

### Exercice 1 :

a.

Risque empirique pour  $EST\_1 = 8 * FP + 1 * FN = 8 + X$ .

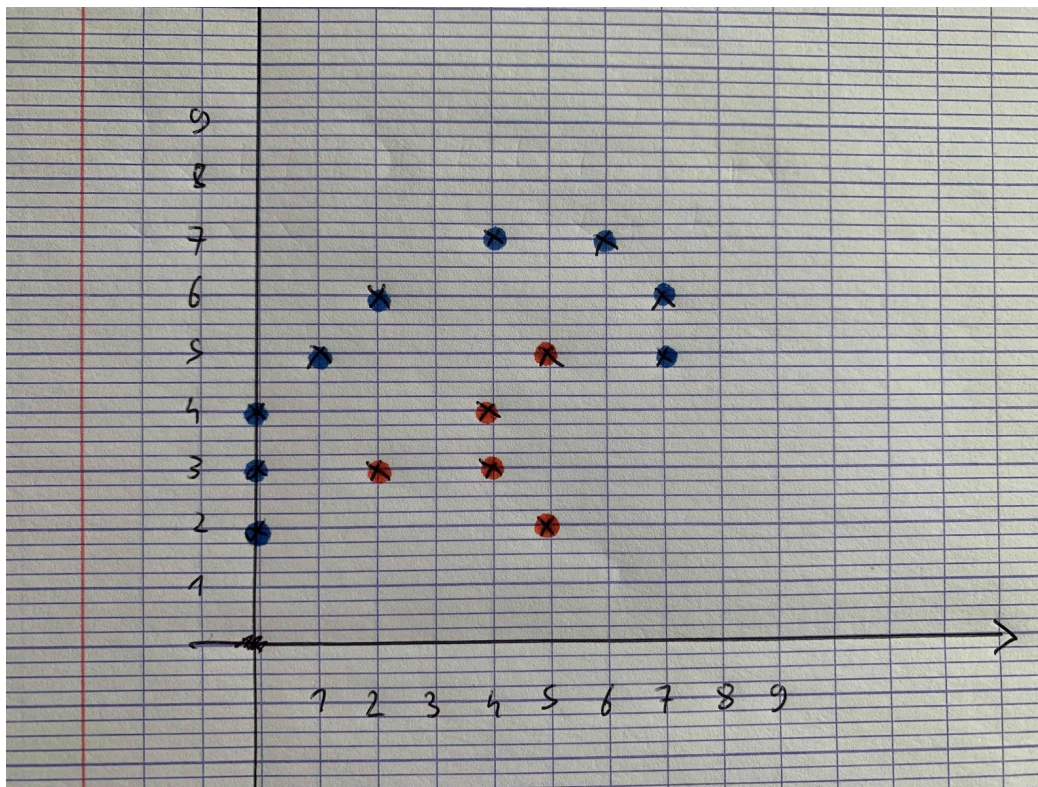
Risque empirique pour  $EST\_2 = 1 * FP + 3 + FN = 3X + 1$ .

On remarque donc que en fonction de la valeur de X un estimateur peut être plus performant que l'autre. Si  $X > 7/2$  alors  $EST\_1$  est plus performant, si  $X < 7/2$  alors  $EST\_2$  est plus performant et si  $X = 7/2$  alors les deux estimateurs sont équivalents.

b.

### Exercice 2 :

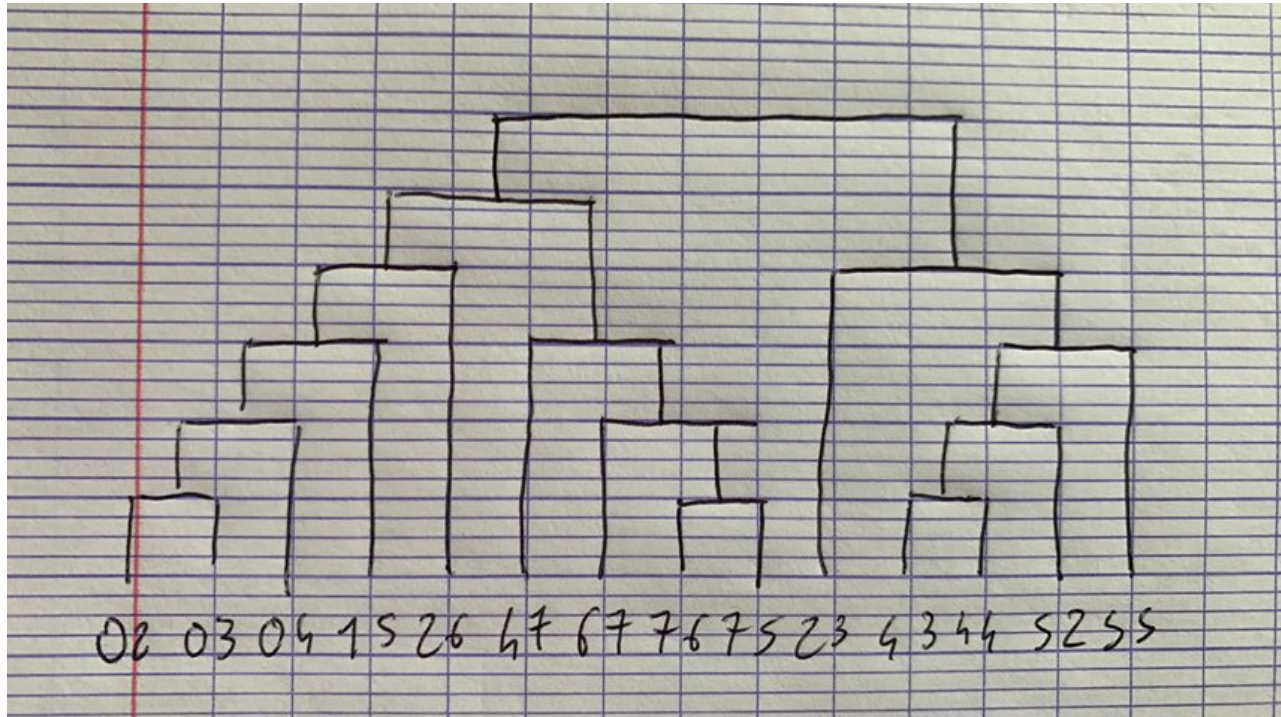
a.



