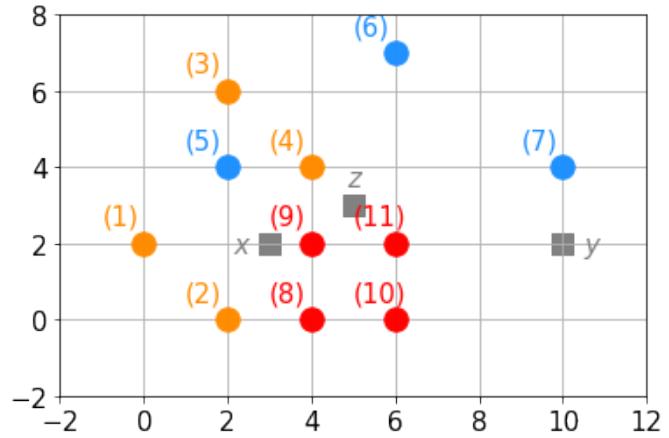


Soient $n = 11$ points en mode supervisé (X, Y) , dont le nuage en dimension $p = 2$ est représenté ci-contre :

n^o	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	\bar{x}	s^2
$t X$	0	2	2	4	2	6	10	4	4	6	6	4.18	6.88
	2	0	6	4	4	7	4	0	2	0	2	2.82	5.23
Y	1	1	1	1	2	2	2	3	3	3	3	c ou $m = 3$	

Sont aussi représentés les points x , y et z dont on souhaite savoir le groupe (classe) auquel les associer (prédirer) à l'aide d'une règle (de classement).

x	3	2
y	10	2
z	5	3



1. Nearest-Prototype Rule

- 1-1) De quoi a-t-on besoin pour appliquer cette règle ?
- 1-2) Réalisez la prédiction pour les points x , y et z à l'aide de la règle du *Plus Proche Prototype* au sens de la distance euclidienne.
- 1-3) Évaluez visuellement si on obtient le même résultat avec les distances de Manhattan et de Chebychev.

2. K-Nearest-Neighbors Rule

Ci-contre, les distances des points x, y et z aux données de X .

- 2-1) Donnez les prédictions par la règle du *Plus Proche Voisin* au sens d'une distance, en précisant les vecteurs d'étiquettes u (ce qu'on appelle parfois abusivement des "probabilités", ex. Orange)

(HW) Pour la/les distances qui reste(nt).

- 2-2) Recommencez avec la règle des 5-*PPV*.
- 2-3) On lit dans certains ouvrages spécialisés ou dans des cours qu'il vaut mieux prendre K impair. Qu'en pensez-vous ?

d_1 :	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
x	3	3	5	3	3	8	9	3	1	5	3
y	10	10	12	8	10	9	2	8	6	6	4
z	6	6	6	2	4	5	6	4	2	4	2
d_2^2 :	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
x	9	5	17	5	5	34	53	5	1	13	9
y	100	68	80	40	68	41	4	40	36	20	16
z	26	18	18	2	10	17	26	10	2	10	2
d_∞ :	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
x	3	2	4	2	2	5	7	2	1	3	3
y	10	8	8	6	8	5	2	6	6	4	4
z	5	3	3	1	3	4	5	3	1	3	1

3. Évaluation de Règle d'Apprentissage

- 3-1) Le *reclassement (resubstitution)* consiste à évaluer la règle sur l'ensemble d'apprentissage, autrement dit prendre $(X_{test}, Y_{test}) = (X_{learn}, Y_{learn})$. Évaluez la règle *NP* selon cette stratégie en donnant la matrice de confusion $\mathcal{C} = [c_{ij}]_{i,j=1,m}$ croisant Y_{test} et Y_{pred} , puis en calculant le score $CA = \text{trace}(\mathcal{C})/n$ (*Classification Accuracy*).
- 3-2) Discutez cette stratégie pour la règle du 1-*NN*. Qu'en est-t-il pour les *K-NN* ?
- 3-3) Composez, selon les règles de l'art, un ensemble de test (X_{test}, Y_{test}) et un ensemble d'apprentissage (X_{learn}, Y_{learn}) en *holdout* à 30% (à peu près).

(HW) Recommencez en *holdout* à 50% (à peu près). Serez-vous en mesure d'avoir le résultat de la 2-validation croisée ?

- 3-4) Quelle autre stratégie est la plus recommandée pour ces données ?
- 3-5) Réalisez visuellement les prédictions par la règle des 3-*NN* avec la distance de Manhattan selon cette stratégie.
- 3-6) Vous donnerez la matrice de confusion, en déduirez la *Classification Accuracy*, puis les scores de *Recall*, *Precision*, et *Classification Accuracy* par classe : $R_j = \frac{TP}{TP+FN}$, $P_j = \frac{TP}{TP+FP}$ et $A_j = \frac{TP+TN}{TP+FN+FP+TN}$. Vous déduirez aussi les scores globaux (donnés par Orange).

(HW) Calculez les autres scores : False Alarm $FA_j = \frac{FP}{TP+FP} = 1 - P_j$ et $F1_j = 2 \frac{R_j \times P_j}{R_j + P_j}$

- 3-7) Pourquoi est-ce plus difficile de réaliser visuellement les prédictions par la règle du *NP* selon cette stratégie ?

(HW) Faites-le, avec la distance de Manhattan.

