

Partiel: FTML

Hicham Madi SCIA 2021
hicham.madi@epita.fr

July 16, 2020

Désolé d'avance pour les photos la mise en page de LaTeX m'a placé des photos qui coupent du texte parfois je m'en suis rendu compte à la toute fin. Les légendes sont là pour que vous puissiez vous y retrouver.

1 Exercice 1

1.1 a)

Le risque empirique de l'estimateur 1 est : $8 + x$ Le risque empirique de l'estimateur 2 est : $1 + 3x$

On peut voir grace à une inéquation simple $(8+x \leq 1 + 3x) \Leftrightarrow (7/2 \leq x)$

Par conséquent a partir du moment ou x est plus petit que $7/2$ l'estimateur 1 est plus pénalisant et inversement quand x plus grand que $7/2$.

1.2 b)

En étudiant l'arbre binaire ci dessous on arrive à l'estimateur suivant:

x_1	x_2	x_1	x_2
x_1	x_2	x_1	x_2
x_1	x_2	x_1	x_2
x_1	x_2	x_1	x_2

Figure 2: Estimateur

1.4 d)

$Y=T$	$X=1$	$X=2$
X_1	5	2
X_2	2	5

$Y=F$	$X=1$	$X=2$
X_1	3	7
X_2	6	4

Figure 3: Tableaux

1.5 e)

On calcule les probabilités suivantes:

Après le calcul de toutes les probabilités on retombe sur le même estimateurs que celui trouvé précédemment.

On calcule la probabilité :

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

$$\frac{P(X_1=1 \mid Y=T) \times P(X_2=1 \mid Y=T)}{P(X_1=1, X_2=1)}$$

Figure 4: Calcul de probabilité

2 Exercice 2

2.1 a)

Il y'a trois méthodes non supervisés qui semblent adaptées à la situation à savoir:

- Agglomerative Clustering

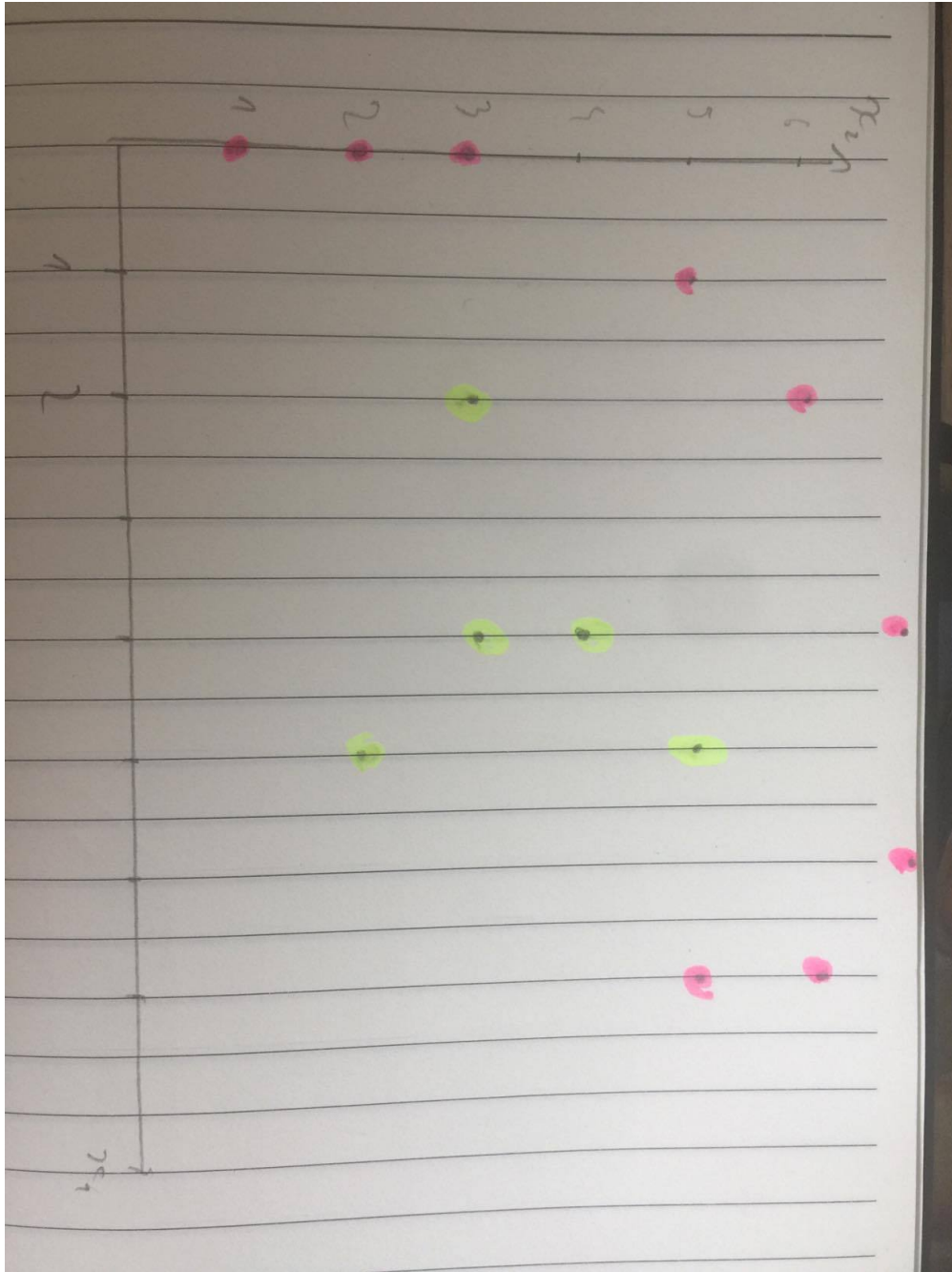


Figure 5: Representation des points

- DBSCAN
- OPTICS

2.2 c)

Non il n'est actuellement pas possible de réaliser une séparation franche des deux groupes avec un SVM linéaire.

