

VALIDATION CROISÉ

- Principe de Validation croisé :
 - La validation croisée permet de mesurer la performance d'un modèle sans qu'il y ait de risque de sur-apprentissage
 - Principe :
 - On coupe la base de données B en K groupes de taille N/K : G_1, G_2, \dots, G_K
 - Pour i allant de 1 à K , on calibre un CART sur $B \setminus G_i$, puis on utilise le modèle calibré pour prédire Y pour les individus de G_i
 - Après la boucle, on dispose d'une prédiction pour tous les individus de B . On peut alors calculer le R^2 de cette prédiction (ou d'autres statistiques)
 - Remarque : On peut répéter la procédure de validation croisée plusieurs fois, et moyenner les résultats pour obtenir une meilleure stabilité des résultats