

ENG-209

# Data science pour ingénieurs avec Python

*Cours 11 - Quiz*

*Jean-Philippe Pellet, Éric Bouillet, Olivier Verscheure*

24 Novembre 2025

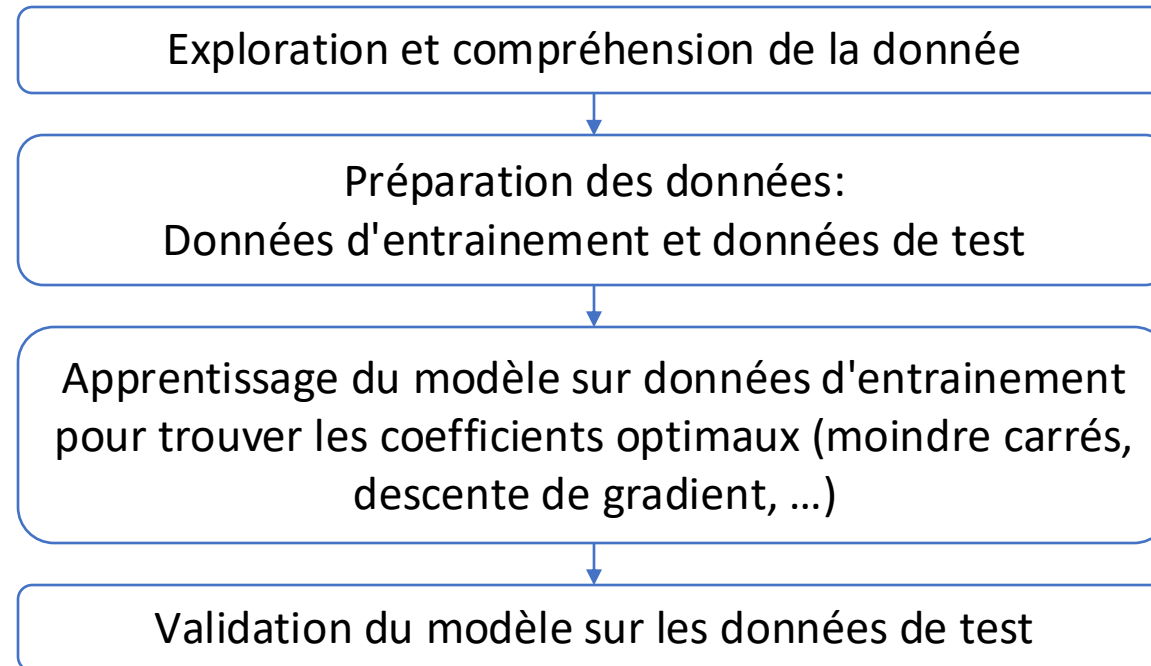
# Sélection de modèles et sélection des caractéristiques

---

- Sélection de modèles: l'hyper-paramétrisation par validation croisée
- Sélection des caractéristiques
- Séance d'exercices
- (Midi): énoncé du projet final #1

