

Exercice 1

Prédit

a)

| | F | T |
|---|---|---|
| F | 2 | 8 |
| T | 1 | 6 |

Y EST - 1

1 quand FP

| | F | T |
|---|---|---|
| F | 9 | 1 |
| T | 3 | 4 |

Y EST - 2

x quand FN

Pour l'estimateur 1 on a:

1 Faux négatif et 8 Faux positif

donc le risque empirique d'après la fonction de perte est :

$$1 + 8 \alpha$$

Pour l'estimateur 2 on a:

3 FN et 1 FP

comme pour l'estimateur 1, on en déduit que le risque empirique est :

$$3 + \alpha$$

$$> \frac{2}{7}$$

Pour α ~~> $\frac{2}{7}$~~ : l'estimateur 1 à un risque plus faible que l'estimateur 2.

$$< \frac{2}{7}$$

Pour α ~~< $\frac{2}{7}$~~ : l'estimateur 2 à un risque plus faible que l'estimateur 1

b) $x = 2$

2/6

On a 4 combinaisons de X_1 et X_2 possibles.

On liste les valeurs de Y du Tableau en fonction de celles-ci.

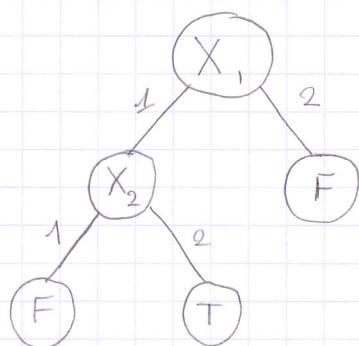
$X_1=1, X_2=2 \quad T F T T$

$X_1=1, X_2=1 \quad T F F$

$X_1=2, X_2=1 \quad F F F T F$

$X_1=2, X_2=2 \quad F F T F$

On choisit de construire un arbre de décision binaire comme estimateur:



- c) Le Bayésien naïf n'est pas utile dans ce cadre car pour chaque entrée possible de X_1 et X_2 , il y a une valeur de Y qui est plus fréquente.

Par exemple pour $X_1=1$ et $X_2=2$,

$$P(Y=T \mid X_1=1 \wedge X_2=2) = \frac{4}{5}$$

Ainsi on ne se retrouve pas en position d'indécision avec autant de T que de F .

C'est donc modélisable par un arbre de décision. On note également que le coût de calcul de l'arbre est plus faible.

d)

$$X_1 = 1 \text{ T T F T T F F T}$$

$$X_2 = 2 \text{ F F F T F F T F F}$$

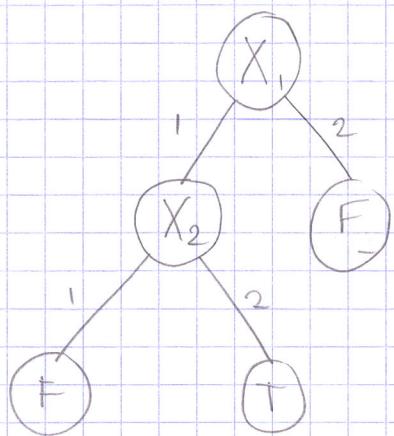
$$X_3 = 1 \text{ T F F F F T F F F}$$

$$X_4 = 2 \text{ T F F T F T T F T}$$

On a les tableaux suivants:

| | X_1 | X_2 | Y |
|-----------|-------|-----------|-----|
| $X_1 = 1$ | T | $X_2 = 1$ | F |
| $X_1 = 2$ | F | $X_2 = 2$ | F |

e) Les tableaux nous incitent à construire un arbre de décision. On a donc le même arbre que celui de la question b), sauf

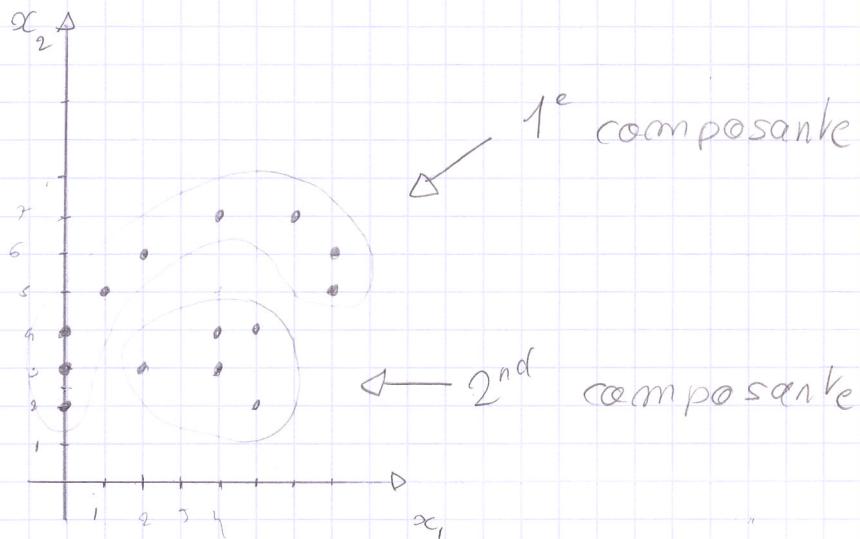


Etant indéfinie à l'estimation trouvée plus haut, il n'y a pas de comparaison possible.