2.3 d)

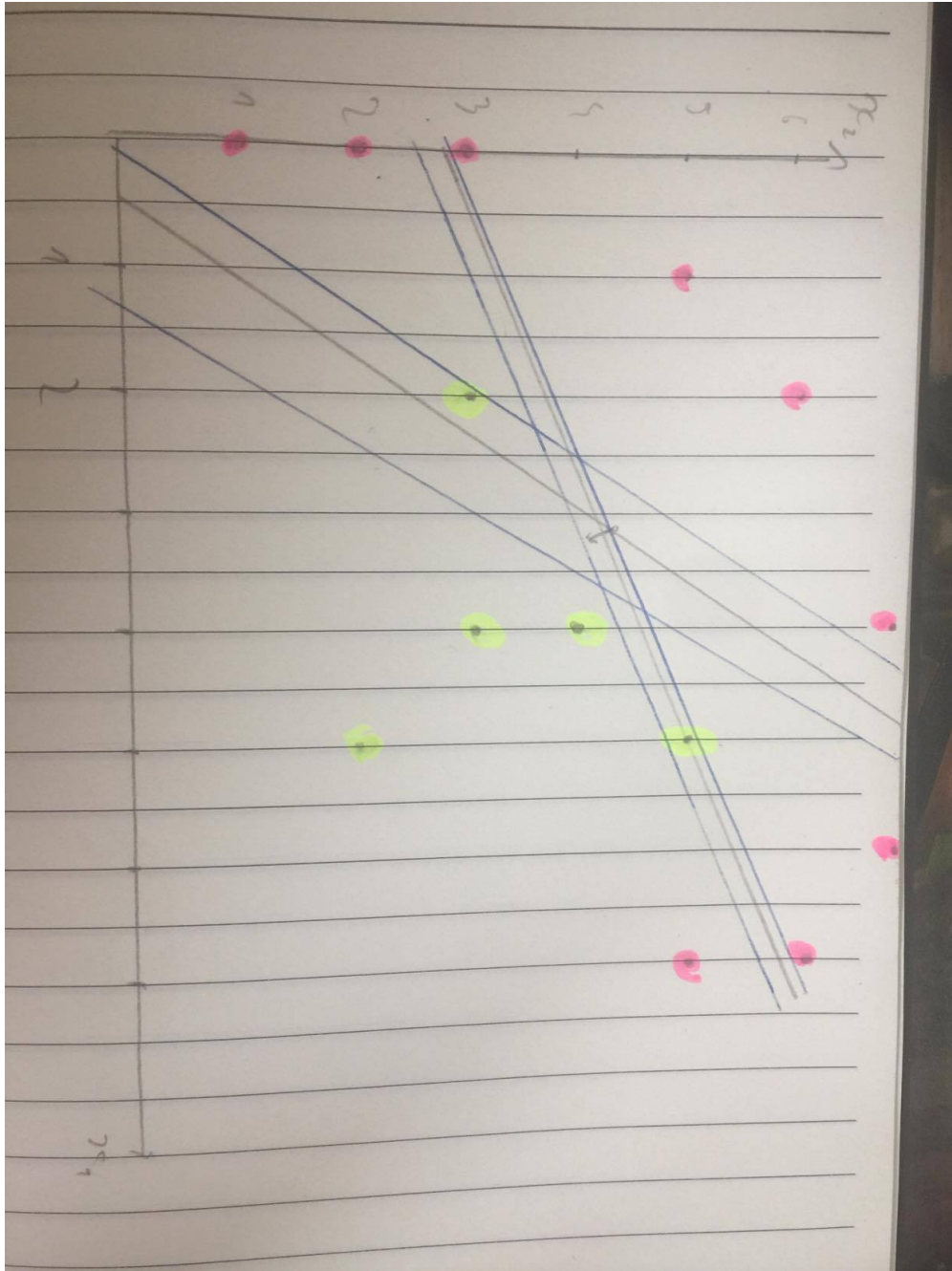


Figure 6: Representation SVM linéaire

2.4 e)

Une kernalisation polynomiale semble propice à la séparation des deux classes.

3 Exercice 3

3.1 a)

Le plus simple est de faire une étude statistique des différentes variables de manière indépendante. Par exemple prendre une part de cas aléatoire et voir si ils satisfont un événement A et prendre un nombre de cas qui satisfont l'événement B et comparer la proportion de cas satisfaisant l'événement A également Si on compare les différentes proportions on sera capable d'observer des différences en cas de dépendances des deux variables aléatoires.

3.2 c)

Le cas du risque est quand les probabilités sont bien connus alors que dans le cas de l'ambiguïté on à pas de chiffre précis.