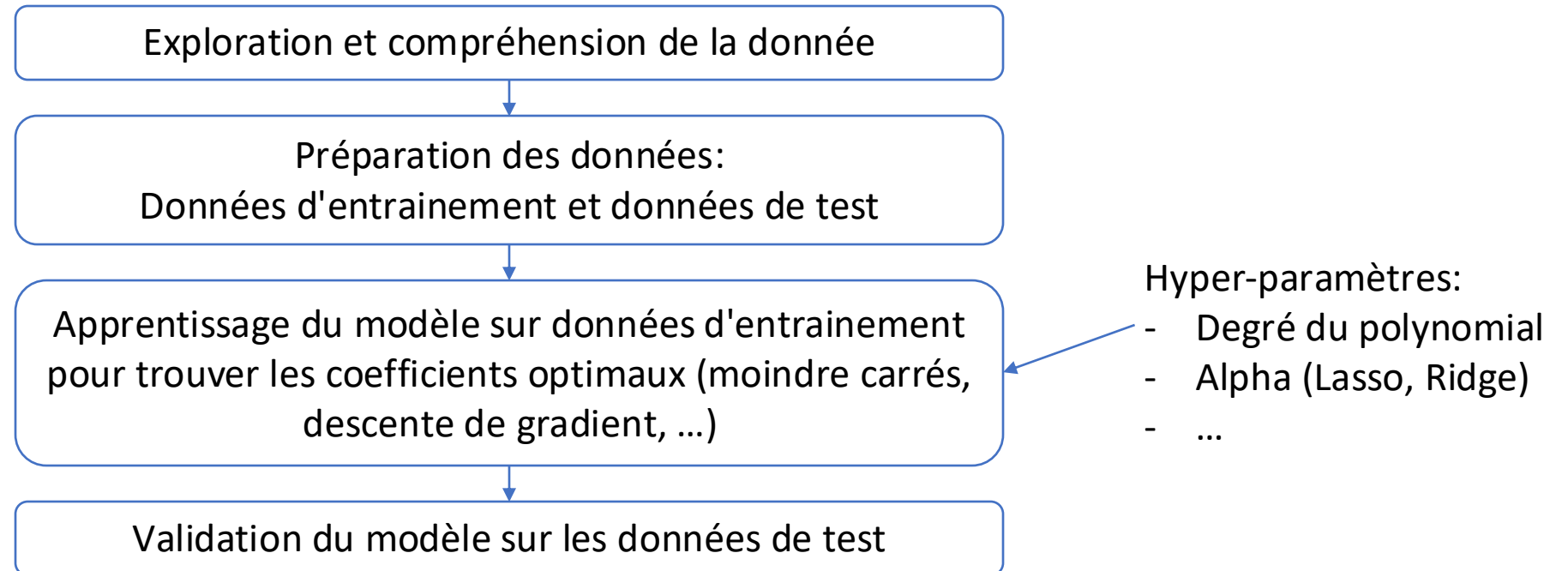
# Les étapes de la modélisation vues en lab

---



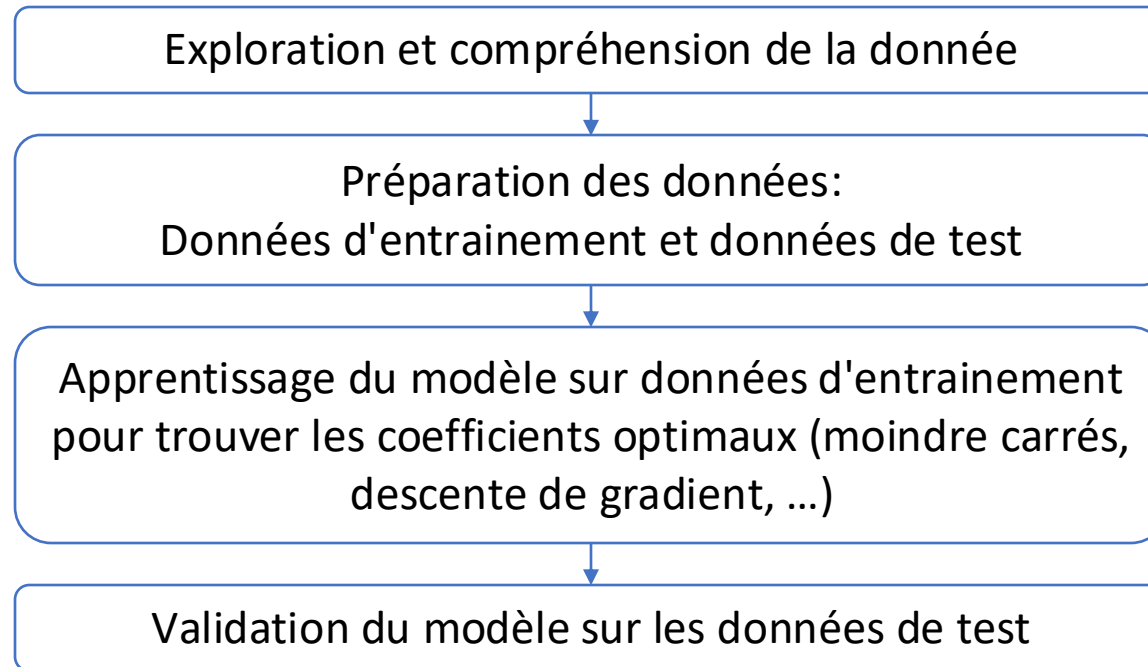
# Les étapes de la modélisation vues en lab

