Nom:			
Prénom:			
Parcours	ISI gr1	ISI gr2	SAR

HMM – 4pts

On considère un HMM à 3 états : s_1=soleil, s_2 = pluie, s_3=nuage. Les observations correspondant à ces états sont directement l'état du ciel. On a donc, pour le premier état $\begin{pmatrix}1\\0\\0\end{pmatrix}$, pour le second, $\begin{pmatrix}0\\1\\0\end{pmatrix}$ et pour le troisième $\begin{pmatrix}0\\0\\1\end{pmatrix}$. A l'instant initial, chaque état est équiprobable. Pour les instants suivants, on a 2 fois plus de chance de rester dans le même état que de changer d'état et si on change d'état, les deux autres états sont équiprobables.

1. **2pts** Déterminer la matrice de transition A, le vecteur de probabilité initiale Π et la matrice d'observation B.

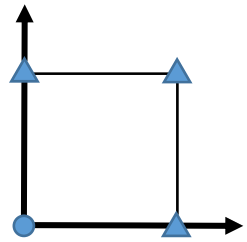
$$A = \begin{bmatrix}2/3 & 1/6 & 1/6\\ 1/6 & 2/3 & 1/6\\ 1/6 & 1/6 & 2/3\end{bmatrix} \quad \pi = \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix} \quad B= \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

2. **2pts** Déterminer la probabilité de la séquence y= « soleil, soleil, nuage, soleil, pluie » en utilisant l'algorithme forward (on posera les calculs sans les résoudre).

	soleil	soleil	nuage	soleil	pluie
S1	1/3	2/9		2/9*1/6*1/6	
S2					2/9*1/6*1/6*1/6
S3			2/9*1/6		

SVM – 6pts

On considère l'ensemble de points, de dimension 2 ci-dessous :

$\begin{pmatrix} 0 & 0 \\ 0 & 2 \\ 2 & 0 \\ 2 & 2 \end{pmatrix}$ de classe $\begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$	

On souhaite résoudre le problème de classification en utilisant le formalisme des SVM.

1. **1 pt** Tracer (sans faire de calcul) sur la figure l'hyper-plan optimal et entourer les vecteurs support.

> Axe comme \
> Les vecteurs support sont les 3 premiers points

2. **2pts** A partir de 2 points de cet hyperplan et sachant qu'il vérifie $\min_i y_i(\boldsymbol{w}^T . \boldsymbol{x}_i + b) = \mathbf{1}$ retrouver les paramètres w_a, w_b et b de l'hyperplan. Tout le raisonnement doit être **clairement** explicité.

> En (1,0), on a $w_a + b = 0$
> En (0,1), on a $w_b + b = 0$
> Le point (2,0) est vecteur support et vérifie donc $y_i(\boldsymbol{w}^T . \boldsymbol{x}_i + b) = \mathbf{1}$. D'où $2w_a + b = 1$
> Les équations 1 et 3 donnent $b = -1$. Il en découle ensuite $w_a = 1$et $w_b = 1$

3. **1pt** Sachant que l'hyperplan optimal a pour équation $w = (1, 1) et\ b = -1,$ donner l'équation de la fonction de décision.

> Elle a pour équation $f(\boldsymbol{x}) = \boldsymbol{w}^T . \boldsymbol{x} + b$=0, soit $x_a + x_b - 1 = 0$

4. **2pts** On ajoute un cinquième point en (-0.5 -0.5) d'étiquette 1 et on résout le problème de classification avec un SVM à marge souple. Le solver renvoie le même hyperplan que précédemment et les 5 variables d'ajustement ξ_i. Donner la valeur de ces 5 variables.

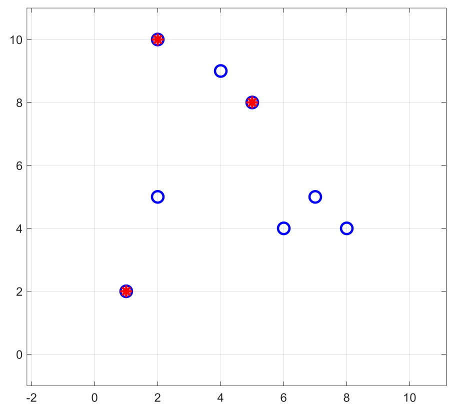
> Les variables d'ajustement sont nulles pour les points qui vérifient la condition de séparabilité. On a donc $\xi_1 = \xi_2 = \xi_3 = \xi_4 = 0$.

