

EXAMEN

Apprentissage

Durée 1 heure- Notes de cours autorisées

1 Voisins

Soit \mathcal{R} la région du plan définie par

$$\mathcal{R} := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq x_1^2 \right\}.$$

1. Tracer dans le plan d'axes x_1 (abscisses), x_2 (ordonnées) la courbe $x_2 = x_1^2$.
2. Le point $\begin{pmatrix} 1/2 \\ 1/8 \end{pmatrix}$ appartient-il à \mathcal{R} ?
3. Représenter la région \mathcal{R} .
4. On suppose que l'on souhaite apprendre la variable $Y = \mathbf{1}_{\mathcal{R}}$ à l'aide des 4 points d'apprentissage $\begin{pmatrix} 1/2 \\ 1/8 \end{pmatrix}$, $\begin{pmatrix} 1/4 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1/2 \\ 3/4 \end{pmatrix}$ et $\begin{pmatrix} 3/4 \\ 1 \end{pmatrix}$. On considère les deux points de validation $\begin{pmatrix} 1/3 \\ 1/27 \end{pmatrix}$ et $\begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}$. Pour l'échantillon de validation déterminer les prédictions au 1 et aux 3 plus proches voisins.

2 Régression linéaire

On considère X_1 et X_2 deux variables aléatoires indépendantes de loi uniforme sur $[-1, 1]$. Soit

$$Y := 2X_1 + X_1X_2^2.$$

Quel est le meilleur prédicteur linéaire \hat{Y} de Y quand on observe uniquement X_1 ? Calculer $\mathbb{E}[(Y - \hat{Y})^2]$.