---



# Les étapes de la modélisation vues en lab

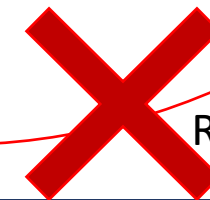
(Régression polynomiale)



Hyper paramètres:

- Degré du polynomial
- Alpha (Lasso, Ridge)
- ...

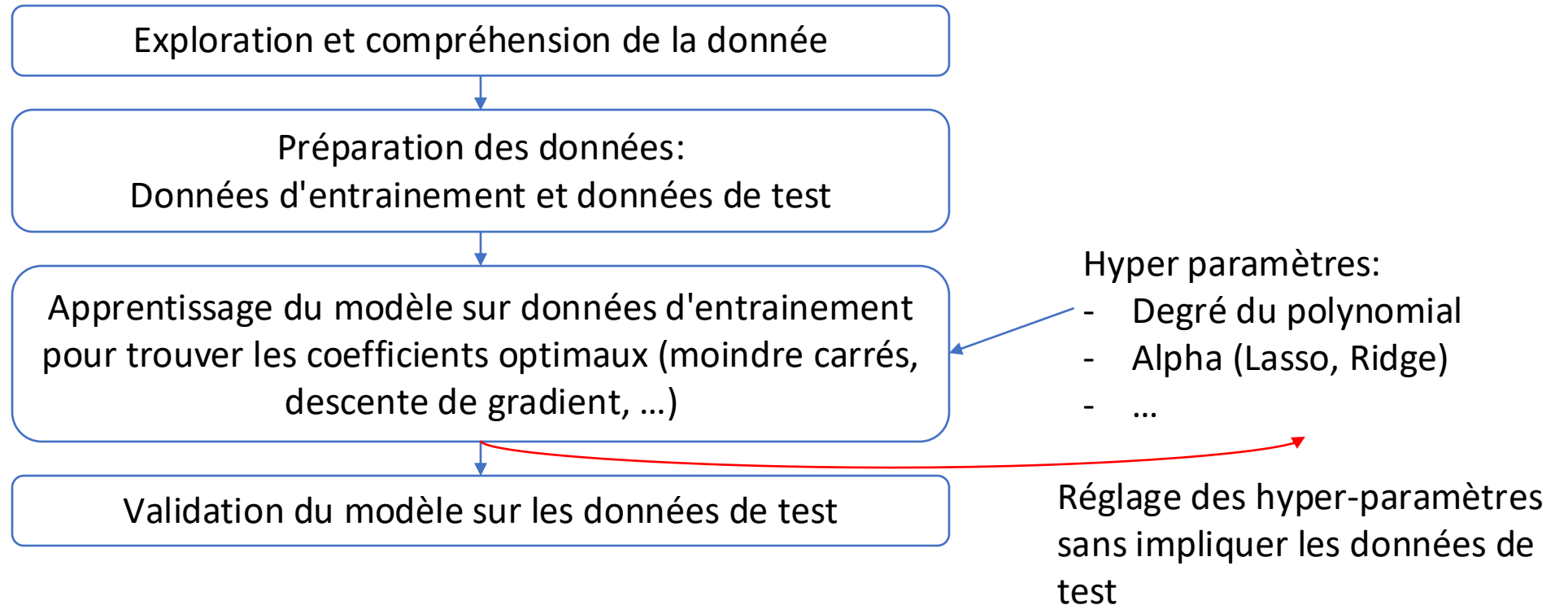
Attention: fuite des données de test dans l'apprentissage, la validation n'est plus fiable



