

Nom:	
Prénom:	

Exercice 1 (5,5 pts)– KPPV

Considérons un problème de classification à deux classes dont la base de référence (2 classes) est représentée par des ronds et des carrés Figure 1.

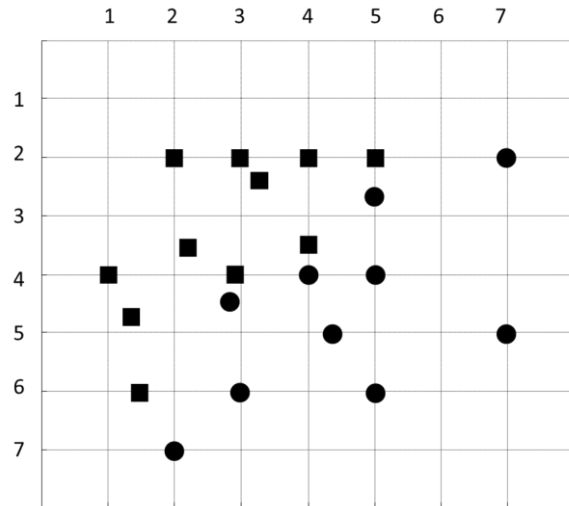


Figure 1

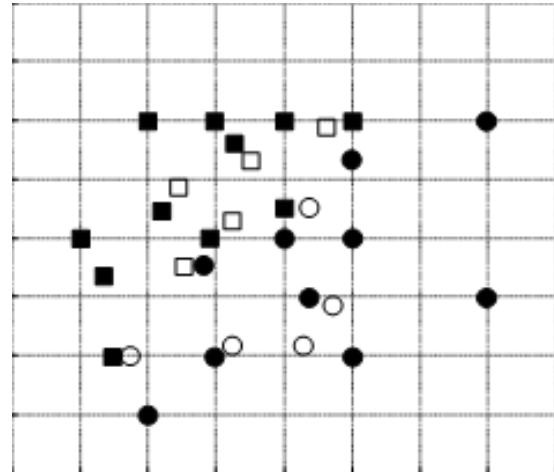


Figure 2

1. (1pt) Tracer Figure 1 les frontières entre classes en utilisant l'algorithme du plus proche voisin.
2. (1pt) Donner la matrice de confusion obtenue en classant les exemples (dessinés en blanc) de la Figure 2 avec l'approche 1ppv.

--

3. (1pt) Sachant que, sur la base de référence, la moyenne des carrés est $(x,y)=(2.75 \ 3.3)$ et celle des ronds $(4.5 \ 4.6)$, tracer Figure 2 la frontière entre classes en utilisant l'approche nearest mean.
4. D'une manière générale,
 - (0.5pt) Quelle est l'influence de k dans la classification par les k -ppv ? Que se passe-t-il si on utilise k petit ? k grand ?

--

- **(0.5pt)** Est-il vrai que plus le nombre d'exemples de référence est grand, plus le temps de calcul est grand pour classer un nouvel exemple ? Justifier votre réponse.

- **(0.5pt)** Est-il vrai que plus le nombre d'exemples de référence est grand, meilleurs seront les résultats (tendance) ? Justifier votre réponse.

- **(0.5pt)** L'algorithme du 1ppv met 0.0005s pour classer un exemple quand il y a 500 exemples de référence par classe (2 classes). Combien de temps mettra l'algorithme nearest-mean pour classer un exemple ?

- **(0.5pt)** Souvent, avant de réaliser l'algorithme des kppv, on procède à une standardisation des données. De quoi s'agit-il ? Concrètement, réalise-t-on exactement les mêmes opérations sur les bases de références et de test ? Pourquoi ? Justifier votre réponse.

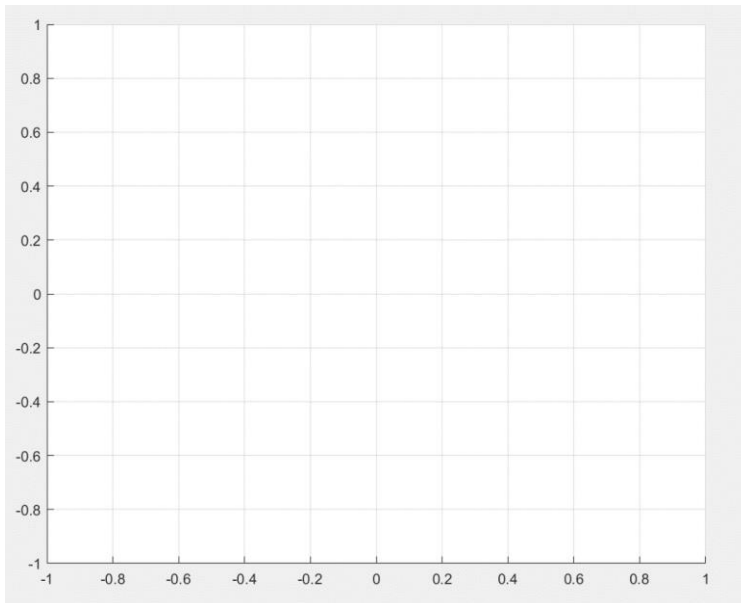
Exercice 2 (4.5 pts) – Courbe ROC

On souhaite mettre en place un processus de reconnaissance de champignons (2 classes). Pour cela, trois paramètres sont mesurés : le diamètre du chapeau, la hauteur du pied et la densité. Or, le classifieur dont nous disposons n'accepte que deux descripteurs en entrée. De plus, un paramètre de contrôle est à régler dans le classifieur, celui-ci modifie les résultats de manière assez opaque. Ci-dessous, les tests réalisés.

