M1 Informatique 19 avril 2012

UE Initiation à l'apprentissage automatique

(2h avec notes de cours, le barème est donné à titre indicatif)

Un certain nombre de figures sont demandées, respectez la numérotation demandée afin d'éviter toute ambigüité dans les réponses.

I Classification non supervisée (5pts)

Un nuage d'observations est formé des 10 points suivants dans R^2 :

$$y_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, y_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, y_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, y_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, y_5 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, y_6 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$y_7 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, y_8 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, y_9 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, y_{10} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

1.1 (3pts) Appliquez l'algorithme de quantification vectorielle pour décomposer cet ensemble de points en deux classes à partir du dictionnaire initial $D_0 = \{ y_6, y_7\}$ (la démonstration peut être faite en donnant une succession de figures illustrant **chacune** des étapes de l'algorithme, sans faire les calculs exacts de distance –**figure 1**-).

1.2 (1pt) Donnez explicitement le dictionnaire final.

1.3 (1pt) Donnez (sans calcul, mais avec justification) un dictionnaire initial qui conduirait à une autre partition de l'espace.

Pour les 3 exercices suivants, les six premiers points de la question précédente appartiennent à la classe ω_1 et les quatre derniers à la classe ω_2.

Notation, une observation est notée $y = \begin{pmatrix} u \\ v \end{pmatrix}$**.**

II Stratégie bayésienne et lois gaussiennes (8pts)

On suppose que les lois conditionnelles des observations sont des lois gaussiennes. Aucune hypothèse n'est faite sur $p(\omega_1)$ et $p(\omega_2)$

2.1 (1pt) Calculez les moyennes de chaque classe en supposant que les points donnés ci-dessus correspondent aux deux ensembles d'apprentissage.

Après quelques calculs, on suppose qu'il est obtenu :

$$\Sigma_1 = \Sigma_2 = \begin{pmatrix} 0{,}5 & 0 \\ 0 & 1 \end{pmatrix}$$

2.2 (2pt) Rappelez la règle du maximum de vraisemblance dans le cas gaussien et simplifiez l'expression au maximum (passage en –Log et élimination des constantes).

2.3 (3pts) Donnez l'équation de la frontière de décision en fonction de u et v (les coordonnées de l'observation y) et **développez les calculs**.

2.4 (1pt) En supposant « $p(\omega_1)= p(\omega_2)$ », tracez sur une nouvelle figure (**figure 2**) les points, les moyennes et la frontière trouvée dans ce cas.

2.5 (1pt) Justifiez la nature et forme de cette frontière. Quelle est l'influence des probabilités a priori ? quelle est l'influence des matrices de covariance ?

III Stratégie bayésienne et règle de décision des kppv (3pts)

3.1 (1pt) Rappelez la règle de décision dite des k plus proches voisins. A quelle classe appartient le point $G = \begin{pmatrix} 0 \\ 1/5 \end{pmatrix}$ avec la règle du 3 plus proche voisin et avec la règle du 1 plus proche voisin ?

3.2 (2pts) Représentez sur une autre figure les zones de décision correspondant à la règle du plus proche voisin (k=1) (**figure 3**).

IV Hyperplan séparateur (4pts)

4.1 (1pt) On considère la droite D_0 d'équation $v = u + 1/2$. Vérifiez que les données initiales sont séparables linéairement et que cette droite est un hyperplan séparateur pour les deux classes ω_1 et ω_2.

4.2 (2pts) Un nouveau point $X = \begin{pmatrix} 0 \\ 0,4 \end{pmatrix}$ est introduit dans l'ensemble d'apprentissage de la classe ω_2. En utilisant l'algorithme du perceptron, les notations du cours et l'hyperplan séparateur D_0 :

4.2.1 Justifiez que $\delta_X = -1$

4.2.2 Justifiez par le calcul que ce point X est mal classé.

4.3 (1pt +1ptBonus) Appliquez une itération de l'algorithme du perceptron pour proposer une nouvelle droite séparatrice, pour un pas de 0.1.