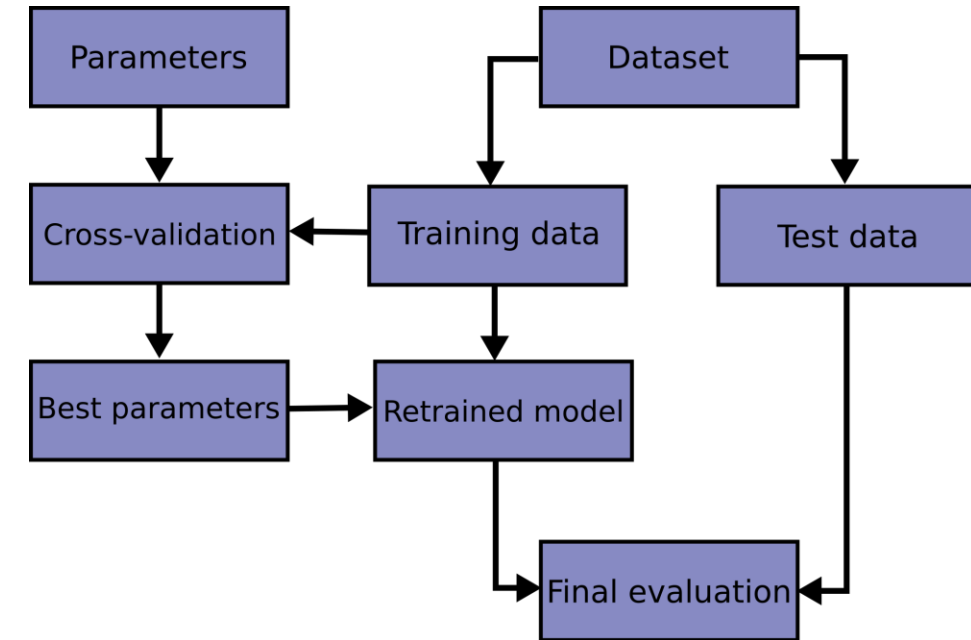
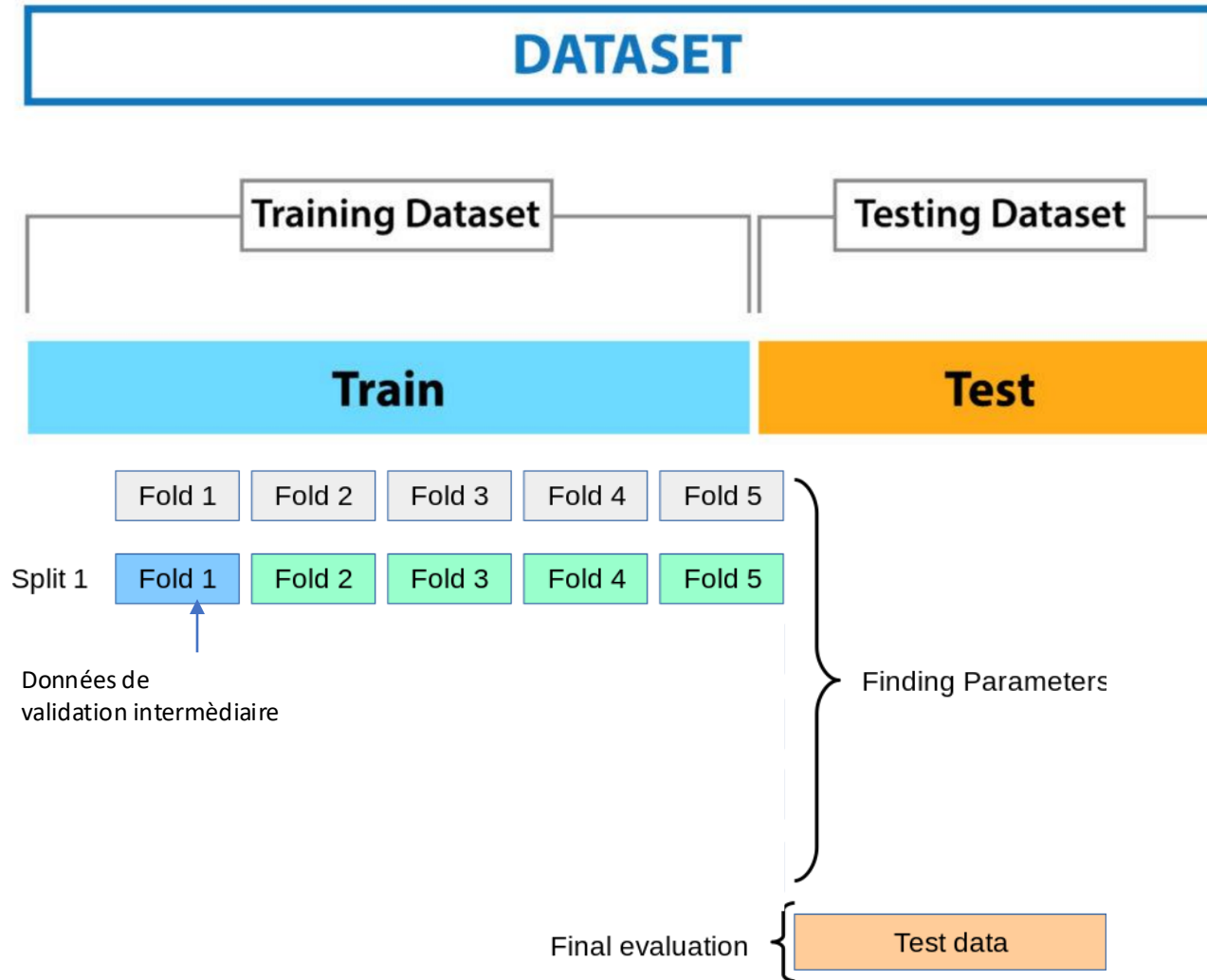
Réglage des hyper paramètres

