SID L3-Learning

Examen du 18 avril 2017

Durée 2h-Documents autorisés: notes et planches de cours

Apprentissage d'une fonction sur $[0,1]$

On souhaite prédire une relation entrée-sortie $Y = F(X)$ où $X \in [0,1]$ et $Y \in \mathbb{R}$. L'échantillon d'apprentissage est:

$$(0.2, F(0.2)),\ (0.43, F(0.43)), (0.62, F(0.62)), (0.75, F(0.75)), (0.87, F(0.87)).$$

On utilise la méthode des plus proches voisins. Soit $\widehat{F}(X)$ la prédiction obtenue au point $X \in [0,1]$.

1. On se place dans le cas d'un seul voisin. Calculer $\widehat{F}(X)$.
2. On suppose que $F(X) = X + 3$. Représenter graphiquement sur un même schéma les fonctions F et $\widehat{F}$.
3. Utiliser la méthode du Jacknife pour évaluer l'erreur quadratique de l'estimateur.
4. Reprendre les 2 questions précédentes pour les 2 plus proches voisins.
5. Quel voisinage (1 ou 2) préconisez-vous pour prédire la fonction $X + 3$?

Beat bits

Soit X et Y deux vecteurs aléatoires indépendants à valeurs dans $\{0,1\}^d$ ($d \in \mathbb{N}$). On suppose que les composantes de X (respectivement de Y) sont des variables aléatoires indépendantes de même loi. On pose $\mathbb{P}(X_j = 1) = \theta_1$ et $\mathbb{P}(Y_j = 1) = \theta_2$, ($j = 1, \cdots, d$ et $0 \leq \theta_1, \theta_2 \leq 1$).

1. Que vaut $\mathbb{P}(X_j = Y_j)$, ($j = 1, \cdots, d$)?
2. Soit $Z = d(X, Y)$ où d est la distance de Hamming sur $\mathbb{R}^d$. Quelle est la loi de Z?
3. On suppose que $\theta_1 = \theta_2$. Que vaut l'espérance de Z? Pour quelle valeur de θ_1, cette espérance est-elle maximale? Comment interpréter ce résultat?