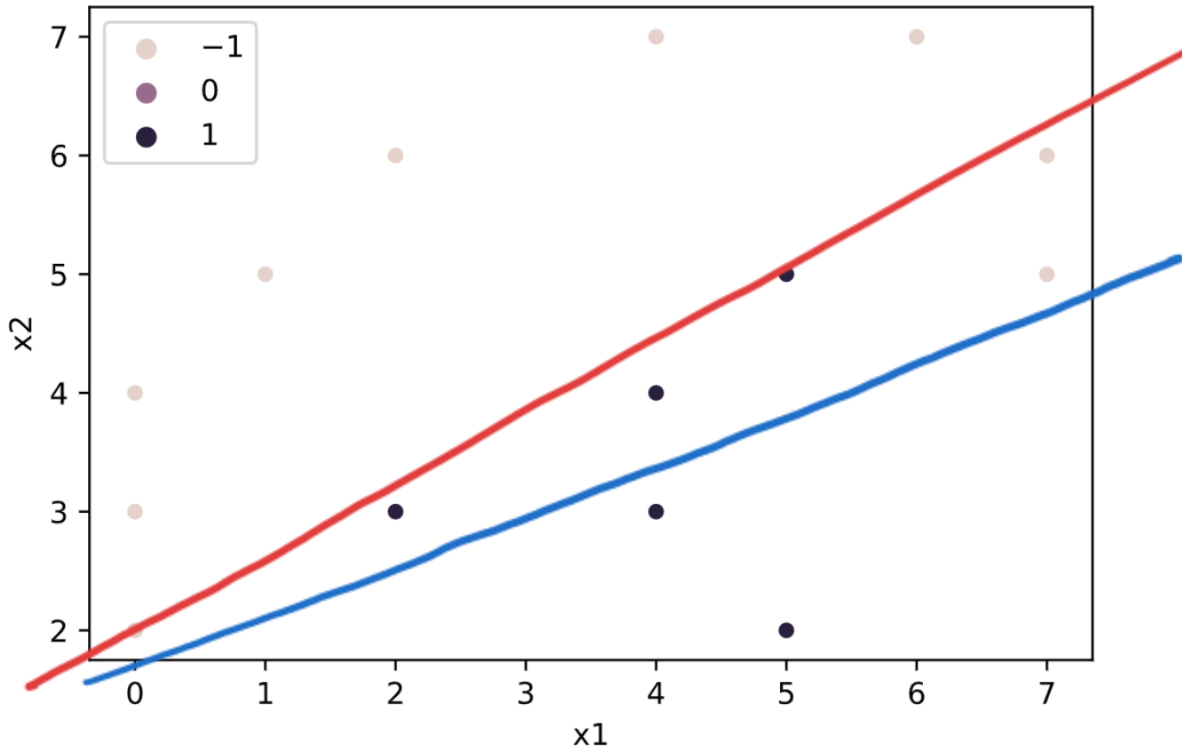


c)

On peut utiliser un SVM linéaire pour séparer le jeux de données en deux classes mais il est claire que la séparation ne serait pas parfaite et que l'on aurait des erreurs de prédiction. En effet un SVM linéaire n'est pas pratique pour séparer deux classe dont l'une est à l'intérieur de l'autre.

d)

On prend une pénalisation C assez grande pour la courbe rouge est moins grand pour la courbe bleue.



e)

Un cercle pourrait parfaitement séparer les deux classes. Ainsi on peut proposer la fonction f(x1,x2)= $x1^2 + x2^2$

Exercice 3

a)

A partir d'un tableau croisées d'effectifs je peut réaliser un test de chi2. Ce test consiste à tester l'existence d'une relation entre deux caractères discrets.

Il tout d'abord à réaliser un tableau d'effectifs théoriques, puis, à partir de ce tableau et du tableau d'effectifs croisées on obtient la matrice de chi2. Ensuite on détermine le degré de liberté (nombre de ligne * nombre de colonne).

La valeur de chi2 s'obtient en sommant les élément de la matrice de chi2.

On regarde alors dans la table de chi2 avec un risque alpha que l'on choisira et le degré de liberté obtenue plus tôt. On pourra alors rejeté ou non l'hypothèse nulle (hypothèse d'indépendance).

Si dans la table, $\sum Chi_{ij}^2$ > Chi^2(alpha, degré de liberté) alors les deux caractères ne sont pas indépendante avec alpha pourcent d'erreur.

b)

Je testerais en tâtonnant en commençant par exemple en essayant de pulvérisé 2, 3 ou 4 points sur la surface de la sphère puis en augmentant petit à petit. J’utiliserais aussi mes connaissance empirique pour mieux juger des nombre de points à tester.

c)

Risque : C'est le fait de connaitre des probabilité à travers des expériences et observations connues et déjà obtenues afin de traiter des probabilité inconnues.

Ambiguïté : On ne connais pas ou du moins ne comprends pas les probabilité des expériences car elles sont inconnues ou sujette à interprétation