

# FTML Exam

Jeremie Roques

July 2020

## 1 Exercice 1

### 1.1 Question a

Pour l'estimateur 1, il y a 8 faux positif et 1 faux négatif, ceci fait donc un risque empirique de

$$\frac{8+x}{17}$$

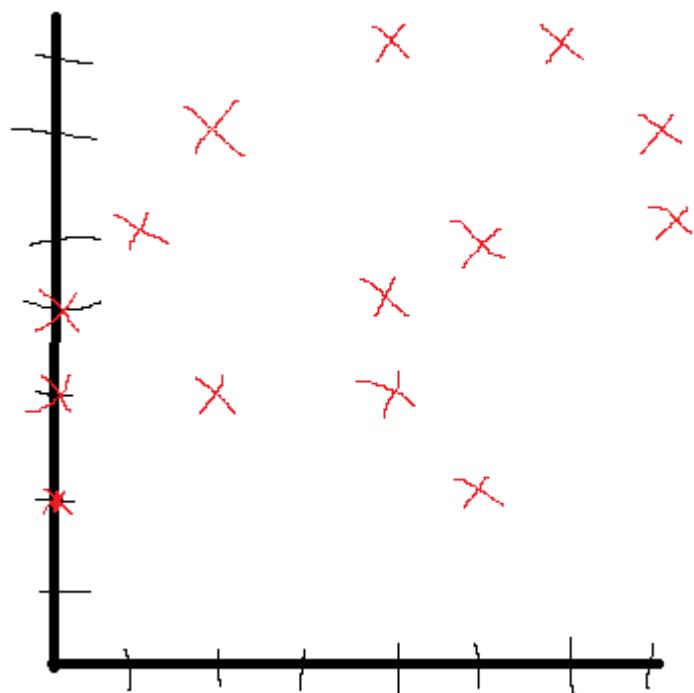
Pour l'estimateur 2, il y a 1 faux positif et 3 faux négatif, ceci fait donc un risque empirique de

$$\frac{1+3x}{17}$$

Plus la valeur de  $x$  sera élevée, plus le risque empirique de EST\_2 sera élevé par rapport à celui de EST\_1. Ils sont de la même valeur quand  $x = 3.5$

