

Examen SCLASS

Durée : 2H

*Seuls documents autorisés : Cours et notes de cours
– Barème indicatif –*

*Ne pas oublier d'indiquer son nom sur chacune des feuilles
et de toute les rendre (même vides)!*

Exercice 1 – k Plus Proche Voisins

[Répondre sur la feuille “ k _PPV”]

On note k _ppv l'algorithme des k plus proches voisins pour un problème de classification à c classes.

Q 1.1 – 1_ppv

Quelle est l'erreur d'apprentissage e_a produit par un algorithme 1_ppv ?

Q 1.2 – n _ppv

Pour $|\Pi_a| = n$, à quoi correspond un algorithme n _ppv ? Quelle est l'erreur d'apprentissage e_a de cet algorithme ?

Q 1.3 – 2_ppv

Donner une méthode pour tracer la frontière de décision de 2_ppv. Tracer (au mieux) cette frontière pour le problème à 2 classes présenté sur le graphe.

Exercice 2 – Classifieur Linéaire Binaire

[Répondre sur la feuille “Classifieur Linéaire Binaire”]

Q 2.1 Le problème à 2 classes présenté sur le graphe est clairement linéairement séparable. Proposer un CLB permettant une décision sans erreur d'apprentissage e_a . Dessiner cette frontière sur le graphe. Quel est le critère de classification pour un point $P(x, y)$?

Q 2.2 Fournir une écriture matricielle (en dimension 2 et 3) de l'équation de la frontière de décision de ce CLB.

Q 2.3 Calculer la distance du point A à votre frontière (ne pas la mesurer!).

Exercice 3 – Perceptron

[Répondre sur la feuille “Perceptron”]

On déroule l'algorithme classique du perceptron sur Π_a présenté ici. À l'itération t , les poids courants sont $W_t = \begin{pmatrix} -105 \\ 20 \\ -60 \end{pmatrix}$ ($\epsilon = 1$).

Q 3.1 Tracer la frontière du CLB de l'étape courante.

Q 3.2 Fournir la prochaine modification des poids dans le déroulement de l'algorithme du perceptron. On rappelle que $\epsilon = 1$.

Q 3.3 Tracer la frontière du CLB correspondant à W_{t+1} .

Exercice 4 – Arbres de classification

[Répondre sur la feuille “Arbre de classification”]

On considère un classifieur de type arbre de décision binaire classique où chaque noeud n’effectue que la comparaison d’une seule variable à un seuil choisi. L’arbre est entraîné avec le critère suivant : on choisit des subdivisions qui minimisent la proportion d’erreurs de classification, et ceci jusqu’à obtenir 0 erreurs sur l’ensemble d’entraînement.

Q 4.1 Dessiner, sur le graphe, la séquence de subdivisions réalisée lors de la construction de l’arbre de décision.

Q 4.2 Dessiner l’arbre de décision correspondant.

Q 4.3 Calculer le gain d’information du noeud racine.

Exercice 5 – Naïf Bayes

Soit une base de données en apprentissage supervisées dont voici la base Π_a

X_1	X_2	X_3	X_4	\mathcal{C}
0	1	0	0	A
1	1	1	0	B
1	1	0	0	A
1	0	1	0	B
1	0	1	0	B
0	0	1	1	A
0	1	0	1	B
0	1	0	0	A
1	0	1	0	B
1	1	1	0	B
1	1	0	0	A
1	0	1	0	B
0	0	1	1	A
0	1	0	1	B

En notant $x = (x_1, x_2, x_3, x_4)$ le vecteur des 4 attributs, on veut construire un classifieur bayésien $f(x)$ maximisant la probabilité de la classe *a posteriori*.

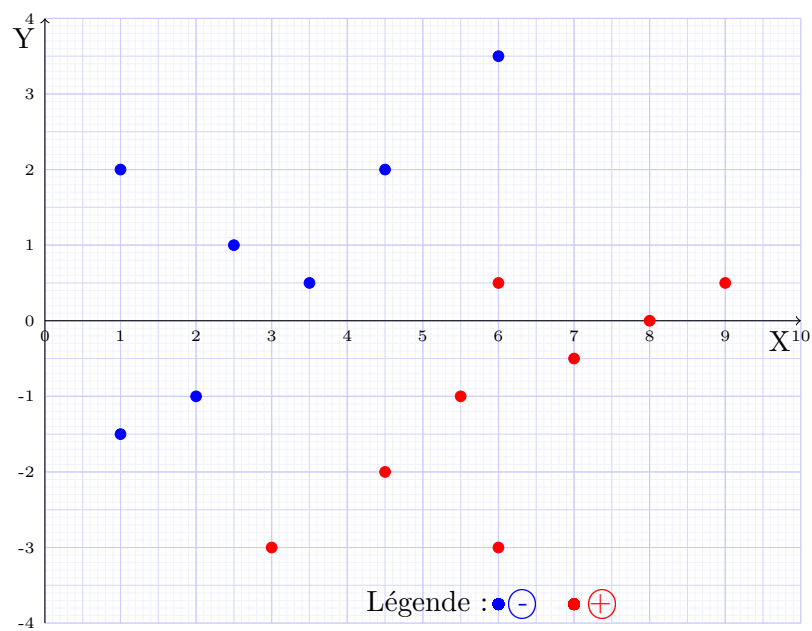
Q 5.1 Quels sont les paramètres à estimer (sur la base) pour pouvoir construire ce classifieur ? À quel moment a-t-on besoin de l’hypothèse du Naive Bayes ?