La méthode non supervisée qui me paraît à même de les distinguer est un regroupement hiérarchique.

b.

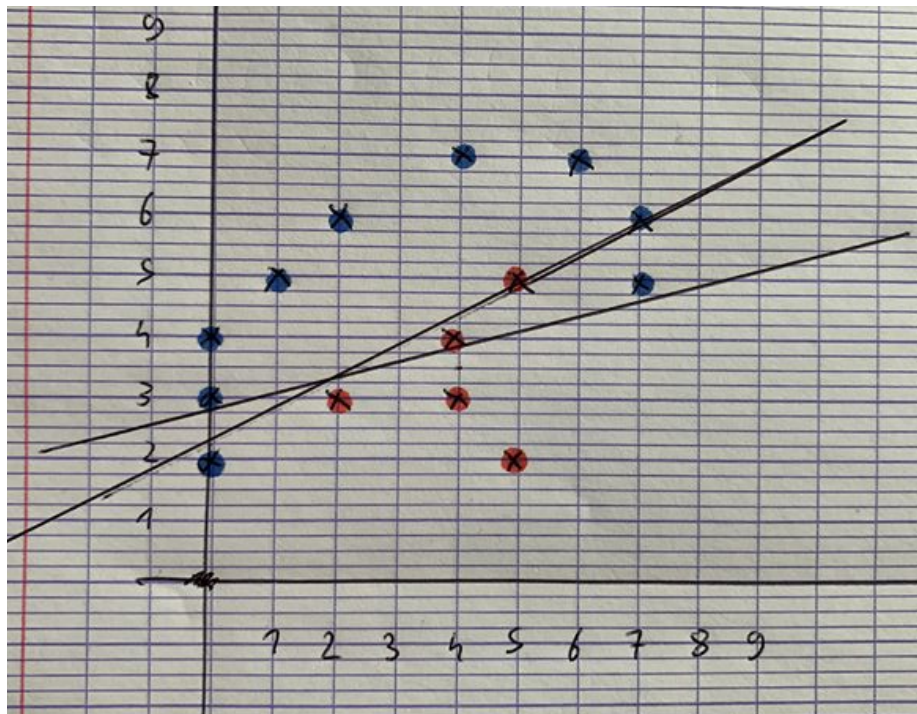
La fonction choisie pour évaluer le coût de l'intégration est la distance euclidienne.



c.

Il n'est pas possible d'avoir une séparation parfaite avec un SVM. Cependant avec un coefficient de pénalisation très faible le SVM pourra séparer approximativement les classes.

d.



La première de nos deux droites a un coefficient moyen (ou au moins plus élevé que la seconde, il n'y a qu'une erreur) alors que la deuxième possède un coefficient très faible qui a permis au SVM de trouver une solution plus approximative avec 3 erreurs.

e.

Pour séparer les deux classes nous pourrions utiliser un kernel de forme circulaire

### Exercice 3 :

a.

Étant donné un tableau croisé d'effectifs, pour savoir si les deux variables concernées sont indépendantes nous utilisons le test d'indépendance avec chi-2. En cas de chi-2 élevé alors on pourra rejeter l'hypothèse d'indépendance. Si le chi-2 est faible on pourra conclure que les deux variables sont indépendantes : il y a pas de lien statistique entre elles (généralement quand la p-value est inférieure ou égale à 0,05).

b.

Pour trouver le maximum de points pulvérisables on part de  $n = 1$  ( $n$  étant notre nombre de point). Il y a  $2^n$  combinaisons possibles et on doit s'assurer de pouvoir tous les séparer. Si jamais on ne peut le faire alors on augmente  $n$  de 1 et on essaie à nouveau de tous les séparer. Si on ne peut pas alors le maximum de point pulvérisable est  $n - 1$ .

c.

L'ambiguïté est lorsque d'une prise de décision il n'est pas possible de choisir avec certitude. On prend un risque partir du moment où on prend une décision dans ce cas là.