## 2 Exercice 2

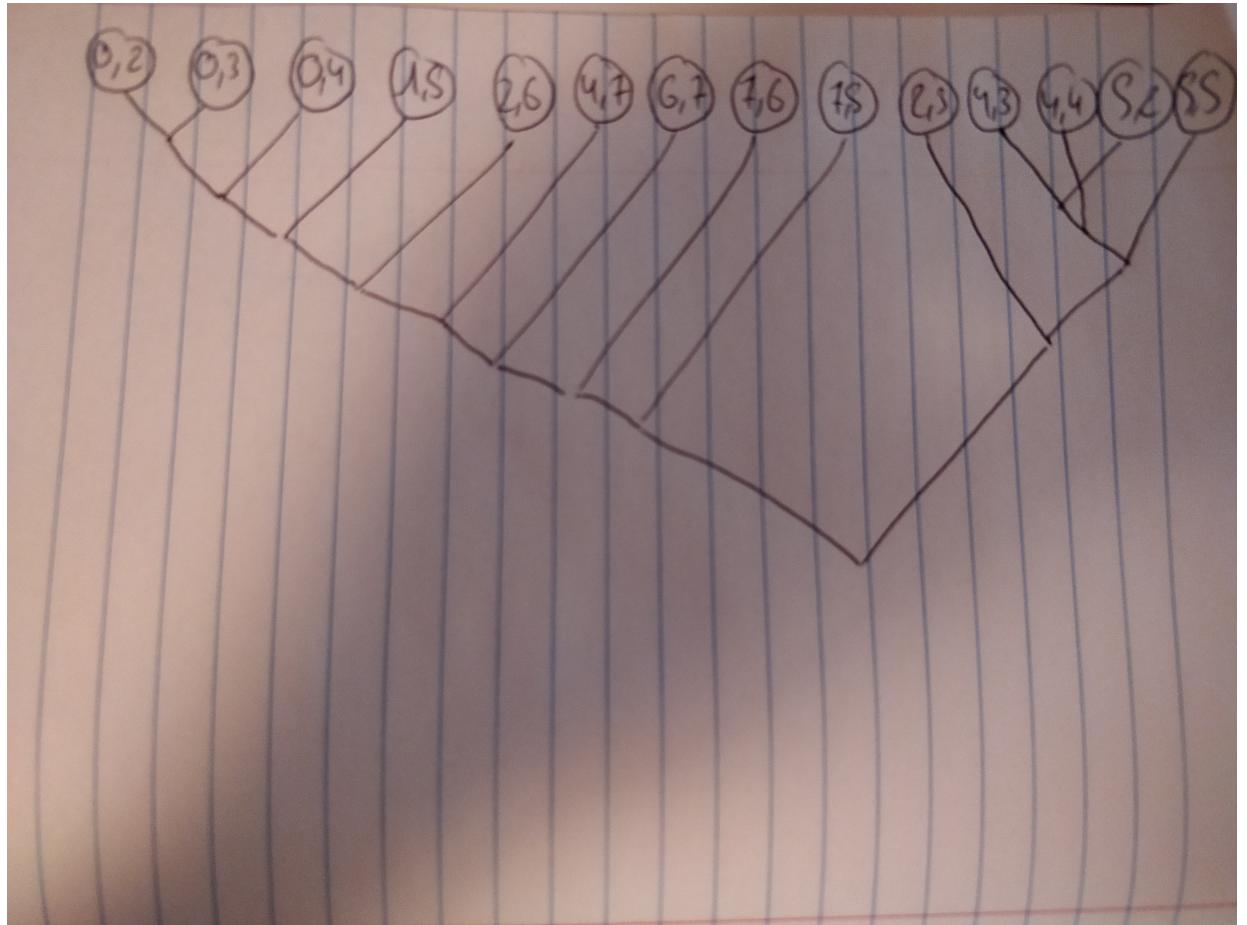
### 2.1 Question a



Les deux composantes sont, premierement l'ensemble des points qui forment l'exterieur ,les 9 premiers, et deuviemement, les 5 points au centres du graphiques Pour distinguer ces deux classes on pourrait utiliser un regroupement hiérarchique, une classification ascendante hiérarchique.

### 2.2 Question b

Voici le dendrogramme du jeu de données, basé en calculant la distance entre deux points:

