

Cross-val

- **Avantages :**
 - **Utilisation efficace des données :** Utilise toutes les données pour l'entraînement et le test, cool si données limitées.
 - **Estimation plus fiable de la performance validation du modèle :** test sur plusieurs sous-ensembles
- **Inconvénients :**
 - **Coût computationnel plus élevé :**
 - **Mise en oeuvre de complexité accrue :**
- **Intérêt par rapport au découpage train/test classique :**
 - La validation croisée : vision plus complète et plus robuste de la performance du modèle, réduisant le risque de surajustement (overfitting) sur un unique ensemble de test.
 - Cela la rend particulièrement utile pour les comparaisons de modèles et la sélection de paramètres, où une évaluation précise de la capacité de généralisation est cruciale.
 - En revanche, un simple découpage train/test peut conduire à une estimation de la performance du modèle qui dépend fortement de la manière dont les données sont divisées, ce qui peut parfois mener à des évaluations optimistes ou pessimistes de la performance réelle du modèle.

Rééchantillonnage aléatoire (bootstrap)

Cette méthode consiste à créer plusieurs échantillons à partir d'un ensemble de données original en sélectionnant des observations de manière aléatoire avec remise. Cela signifie qu'une même observation peut apparaître plusieurs fois dans un même échantillon. À partir de ces échantillons rééchantillonnés, le modèle est entraîné et testé, permettant ainsi d'évaluer sa performance et de mesurer l'incertitude de ses prédictions.

Avantages :

- Plus grande variance dans les folds, encore mieux pour réduire les biais

Inconvénients :

- Biais possible à cause des données dupliqués
- Coût computationnel comme cross val

Intérêt par rapport au découpage train/test classique :

- Estimation plus robuste de la variance des métriques de performance du modèle.
- évitant ainsi le problème de "gaspillage" de données qui peut se poser avec un découpage train/test fixe, où une partie des données n'est utilisée que pour le test et ne contribue pas à l'entraînement du modèle.

Cross-val vs. Bootstrap :

L'intérêt du bootstrap par rapport à la validation croisée, et inversement, dépend des objectifs spécifiques de l'analyse, de la nature des données disponibles, et des contraintes pratiques comme le temps de calcul et la puissance de calcul disponible. Voici une comparaison des deux méthodes selon différents critères :

Intérêt du Bootstrap par rapport à la Validation Croisée

bootstrap

- Estimation de l'incertitude parfait: Le bootstrap excelle dans l'estimation de l'incertitude des estimations, comme l'erreur standard, les intervalles de confiance et la variance des prédictions. Cette caractéristique est particulièrement utile pour les études où comprendre la variabilité des estimations est aussi important que les estimations elles-mêmes.
- **Flexibilité dans les tailles d'échantillon :**
- **Simplicité pour certains types de données :** Pour des données ou des modèles où la validation croisée peut être difficile à mettre en œuvre (par exemple, avec des séries temporelles fortement dépendantes), le bootstrap peut offrir une alternative plus simple, bien que des adaptations soient nécessaires pour ces cas spécifiques.
- Validation croisé
 - **Réduction du biais :** La validation croisée tend à introduire moins de biais dans l'estimation de l'erreur de généralisation, car chaque observation est utilisée à la fois pour l'entraînement et le test exactement une fois dans le cadre d'une validation croisée k-fold. Cela contraste avec le bootstrap, où certaines observations peuvent ne jamais être sélectionnées dans un échantillon rééchantillonné, tandis que d'autres peuvent apparaître plusieurs fois.
 - **Meilleure estimation de la performance de généralisation :** La validation croisée est souvent considérée comme fournissant une meilleure estimation de la capacité d'un modèle à généraliser à de nouvelles données, car elle force le modèle à prouver sa performance sur l'ensemble des données à travers les différents plis.
 - **Adaptabilité à différents types de données :** La validation croisée stratifiée et d'autres variantes permettent d'adapter la méthode à des situations spécifiques, comme les ensembles de données déséquilibrés ou les problèmes de classification, où maintenir la proportion des classes dans chaque pli est important.

