

Examen semestriel

Module "Datamining" - "Modèles pour le Datamining"

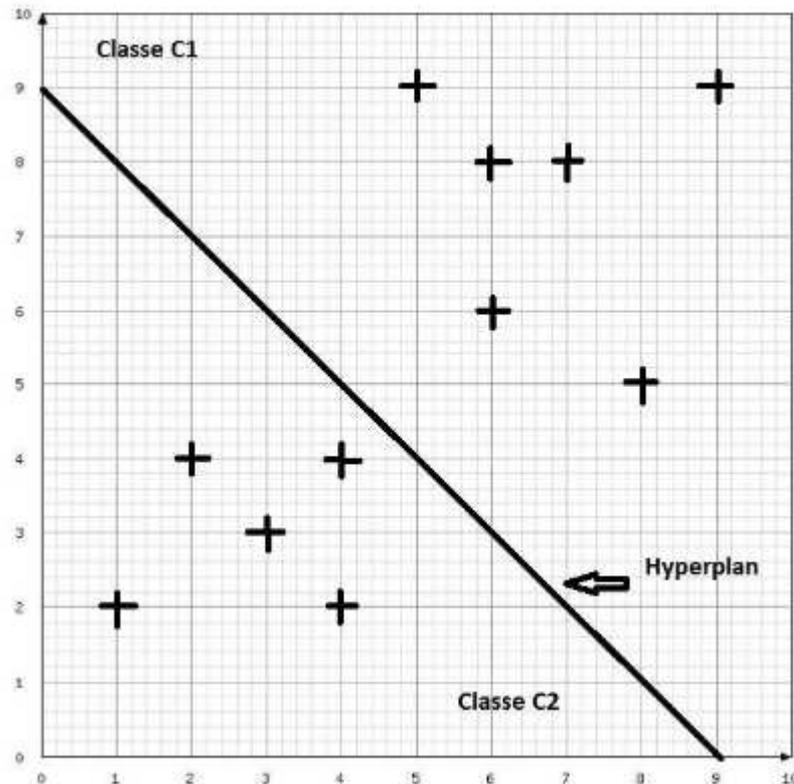
Durée : 01H30

Exercice 1 (04 points) :

Représentez sous forme d'un arbre de décision l'expression logique suivante : A et (B ou C).

Exercice 2 (04 points) :

La figure suivante représente la classification par SVM d'un ensemble de données. On y a représenté l'hyperplan séparateur entre les deux classes C1 et C2. Mais, deux points ont été volontairement masqués (cachés) : un point de la classe C1 et un point de la classe C2.



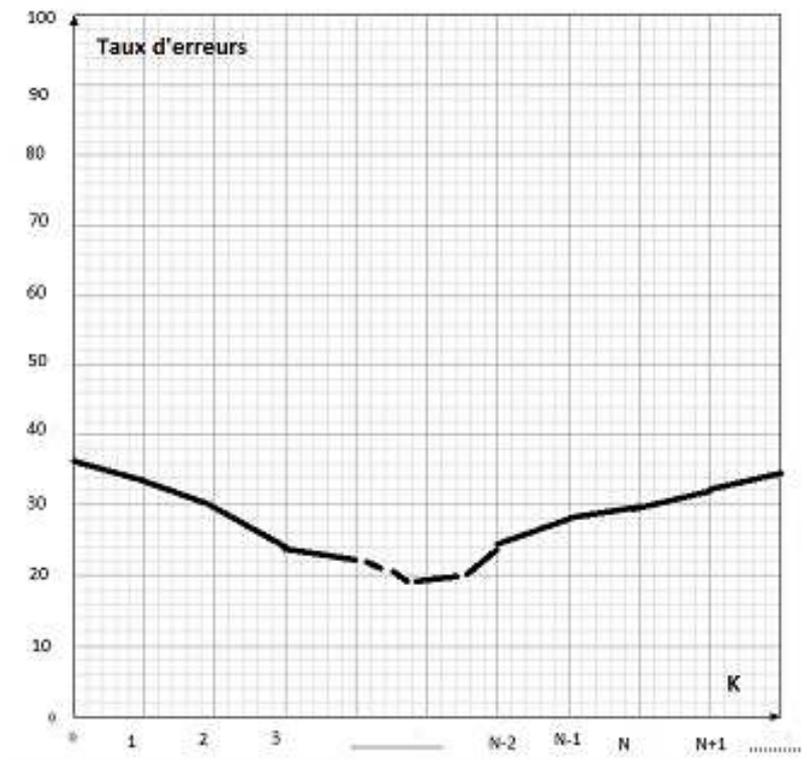
Question 1 : Donnez l'équation de l'hyperplan séparateur de la figure.

Question 2 : Pouvez-vous deviner quel est le point qui a été masqué de la classe C1 ? . Si oui, donnez ses coordonnées exactes, sinon situez-le approximativement .

Question 3 : Même question pour le point masqué de la classe C2.

Exercice 3 (04 points) :

La figure suivante représente le taux d'erreurs de classification, en utilisant l'algorithme KNN, en fonction de K sur un ensemble de N données.



Question : La courbe vous semble-t-elle normale ? Si oui donnez une explication de ses variations, sinon expliquez ce qui est anormal.

Exercice 4 (08 points) :

Dans un magasin de produits informatiques, on a enregistré les achats suivants faits par les clients (transactions) :

N°	Articles
01	CD, DVD, Jeu, Antivirus
02	CD, Jeu, Souris, Caméra, Antivirus
03	Antivirus, FlashDisk, CD, Jeu, Caméra
04	Jeu, DVD, Tablette
05	FlashDisk, Antivirus, Tablette, Souris

- 1/ Trouver tous les ensembles d'items fréquents vérifiant $minsup \geq 60\%$, en appliquant l'algorithme Apriori .
- 2/ Dédire toutes les règles d'associations .
- 3/ Quelles sont les règles ayant une confiance $\geq 75\%$?.