# Hyper-paramétrisation – ce qu'il faut faire

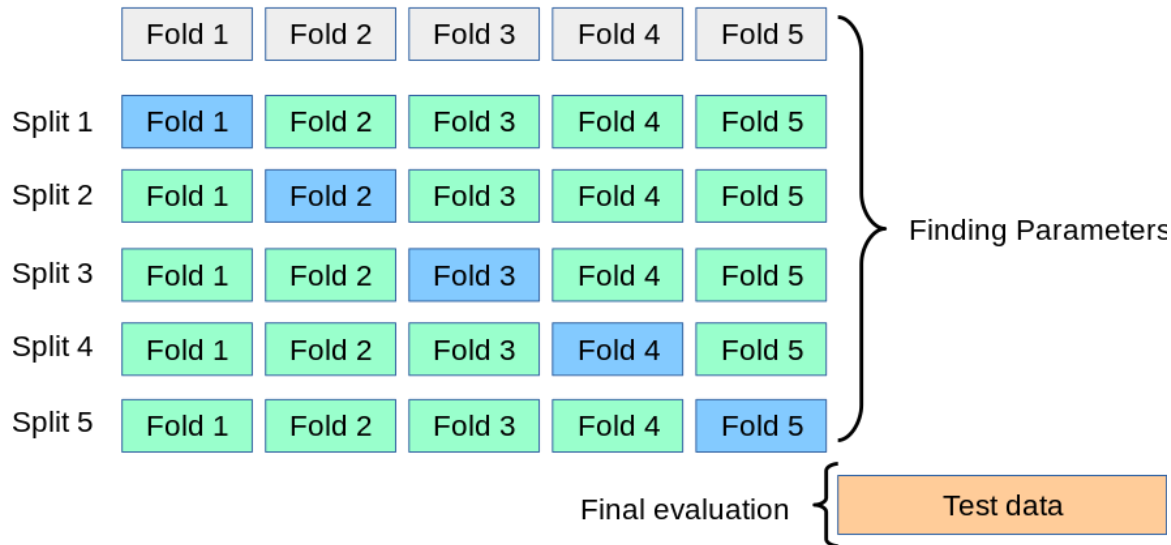