Choix entre les deux méthodes

- **Nature des données et objectifs de l'étude :** Si l'objectif est d'estimer avec précision l'erreur de généralisation et de minimiser le biais, la validation croisée est souvent préférée. Si l'objectif est d'évaluer l'incertitude autour des estimations ou de travailler avec un ensemble de données de petite taille, le bootstrap peut être plus approprié.
- **Complexité computationnelle :** Pour les grands ensembles de données ou les modèles complexes, le coût computationnel de la validation croisée peut être prohibitif, surtout si un grand nombre de plis est utilisé. Le bootstrap peut parfois être plus rapide, bien que cela dépende fortement de la taille de l'échantillon rééchantillonné et du nombre de réplications.

En résumé, le choix entre le bootstrap et la validation croisée dépend de l'équilibre entre la précision de l'estimation de performance, la compréhension de l'incertitude des estimations, et les contraintes pratiques de l'étude.

Leave-one out :

- Choix LOOCV : taille du dataset, la complexité du modèle, et des ressources computationnelles
- Intérêt du Leave-One-Out par rapport à la Validation Croisée Classique
 - La LOOCV peut être préférable pour les petits ensembles de données ou quand une estimation précise du biais est critique
 - **Minimisation du biais :** utile pour les petits ensembles de données
 - **Variabilité de l'estimation :** Réfléter la variabilité de la performance du modèle sur différents sous-ensembles de données.
 - **Pas de choix arbitraire du nombre de folds :**

Influence du Nombre de Folds

- **Validation croisée classique :** influencer à la fois la biais et la variance de l'estimation de la performance du modèle. Avec moins de folds, chaque fold d'entraînement est plus grand, ce qui peut réduire le biais mais augmenter la variance de l'estimation de l'erreur, car moins de répétitions sont utilisées pour évaluer la performance. À l'inverse, un plus grand nombre de folds augmente le coût computationnel mais peut réduire la variance de l'estimation de l'erreur, tout en introduisant un biais potentiellement plus élevé si le modèle est très sensible aux variations dans les données d'entraînement.
- **LOOCV :** En utilisant (N) folds, la LOOCV minimise le biais en maximisant la taille de l'ensemble d'entraînement, mais peut avoir une variance élevée dans l'estimation de l'erreur de test, surtout pour des modèles très flexibles ou lorsque les données sont très hétérogènes.

Discussion

- **Coût computationnel :**
- **Sensibilité aux outliers :** peut fortement influencer l'estimation de la performance du modèle.
- **Choix du modèle :**

Un modèle a été déployé en production, après avoir montré des performances satisfaisantes lors de sa mise au point. Proposez une méthode pour s'assurer qu'il continue à fonctionner correctement pendant toute sa durée de vie ?

- Surveillance en Temps Réel
 - Suivi des performances
 - Détection des anomalies
- Validation Continue
 - Test A/B: Si possible, réaliser des tests A/B en dirigeant une partie du trafic vers le nouveau modèle et une autre partie vers l'ancien modèle ou un modèle de contrôle, afin de comparer les performances en conditions réelles.
 - Retest régulier avec de nouvelles données
- Gestion des Données
 - Vérification de la qualité des données (dégradation possible): valeurs manquantes outliers, changements dans la distribution des données, (concept drift)
- Mise à Jour et Ré-entraînement (automatique)
- Feedback des Utilisateurs

Comment faire un modèle puissant lorsque peu de données sont disponibles ?

- Transfert d'apprentissage (Transfer Learning)
- Apprentissage par renforcement (Few-Shot Learning) ou One-Shot Learning
- Augmentation des données (Data Augmentation)
- Régularisation et Architectures de Modèle Simplifiées Domain adaptation