Exercice 2

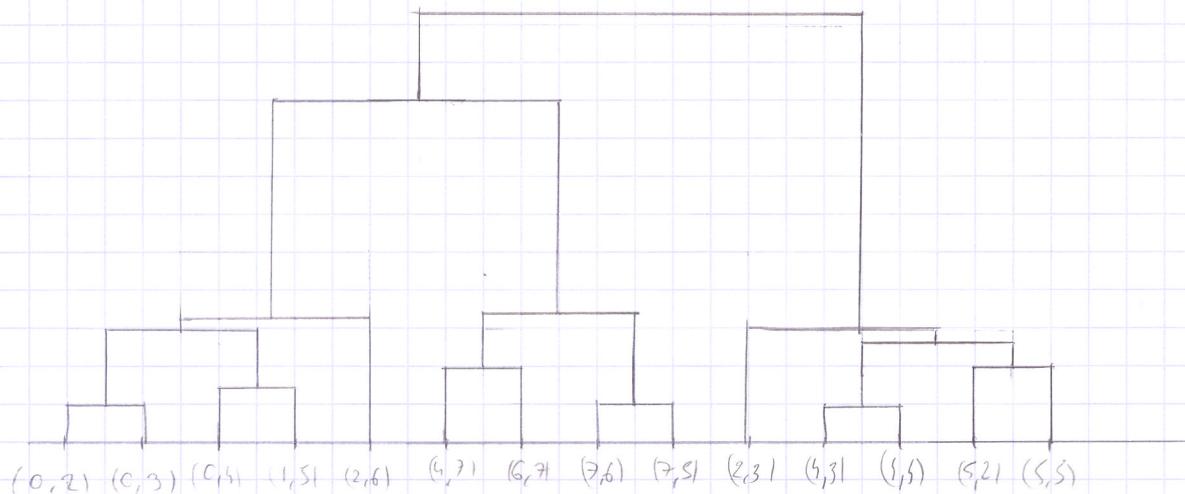
4/6

a)



On peut penser qu'un Regroupement hiérarchique permettrait d'identifier les 2 composantes.
(Agglomerative Hierarchical Clustering)

b)



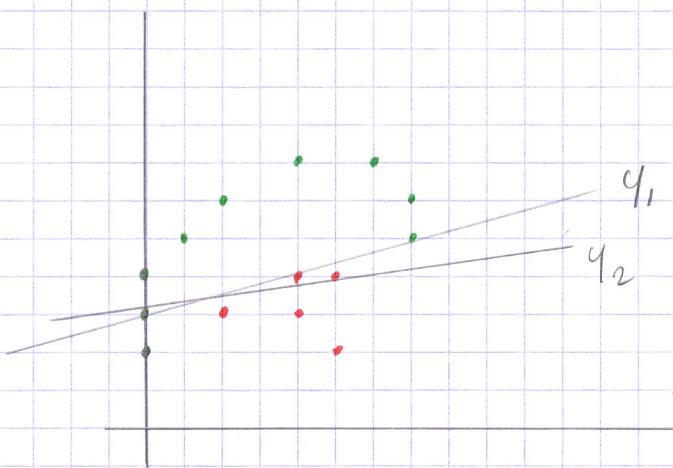
On a pris ici la distance comme fonction de coût.

c)

On ne peut séparer parfaitement les composantes par une droite, et donc, non-plus par un sum Linéaire. On peut en revanche s'approcher très fortement de cette séparation. Plusieurs points sont sur une même droite de séparation ce qui rend la décision complexe.

Exercice 2 - (Suite)

d)



y_1 correspond à un SVM avec un C assez grand et n'autorise que peu d'erreurs.
Ex: $C = 1000$

y_2 autorise plus d'erreurs car la pénalité est plus faible. Ex: $C = 1$

e)

On peut penser à 2 kernelisations, qui se ressemblent, permettant de séparer les composantes.

- Kernel polynomiale avec un degré ≥ 2 .
- Kernel sinusoïdale, qui dans notre cas s'adapte aussi à la séparation entre les composantes.

Exercice 3:

a)

Pour tester l'indépendance de 2 variables, on utilise généralement le test du Khi-2 (χ^2).
Dans la pratique on utilise communément le Test du Khi-2 de Pearson.

b) Pour déterminer le nombre maximum de point pouvérables à la surface d'une sphère, on cherchait à déterminer la dimension de Vapnik-Chervonenkis.

Après avoir poser le groupe d'objets sur lequel nous voulons pouvénir les point de la surface de la sphère, nous essayons de trouver un contre exemple qui nous servirait de "dimension bonne supérieure". Il s'agit ensuite de trouver le plus grand nombre de points pouvérable et inférieur à notre bonne.

c) Le risque est la possibilité de la réalisation d'un événement connu, et indéfinable.
L'ambiguité représente une certain manque d'information concernant les conditions de réalisation d'un événement.