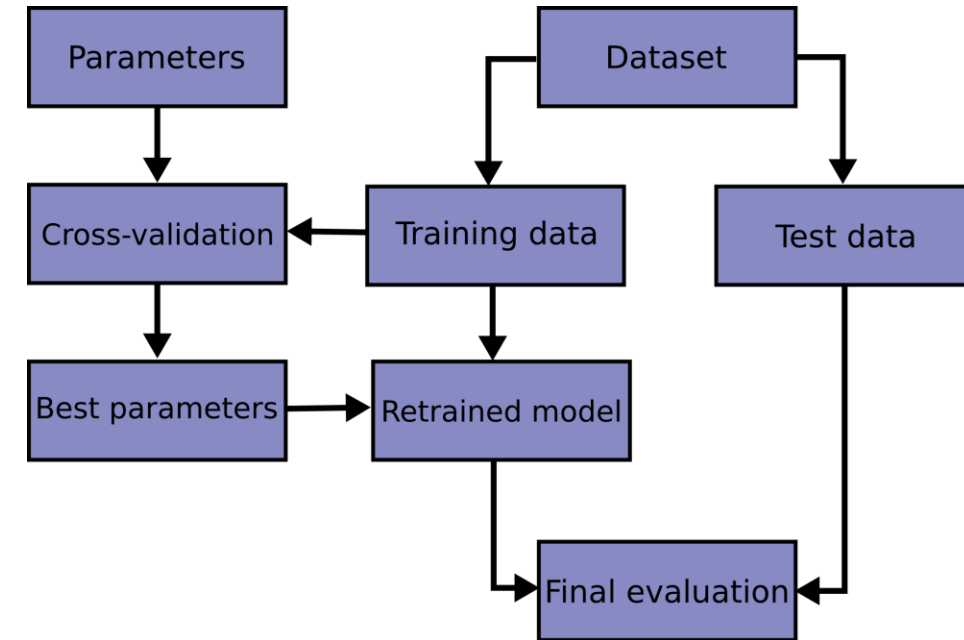
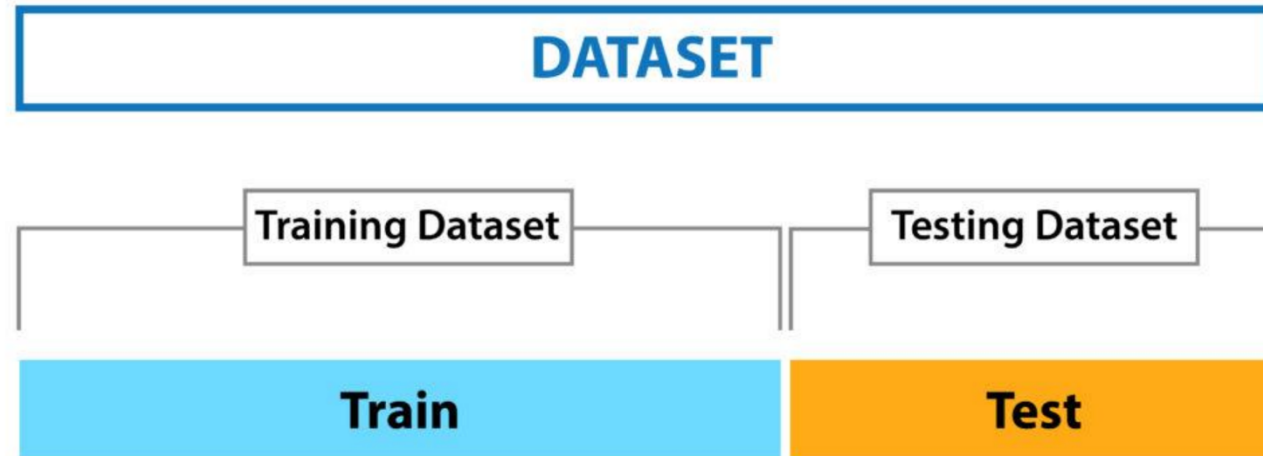
---



