Exam FTML 2020

Exercice 1

a. $LF(y, y') = 1$ s'il y a FP et $LF(y, y') = x$ s'il y a FN.

Pour EST_1:

- La matrice de confusion d'EST_1 est:

Pred\real	**T**	**F**
T	6	7
F	1	2

$LF(y, y') = 7 * 1 + 1 * x = 7 + x.$

- Pour EST_2:

Pred\real	**T**	**F**
T	4	0
F	3	9

$LF(y, y') = 0 * 1 + 9x = 9x.$

Pour $x > 7/8$, l'estimateur EST_2 est plus avantageux que l'estimateur EST_1.

---

b. Soit EST_3 cet estimateur, tel que $EST\_3 : f(x_1, x_2) = T$ si $x_1 = 1$ et $f(x_1, x_2) = F$ sinon.

Y	**Y_EST_3**	
T	T	
F	F	
T	T	
F	T	FP
T	T	
F	F	
F	F	
T	F	FN
F	F	
F	F	

Y	Y_EST_3	
F	T	FP
T	F	FN
F	T	FP
F	F	
F	F	
T	T	

$LF(y, y') = 3 + 2x = 3 + 4 = 7$

EST_1 et EST_2 ont un risque empirique de 9 et 18 respectivement.

Cet estimateur minimise plus le risque empirique.

---

c. Il n'est pas necessaire de se placer ici dans un cadre de baysien naif car on a assez de variable et car on ne sait pas si elles sont statistiquement indépendantes.

---

d.

Y=T	x1	x2
1	4	2
2	2	5

Y=F	x1	x2
1	3	6
2	7	4

---

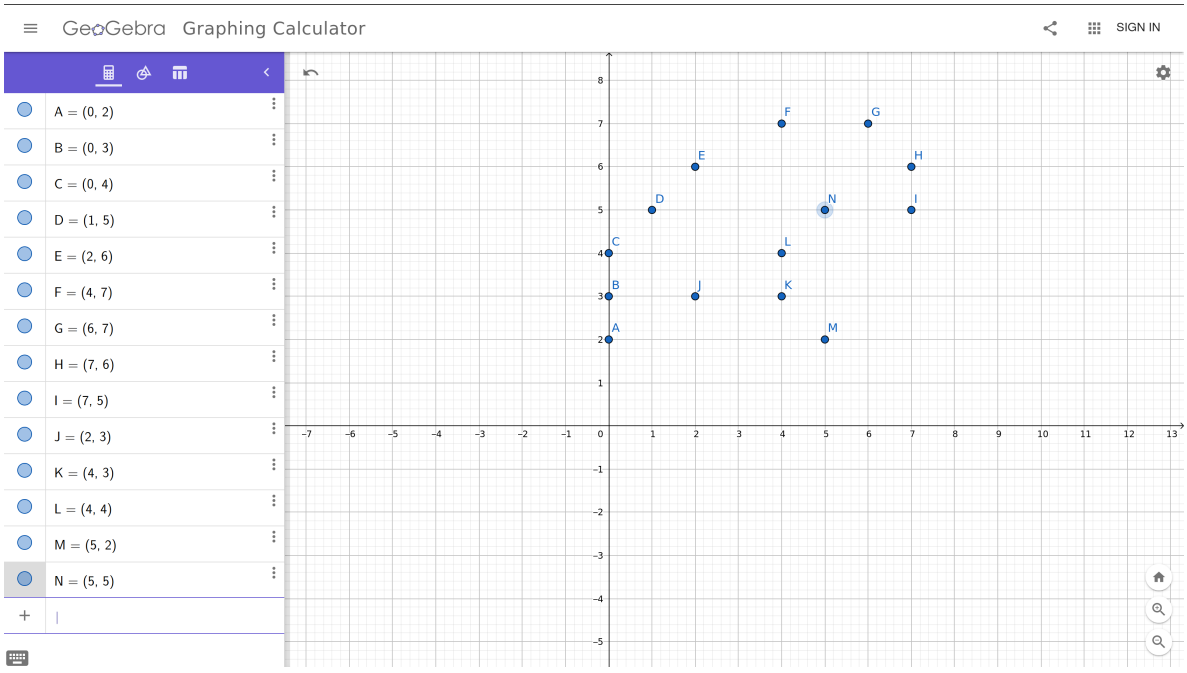
e. L'estimateur bayesien naïf serait:

$\tilde{f}(x) = argmax\ P(y = k) \prod_{j \in \{1,2\}} P(X_j | Y = k)$

Comparer?

