



Partiel avec documents autorisés

Barème indicatif : Ex. 1 (Partie 1) (7pts), Ex. 1 (Partie 2) (7pts), Ex. 2 (8pts)

Exercice 1

On considère un problème de classification à deux classes équiprobables ω_1 et ω_2 avec la base d'apprentissage

$$\begin{aligned} \omega_1 &: \mathbf{x}_1 = (1, 1)^T \\ \omega_2 &: \mathbf{x}_2 = (-1, -1)^T, \mathbf{x}_3 = (1, 0)^T, \mathbf{x}_4 = (0, 1)^T \end{aligned}$$

Partie 1

1) **Classifieur Bayésien** : on suppose que les lois conditionnelles associées aux classes ω_1 et ω_2 sont définies comme suit

$$f(\mathbf{x} | \omega_1) = \mathcal{N}(\mathbf{m}_1, \sigma^2 \mathbf{I}_2) \text{ et } f(\mathbf{x} | \omega_2) = \mathcal{N}(\mathbf{m}_2, \sigma^2 \mathbf{I}_2)$$

où \mathbf{m}_1 et \mathbf{m}_2 sont deux vecteurs de \mathbb{R}^2 et où \mathbf{I}_2 est la matrice identité de taille 2×2 . Développer la règle de décision Bayésienne associée à ce problème et représenter les régions du plan associées aux décisions $d(\mathbf{x}) = \omega_1$ (on affecte le vecteur \mathbf{x} à la classe ω_1) et $d(\mathbf{x}) = \omega_2$ (on affecte le vecteur \mathbf{x} à la classe ω_2) en choisissant deux valeurs arbitraires de \mathbf{m}_1 et \mathbf{m}_2 . Quelle règle classique obtient-on ?

2) **Apprentissage** : comment peut-on estimer les vecteurs \mathbf{m}_1 et \mathbf{m}_2 à l'aide de la base d'apprentissage ? Représenter les régions du plan associées aux décisions $d(\mathbf{x}) = \omega_1$ et $d(\mathbf{x}) = \omega_2$ lorsque \mathbf{m}_1 et \mathbf{m}_2 sont remplacés par leurs estimateurs. Que pensez vous de ce classifieur ?

3) **Prétraitement** : on désire projeter les données sur une droite adaptée à la classification. Déterminer les matrices de dispersion intraclasse et interclasse associées à ce problème. En déduire un vecteur directeur de la droite résultant de la maximisation du critère de Fisher. Représenter les projections des vecteurs $\mathbf{x}_i, i = 1, \dots, 4$ sur cette droite.

Partie 2

1) Représenter les données d'apprentissage dans le plan ainsi que l'hyperplan séparateur donné par le classifieur SVM. Quels sont les vecteurs supports ?

2) On désire retrouver le résultat de la question précédente par le calcul.

2.1) Déterminer l'expression analytique de la fonction $U(\boldsymbol{\alpha})$ à optimiser, où $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_4)^T$ est le vecteur constitué des multiplicateurs de Lagrange associés au problème d'optimisation associé au classifieur SVM.

2.2) Quelles sont les contraintes que doivent satisfaire les coefficients $\alpha_i, i = 1, \dots, 4$?

2.3) En utilisant la contrainte d'égalité liée au classifieur SVM ainsi que le fait que les multiplicateurs de Lagrange associés aux vecteurs non supports sont nuls, montrer que la détermination de l'hyperplan séparateur peut se ramener à l'optimisation d'un critère $V(\alpha_3, \alpha_4)$.

2.4) Déterminer le vecteur $(\alpha_3, \alpha_4)^T$ maximisant $V(\alpha_3, \alpha_4)$.

2.5) En déduire l'équation de l'hyperplan séparateur et retrouver le résultat de la question 1).

Exercice 2 : Questions portant sur l'article

- 1) Dans le problème d'optimisation défini par (1), que traduit la contrainte $\|\alpha_l\|_0 \leq L$? Déterminer $\|\alpha_l\|_0$ lorsque $K = 5$ et $\alpha_l = (1, 0, 0, \sqrt{3}, 0)^T$.
- 2) Que représente $R^*(\mathbf{x}_l, D_j)$ dans (2) ?
- 3) Dans la référence [21], la fonction de coût $c_i^\lambda(y_1, \dots, y_N)$ utilisée dans (2) est précisée

$$c_i^\lambda(y_1, \dots, y_N) = \log \left[\sum_{j=1}^N e^{-\lambda(y_j - y_i)} \right].$$

Quand est ce que le terme

$$c_i^\lambda [R^*(\mathbf{x}_l, D_1), \dots, R^*(\mathbf{x}_l, D_N)]$$

est "petit" ?

- 4) Expliquer comment les dictionnaires représentés sur la figure 2 ont été obtenus.
- 5) Comment les auteurs de l'article proposent-ils de détecter les contours en utilisant la transformée en ondelettes ?
- 6) Après avoir déterminés les deux dictionnaires D_1 et D_2 associés respectivement aux images de texte et de background, comment peut-on détecter les zones de texte dans une image ?
- 7) Expliquer l'analyse effectuée pour montrer que la méthode de l'article est sensible à la taille des caractères et la modification apportée par les auteurs pour que la méthode fonctionne pour différentes tailles de caractères.