# Hyper-paramétrisation par validation croisée

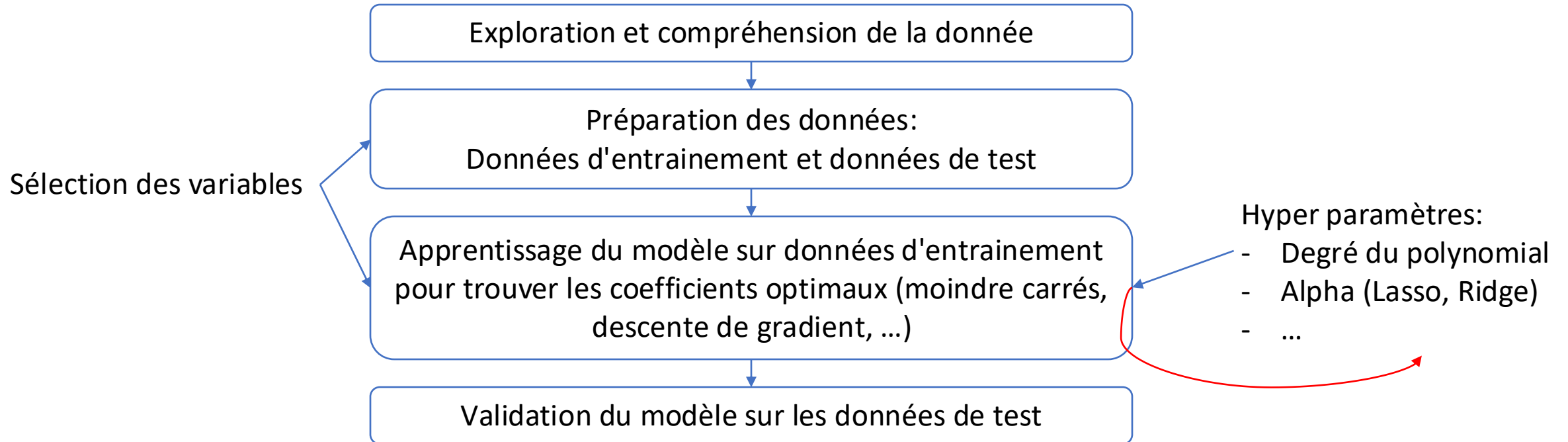


# Hyper-paramétrisation par validation croisée



- Validation croisée en k blocs (k-Fold Cross-Validation)
- Validation croisée "Leave-One-Out" (LOOCV)
- k-Fold stratifié
- Division temporelle (Time-Series Split)

# Sélection des variables indépendantes



La sélection des variables indépendantes peut se faire à deux moments, au cours de la préparation des données ou au moment de l'apprentissage.

# Sélection des variables indépendantes - RFE

---

- **Recursive Feature Elimination**

- ✦ Méthode de Sklearn

- ✦ Invocation: `RFE(model)`, par exemple `RFE(LinearRegression())`

- ✦ Méthode itérative:

- Après chaque itération, élimine N caractéristique

- Entraîne le modèle avec ces caractéristiques

- Répète jusqu'à ne garder que le nombre de caractéristiques désiré

- **C'est une méthode parmi d'autres méthodes**

- SelectKBest, etc.