Exercice 2

a.



Les deux composantes probables sont être (A, B, C, D, E, F, G, H, I) et (J, K, L, M, N). Il me paraît que la méthode non supervisé, parmi d'autres, qui pourrait les distinguer est la classification hiérarchique ascendante (CHA).

Les autres étant des méthode comme DBSCAN, Agglomerative clustering, spectral clustering, etc. (cf cours).

---

b. Pour évaluer le coût de l'intégration, on choisit : $f = \sum_K \sum_{i,j \in K} x(i,j)$.

Dendogramme pas fait par faute de temps.

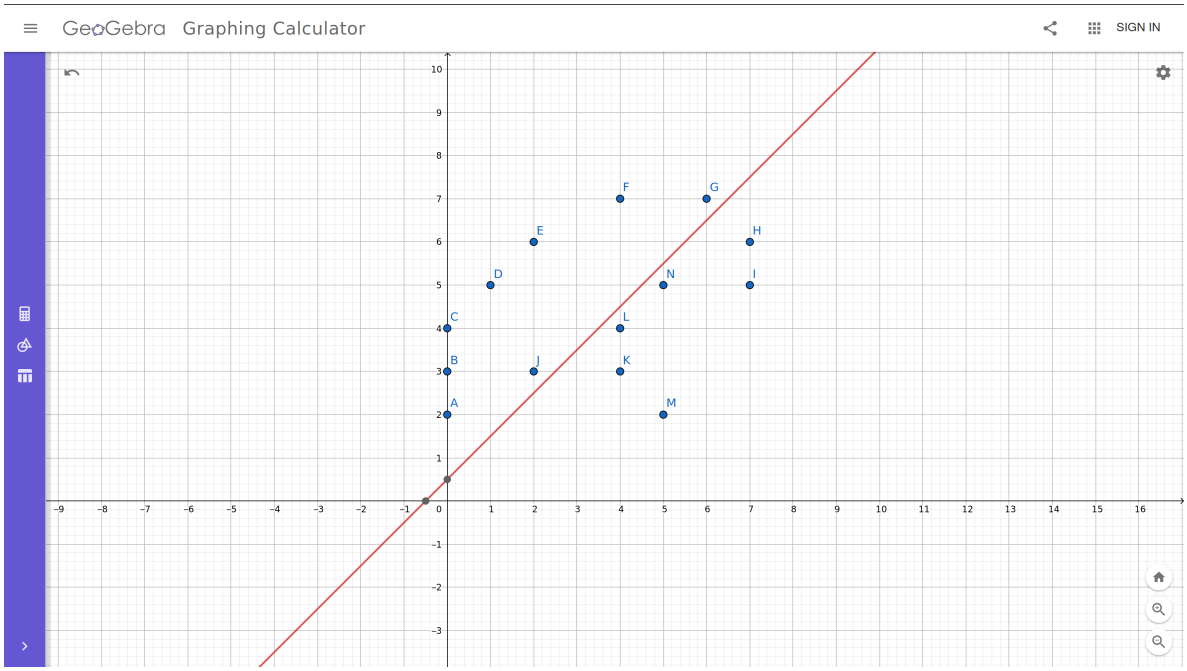
---

c. Non, il n'est pas possible de les séparer par un SVM linéaire. En effet, on est en 2D, donc le séparateur sera une droite. Ceci fait que les ensembles 1 et -1 doivent Être séparables par une droite, ce qui n'est pas possible dans ce cas.