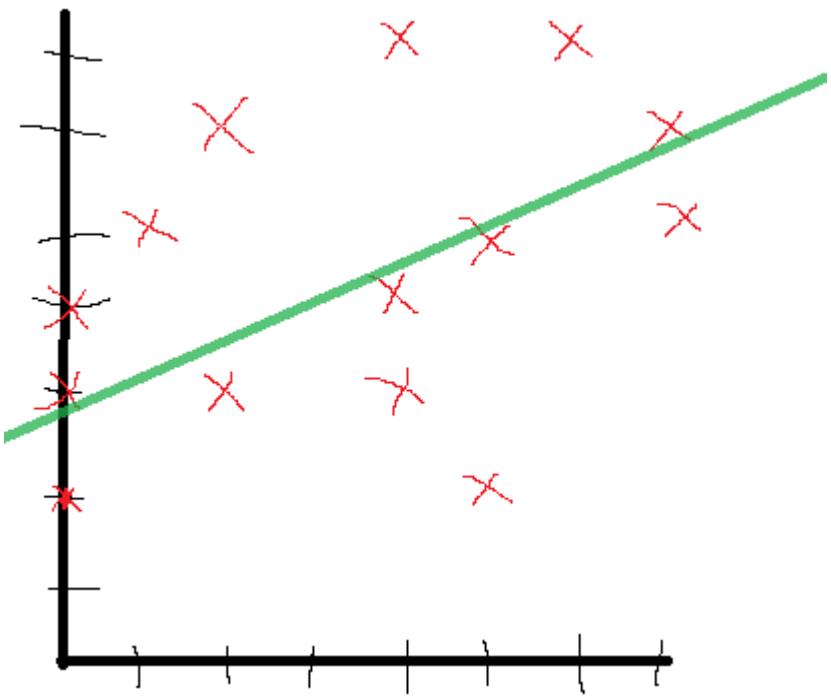


### 2.3 Question c

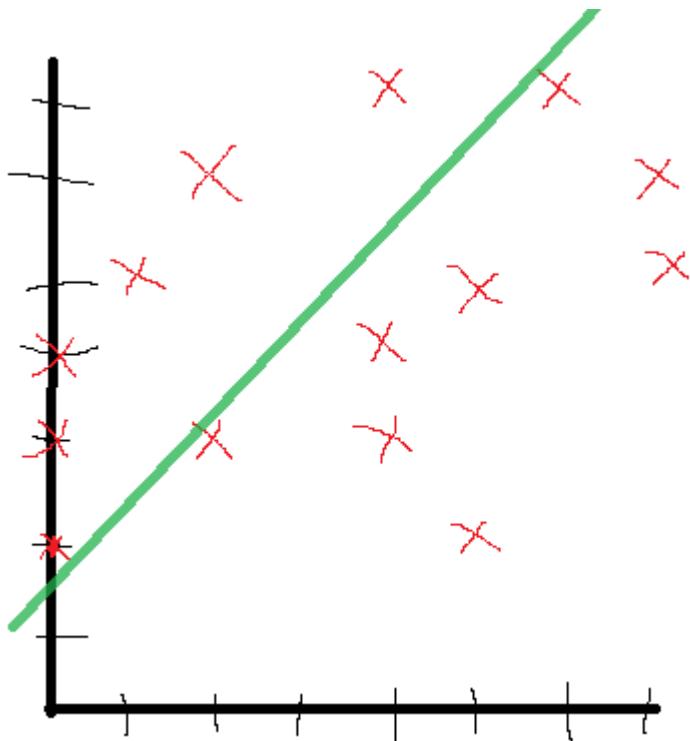
Il n'est pas possible parfaitement à l'aide d'un SVM linéaire. Cependant il est possible je pense d'avoir une vague approximation à l'aide d'un SVM linéaire avec un coefficient de pénalisation très permisif (donc faible)

### 2.4 Question d

Si l'on fixe un niveau de pénalisation élevé, voici à quoi pourrait correspondre la droite du svm :



Si l'on fixe un niveau de pénalisation plus faible que le précédent, voici à quoi pourrait correspondre la droite du svm :



### 3 Exercice 3

#### 3.1 Question a

Afin de vérifier l'indépendance de deux variables avec un tableau croisée d'effectifs, il faut effectuer le calcul du khi-2. Si le résultat de ce calcul est élevé, les deux variables ne sont pas indépendantes, si au contraire le résultats de ce calcul est faible, on peut imaginer supposer qu'il n'existe aucun lien entre les deux variables. En général le seuil d'indépendance est de 0.05. Avec le tableau des effectifs on calcul un tableau des contributions au khi-2. La valeur du khi-2 est donc la somme de toutes ses contributions.

#### 3.2 Question b

Pour connaître le nombre maximum de point de la surface d'une sphère pulvérisable par un autre ensemble, on va commencer par prendre deux point appartenant à la surface d'une sphère. Les points seront choisis et on leur attribuera une valeur (-1 et +1 par exemple) aléatoirement. Le but de la pulvérisation est d'arriver à regrouper tous les points d'une même valeur à l'aide de l'ensemble pulvérisant. Si les deux points arrivent à être séparés à l'aide de cette ensemble,

alors deux points de la surface de la sphère sont pulvérisable. On va donc continuer itérativement en rajoutant un point à chaque fois jusqu'à ne plus réussir à pulvériser n point. n-1 sera donc le nombre de point maximum pulvérisable par l'ensemble pulvérisant.

### 3.3 Question c

L'ambiguité à un rapport avec l'hésitation entre deux choses, par exemple un modèle de classification qui hésite entre deux classes.