Pour le cinquième point, on a
$y_i\left(\boldsymbol{w}^T.\boldsymbol{x}_i + b\right) = 1 - \xi_i$ soit $-0.5w_a - 0.5w_b + b = 1 - \xi_5$ et en remplaçant w_a, w_b et b par leur valeur, -2 = $1 - \xi_5$ et donc, $\xi_5 = 3$

k-means – 3.5pts

On considère les 8 points ci-contre que l'on souhaite regrouper en 3 ensembles en utilisant l'algorithme des k-means et la distance euclidienne.
Les points 1, 4 et 7 (représentés en rouge ci-contre) ont été tirés aléatoirement pour l'initialisation des centres.

2	10
2	5
8	4
5	8
7	5
6	4
1	2
4	9

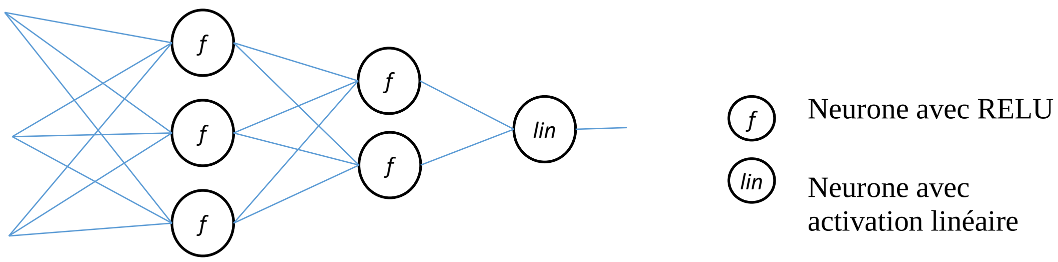


Quelle seront les valeurs des 3 nouveaux centres suite à la première itération ?

1er centre : (2,10)
2ieme centre (6,6)
3ieme centre (1.5 3.5)

Réseaux de neurones – 6.5pts

1. **2pts** On considère le réseau de neurone suivant :



Combien de paramètre y a-t-il à estimer ?

A quoi peut servir ce réseau (quel problème peut-il résoudre) ?

Si on suppose que tous les poids et bais valent 1, quelle sera la sortie du réseau à l'entrée (1, -6, 3) ?

$12 + 8 + 2 + 1 = 23$ paramètres
Il peut servir à faire de la régression (1 seul neurone de sortie, activation linéaire)
A la sortir de la première couche, on aura que des zéros. A la sortir de la seconde couche, on aura que des 1 et en sortie, on aura 3

2. **1.5pts** L'image ci-dessous est passée en entrée d'une couche convolutionnelle. Représentez la réponse au filtre F, avec un stride de 1 et un padding de 0.

3	3	3	1	6
3	3	3	1	6
3	3	3	1	1
1	1	1	4	4
5	5	1	4	4

$$F = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

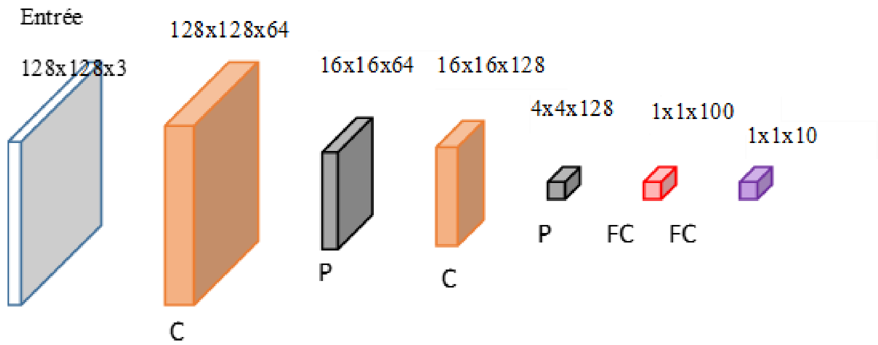
18	12	22
14	13	18
14	18	14

3. **1.5pts** En considérant l'image d'entrée ci-dessous et une couche de max-pooling de taille 3x3 avec un stride de 2 et un padding de 0, donner l'image en sortie.

1	2	4	1	4	0	1
0	0	1	6	1	5	5
1	4	4	5	1	4	1
4	1	5	1	6	5	0
1	0	6	5	1	1	8
2	3	1	8	5	8	1
0	9	1	2	3	1	4

4	6	5
6	6	8
9	8	8

4. **1.5pts** On considère l'architecture suivante composée de couches convolutionnelles C, de pooling P et fully connected FC.



Les couches convolutionnelles sont réalisées avec des filtres 3x3, avec un padding de 1 un un stride de 1. De combien de filtres est composée la seconde couche convolutionnelle ? Combien y a-t-il de paramètres à estimer sur cette couche ?

128 filtres
128*(3*3*64+1)