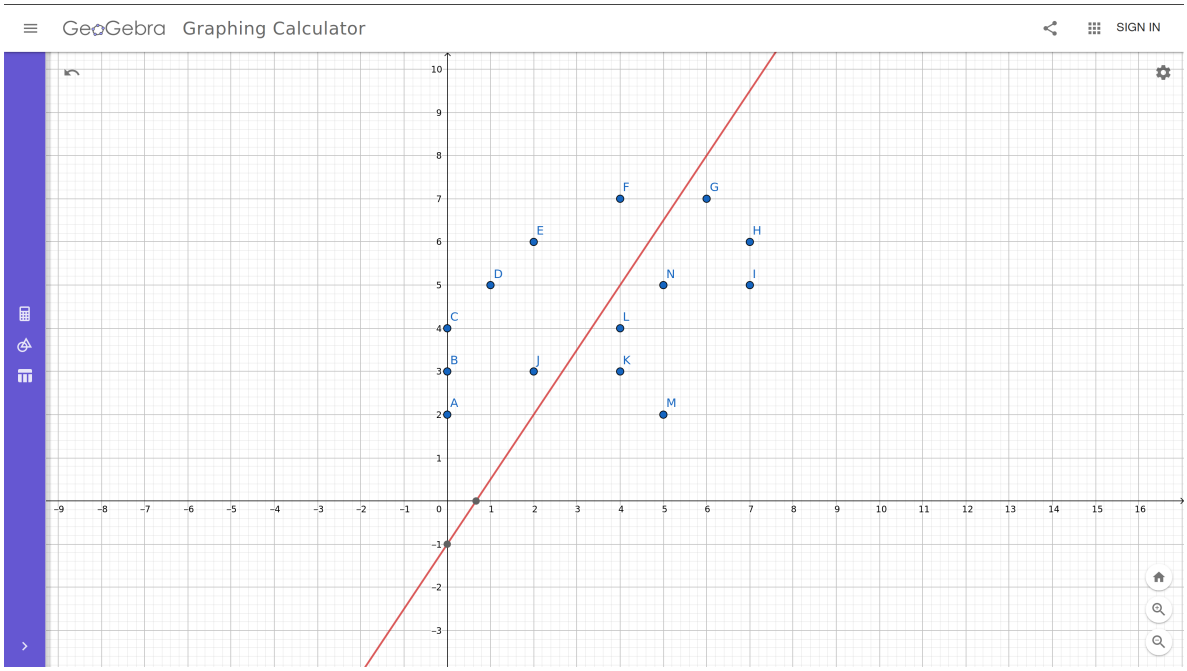
---

d.

- Forte pénalisation (C=10 par exemple):



- Plus faible pénalisation (C=1, par exemple):



---

e. On veut une fonction qui dérive un y d'un x qui va faire basculer la classe -1 "vers devant" et la classe 1, "vers derrière".

On choisit le cercle passant par M, K, L, et N: $(x_1 - a)^2 + (x_2 - b)^2 = r$ où (a, b) est le centre (~(3, 5, 6)) et r le rayon (~2).

Exercice 3

a. J'effectue une étude du type contigence de χ^2. On pose comme hypothèse nulle la dépendance des variables. Après avor effectué notre étude de χ^2 en utilisant par exemple chi2_contigency de sklearn, on regarde la p-valeur obtenue. Si elle est assez petite (0.05), on pourra rejeter l'hypothèse de la dépendance. Les deux variables seront indépendantes.

---

b. On commence par déterminer par quoi on veut pulvériser notre famille de points. Puis on détermine la borne inférieur du nombre n de points avec laquelle on peut pulvériser la famille quelque soit la disposition des points. Puis on montre que cette borne inférieur est aussi la borne supérieur et tout nombre de points plus grand n'est pas pulvérisable. Ainsi on a n.

---

c. Le risque implique la connaissance de probabilité comme par exemple quand on lance un dé. On a un risque de 5/6 de ne pas tomber sur la face 1.

L'ambiguité, par contre, implique de faire une décision quand on ne connait pas les probabilités. Il y a une ambiguité quand on est dans une situation inconnue.