Q 5.2 Classifier le vecteur : $x = (0, 1, 1, 0)$.

k_{PPV}

Nom :

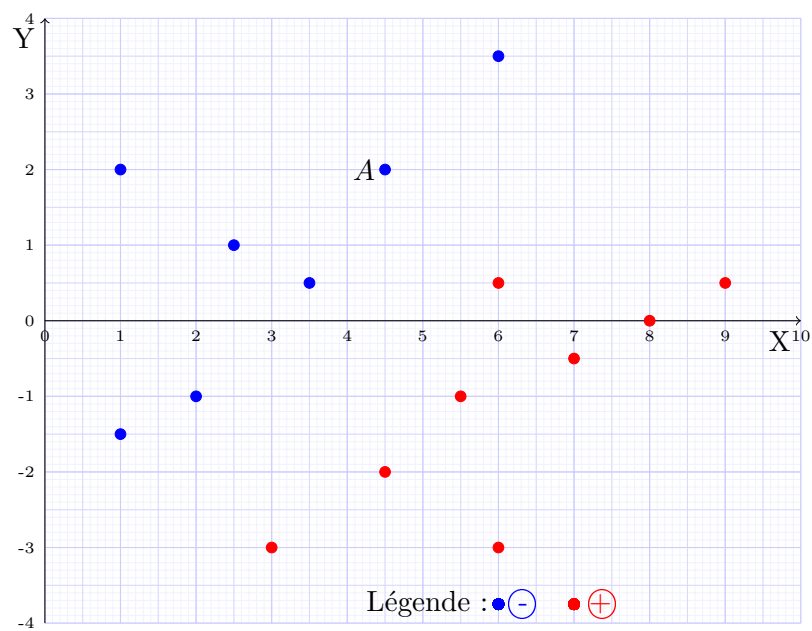
Prénom :



Classifieur Linéaire Binaire

Nom :

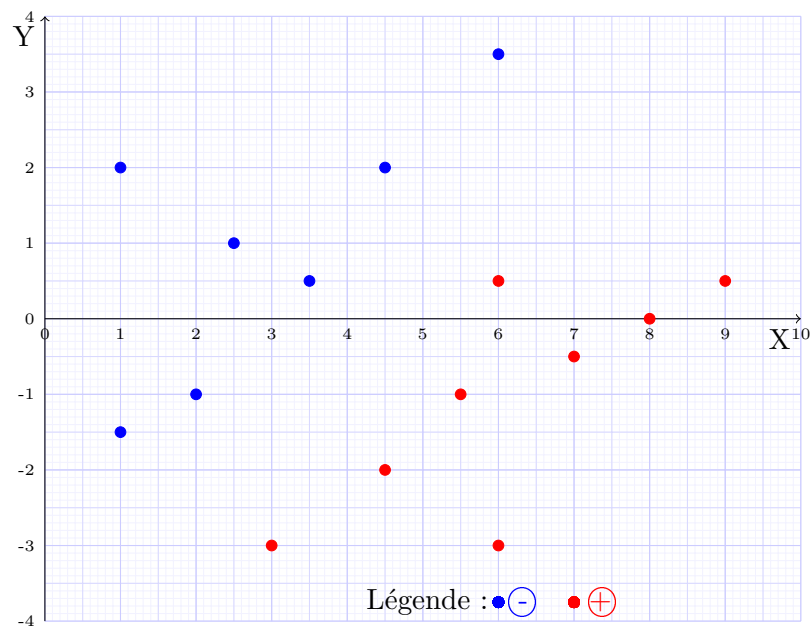
Prénom :



Perceptron

Nom :

Prénom :



Arbres de classification

Nom :

Prénom :

