

M1 MABS – Examen de Fouille de Données – Janvier 2013

durée : 2h barème à titre indicatif (susceptible d'évoluer) Documents et téléphone portable interdits

Classification

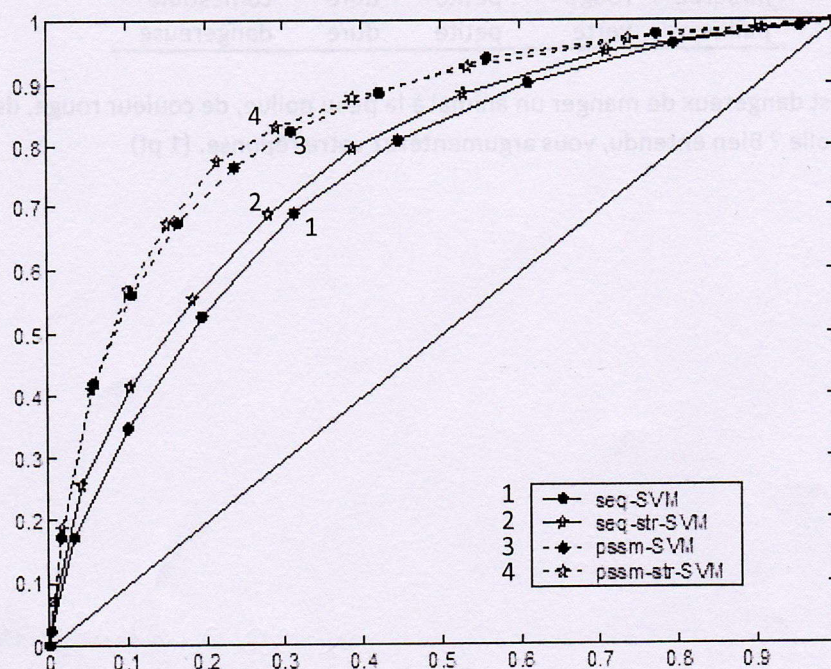
1. Faire une synthèse des différentes méthodes de classification abordées. Vous introduirez d'abord le but de cette tâche ainsi que les différentes étapes de sa mise en œuvre. Puis, pour chaque méthode, vous donnerez son principe, ses avantages et ses limitations. (6 pts)

2. Comment peut-on évaluer les performances d'un classificateur, d'une méthode de classification (éventuellement par rapport à d'autres), et de leurs paramètres ? (1 pt)

3. Qu'est-ce que la robustesse ? (1 pt)

4. Vous avez ci-dessous une courbe ROC. Expliquez le principe de son tracé et notamment à quoi correspondent les abscisses et les ordonnées ainsi que la diagonale qui n'est pas dans la légende ? A quoi peut donc servir ce type de graphique ? (1 pt)

5. Les 4 courbes correspondent à différentes méthodes et de jeux de paramètres pour la prédiction de sites de liaison sur une séquence ADN. Laquelle choisiriez-vous pour effectuer une prédiction et pourquoi ? (1 pt)



Clustering

6. Pourquoi normaliser les données avant de réaliser un partitionnement ? (0.5 pt)

7. Citez les différents types de données que l'on peut rencontrer et les mesures de dissimilarité que l'on peut utiliser pour chacun d'eux afin de calculer la dissemblance de 2 éléments. (3.5 pts)

8. Citez les méthodes et leur principe pour le calcul de dissimilarité entre 2 ensembles d'éléments (2 clusters). (2 pts)

Méthodes bayésiennes

9. Quelle est l'hypothèse faite pour un classificateur bayésien naïf ? Pourquoi est-elle (souvent) nécessaire ? (1 pt)

10. A partir du jeu de données d'apprentissage suivant, donnez le modèle (table de probabilités) construit par un classificateur bayésien naïf. (2 pt)

Peau	Couleur	Taille	Chair	Classe
poilue	brune	grande	dure	comestible
poilue	verte	grande	dure	comestible
imberbe	rouge	grande	molle	dangereuse
poilue	verte	grande	molle	comestible
poilue	rouge	petite	dure	comestible
imberbe	rouge	petite	dure	comestible
imberbe	brune	petite	dure	comestible
poilue	verte	petite	molle	dangereuse
imberbe	verte	petite	dure	dangereuse
poilue	rouge	grande	dure	comestible
imberbe	brune	grande	molle	comestible
imberbe	verte	petite	molle	dangereuse
poilue	rouge	petite	molle	comestible
imberbe	rouge	grande	dure	dangereuse
imberbe	rouge	petite	dure	comestible
poilue	verte	petite	dure	dangereuse

100. Est-ce qu'il est dangereux de manger un animal à la peau poilue, de couleur rouge, de grande taille, à la chair molle ? Bien entendu, vous argumenterez votre réponse. (1 pt)