(HW) Dans un espace réduit

Soit le tableau de données du ressenti de 10 apprenants sur l'utilisation de logiciels (Blue et Yellow) de *Machine Learning* et leur association Green du TD précédent. On donne ci-contre les vecteurs directeurs des 2 premiers (seuls ?) axes discriminants, résultats de l'AFD.

-0.0656	0.5379	0.8404	${}^t u_1$
0.8693	-0.4441	0.2172	${}^t u_2$

- ${}^t \bar{x}_{Blue} = (11.23, -1.58)$
- ${}^t \bar{x}_{Green} = (21.87, 6.83)$
- ${}^t \bar{x}_{Yellow} = (7.73, 11.20)$

(HW) Projetez les 10 apprenants dans l'espace discriminant.

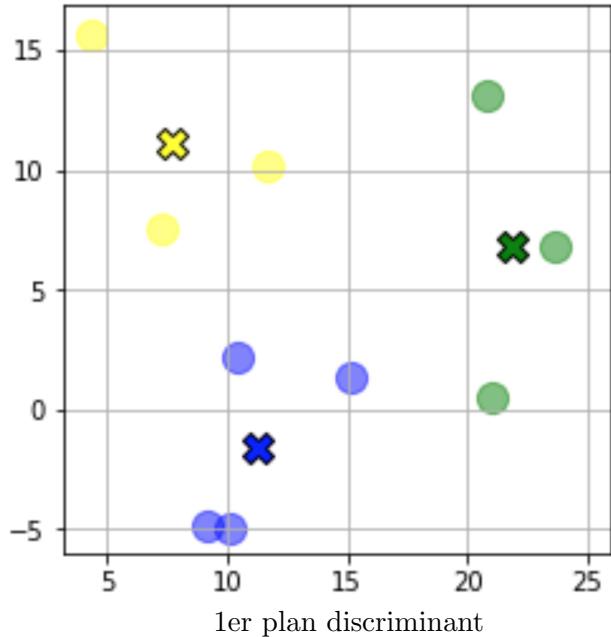
(HW) Vérifiez que les centres des groupes sont bien ceux donnés ci-contre.

(HW) À l'aide de la règle *NP* ou *1-NN*, prédisez le mode d'utilisation d'un nouvel apprenant qui aurait donné comme évaluation ${}^t x = (10, 10, 10)$.

(HW) Évaluez (visuellement) les règles *NP* et/ou *1-NN* par *Leave-One-Out* dans l'espace discriminant.

(HW) Aurait-on d'autant bons, voire meilleurs, résultats si on se plaçait dans le premier plan principal (ACP) ?

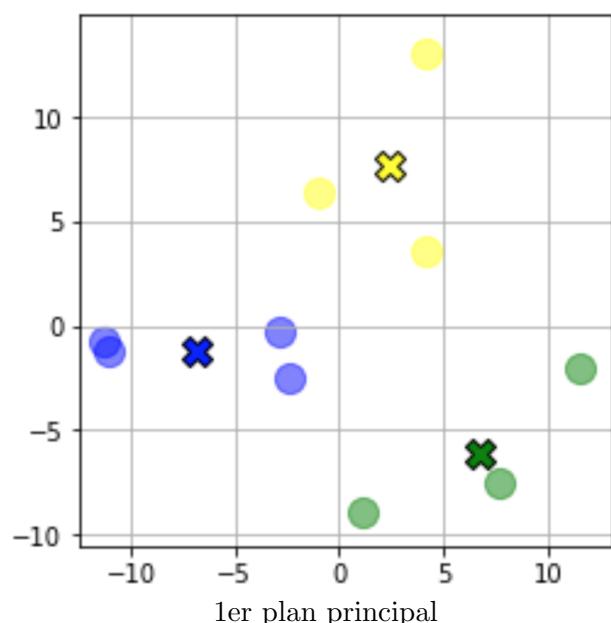
Espace réduit par AFD



		D1	D2	_sans amel.	amel.+	logiciel
0	10.425	2.134	2	4	10	Blue
1	10.156	-4.933	2	16	2	Blue
2	7.246	7.539	11	7	5	Yellow
3	20.836	13.149	16	11	19	Green
4	4.267	15.608	19	4	4	Yellow
5	21.084	0.527	3	13	17	Green
6	11.685	10.155	12	6	11	Yellow
7	23.684	6.819	10	14	20	Green
8	9.146	-4.914	1	14	2	Blue
9	15.173	1.378	7	15	9	Blue

- ${}^t \bar{x}_{Blue} = (11.23, -1.58)$
- ${}^t \bar{x}_{Green} = (21.87, 6.83)$
- ${}^t \bar{x}_{Yellow} = (7.73, 11.20)$

Espace réduit par ACP



		C1	C2	_sans amel.	amel.+	logiciel
0	-2.815	-0.277	2	4	10	Blue
1	-11.004	-1.185	2	16	2	Blue
2	-1.046	6.452	11	7	5	Yellow
3	11.541	-2.004	16	11	19	Green
4	4.153	13.081	19	4	4	Yellow
5	1.066	-8.926	3	13	17	Green
6	4.139	3.586	12	6	11	Yellow
7	7.670	-7.525	10	14	20	Green
8	-11.263	-0.714	1	14	2	Blue
9	-2.442	-2.489	7	15	9	Blue

- ${}^t \bar{x}_{Blue} = (-6.88, -1.17)$
- ${}^t \bar{x}_{Green} = (6.76, -6.15)$
- ${}^t \bar{x}_{Yellow} = (2.42, 7.71)$