

Exercice - Session 2

Le pourcentage de viande maigre de carcasses de porc (LMP pour Lean Meat Percentage) est le principal critère sur lequel la valeur commerciale de la viande est calculée. Dans certains abattoirs, ce critère est prédit à partir des données issus d'un scanner à imagerie tomographique qui fournit un profil du pourcentage de viande maigre le long de la carcasse. Pour estimer un modèle de prédiction du pourcentage de viande maigre à partir de ce profil, on dispose de données expérimentales consistant en 117 observations conjointes du profil fourni par le scanner et du pourcentage de viande maigre. Ces données sont disponibles dans le fichier `scanner.txt`

La problématique est donc la suivante : comment prédire le pourcentage de viande maigre à partir des données du scanner ?

1. Proposez une méthode de sélection des meilleures variables prédictrices, parmi celles données par le scanner.
2. La méthode mise en oeuvre à la question précédente garantit-elle la sélection d'un modèle performant en matière de précision de la prédiction ?
3. La liste des variables sélectionnées est-elle reproductible avec de nouvelles données ? Argumentez votre réponse.
4. Évaluez la précision de la prédiction par la méthode mise en oeuvre à la question 1.