Descripteur 1+2						Descripteur 1+3						Descripteur 2+3					
	Test1	Test2	Test3	Test4	Test5		Test1	Test2	Test3	Test4	Test5		Test1	Test2	Test3	Test4	Test5
TP	100	99	86	56	17	TP	100	99	86	56	17	TP	96	89	57	15	3
FN	0	1	14	44	83	FN	0	1	14	44	83	FN	4	11	43	85	97
FP	96	89	57	15	3	FP	94	68	26	10	3	FP	94	68	26	10	3
TN	4	11	43	85	97	TN	6	32	74	90	97	TN	6	32	74	90	97

1. **(3.5 pts)** Tracer les courbes ROC des 3 classifieurs afin de déterminer le meilleur couple de descripteurs à utiliser indépendamment du réglage du classifieur. Expliquer votre façon de procéder. Conclusion.

2. (1 pt) Pour le meilleur couple de descripteurs, situer graphiquement sur la courbe ROC le point EER (Equal Error Rate, même taux de faux positifs et de faux négatifs). Donner le taux de reconnaissance que l'on a alors.



Exercice 3 (3 pts) – Performance d'un classifieur

Un algorithme de classification amène à la matrice de confusion $\begin{bmatrix} 977 & 50 & 30 \\ 1 & 1128 & 70 \\ 4 & 25 & 1015 \end{bmatrix}$. Pour la **première classe**, que valent la précision et le rappel ?

Exercice 3 (7 pts) – Réseau de neurones

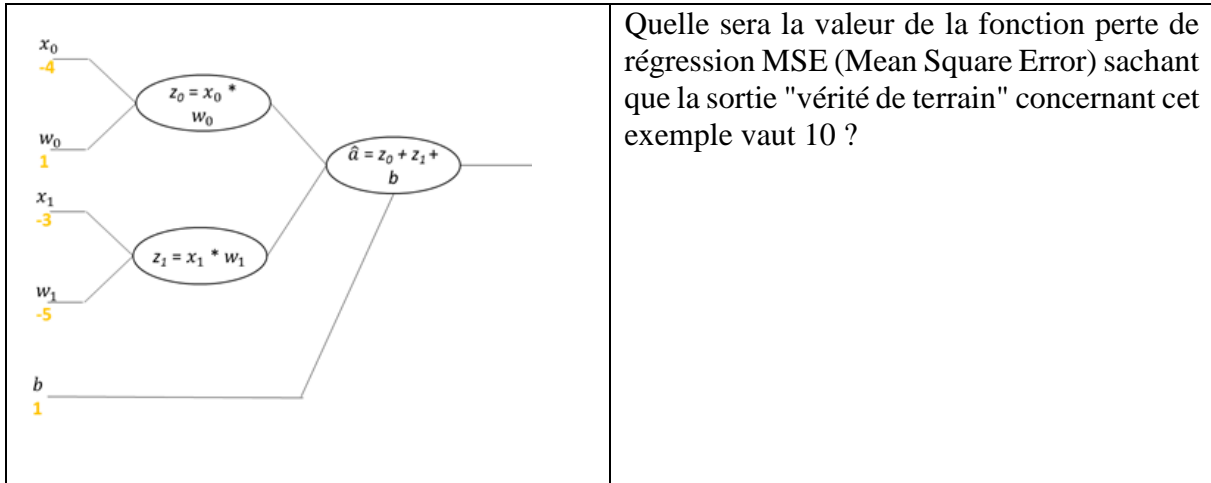
1. **(1 pt)** Un CNN est appris pour reconnaître la classe d'images de taille 256x256x3. La première couche de ce CNN est une couche convolutionnelle composée de 12 filtres 5x5. Combien y a-t-il de paramètres à estimer sur cette première couche ?

2. **(1 pt)** La sortie de la dernière couche d'un CNN réalisant une classification à deux classes vaut (1.3 0.9). On applique un soft-max à cette sortie. Quelle sera la sortie après soft-max ?

3. **(1 pt)** Considérons une image de taille 16x16x3 à laquelle on applique une couche convolutionnelle composée de 6 filtres 3x3, avec un stride de 1 et un padding de 0. Quelle sera la dimension de l'image de sortie ?

4. **(1 pt)** Un MLP est appris pour reconnaître des images monochromes de taille 12x12 pixels. Il y a 10 classes. Le CNN est composé d'une couche cachée de 48 neurones. Combien le réseau aura de paramètres à estimer lors de l'apprentissage ?

5. (0.5 pt) On considère le petit réseau suivant :



6. (1 pt) Lors de l'apprentissage d'un réseau de neurones, on observe les courbes d'erreur sur les bases de test et d'apprentissage. Donner, pour ces deux bases, des courbes résultant d'un trop fort pas d'apprentissage, d'un sur-apprentissage ou d'un sous-apprentissage.

Trop fort pas d'apprentissage	Sur-apprentissage	Sous-apprentissage

7. (0.5 pt) Rappeler en quoi consiste le dropout. Est-il utilisé en test et en apprentissage ? Justifiez votre réponse.

8. (0.5 pt) Rappeler en quoi consiste la régularisation L2 .

9. (0.5 pt) Le module inception est donné par :

