

Feuille d'exercice Machine Learning

Exercice 1

On considère un couple de variables aléatoires (X, Y) réelles. On se donne la fonction de perte suivante

$$\ell(Y, \hat{Y}) = \begin{cases} k_2(Y - \hat{Y}) & \text{si } Y > \hat{Y} \\ k_1(Y - \hat{Y}) & \text{sinon.} \end{cases}$$

Déterminez le prédicteur de Bayes à l'aide de la loi de $Y|X$

Exercice 2 *

Supposons que l'on observe N observations Y_i d'un vecteur aléatoire dans \mathbb{R}^d . On définit l'estimateur LASSO

$$\hat{\beta} = \operatorname{argmin}_b \sum ||Y_i - b||^2 + \lambda ||b||_1$$

Donnez une formule explicite des coordonnées de $\hat{\beta}$. (Plus facile : même exercice sans pénalisation, et pour la pénalisation ridge vu en cours $\lambda ||b||_2^2$.)

Exercice 3

Trouver la VC-dimension des classificateurs suivants :

1. $S = \{(-\infty, b] : b \in \mathbb{R}\}$.
2. $S = \{(a, b] : a, b \in \mathbb{R}\}$.
3. $S = \{(-\infty, b_1] \times (0, b_2] : b_1, b_2 \in \mathbb{R}\}$
4. $S = \{[a, b] \times [c, d] : a, b, c, d \in \mathbb{R}\}$

Exercice 4

Démontrer que, en reprenant les notations du cours

$$\mathcal{R}(f_{\text{erm}, n, \mathcal{F}}) - \min_{f \in \mathcal{F}} \mathcal{R}(f) \leq 2 \sup_{f \in \mathcal{F}} |\mathcal{R}(f) - \mathcal{R}_n(f)|$$