

Vérification ponctuelle des algorithmes de RÉGRESSION

+ sélection de modèles

Atelier #11





Qu'est-ce que la vérification ponctuelle (*"spot-check"*)?

Définition de vérification ponctuelle (*"spot-check"*)

C'est une méthode qui compare plusieurs algorithmes entre eux dans le but d'identifier celui ou ceux qui performant le mieux pour la problématique d'apprentissage automatique à solutionner et les données disponibles.



Pourquoi faire une vérification ponctuelle?

- Très difficile de savoir à l'avance quel algorithme est le meilleur pour un problème d'apprentissage automatique.
- Pour se concentrer dès le début sur un ou quelques algorithmes qui sont les plus prometteurs.

Algorithmes de régression

Terminologie		Scikit-learn		
Français	Anglais	Module	Documentation	Algorithme
Régression Linéaire	<i>Linear Regression</i>	linear_model	https://scikit-learn.org/stable/modules/linear_model.html#ordinary-least-squares	<i>LinearRegression</i>
Régression de Crête	<i>Ridge Regression</i>	linear_model	https://scikit-learn.org/stable/modules/linear_model.html#ridge-regression-and-classification	<i>Ridge</i>
Régression Linéaire Lasso	<i>Lasso Linear Regression</i>	linear_model	https://scikit-learn.org/stable/modules/linear_model.html#lasso	<i>Lasso</i>
Régression Elastic Net	<i>Elastic Net Regression</i>	linear_model	https://scikit-learn.org/stable/modules/naive_bayes.html#gaussian-naive-bayes	<i>ElasticNet</i>
K plus proches voisins	<i>k-Nearest Neighbors (KNN)</i>	neighbors	https://scikit-learn.org/stable/modules/neighbors.html#nearest-neighbors-regression	<i>KNeighborsRegressor</i>
Arbres de décision	<i>Decision Trees</i>	tree	https://scikit-learn.org/stable/modules/tree.html#regression	<i>DecisionTreeRegressor</i>
Machines à vecteurs de support (MVS)	<i>Support Vector Machine (SVM)</i>	svm	https://scikit-learn.org/stable/modules/svm.html#regression	<i>SVR</i>



Sélection de modèles

Définition : Comparer plusieurs modèles entres eux en utilisant un métrique commun dans le but de sélectionner le ou les meilleur(s) modèle(s).

Étapes :

1. Définir les algorithmes à tester.
2. Sélectionner le métrique approprié à la problématique.
3. Faire rouler les modèles de base, sans hyperparamétrisation, tous ensemble.
4. Classer les modèles du meilleur au pire.
5. Sélectionner le modèle le plus performant.
 - Si plusieurs modèles ont une performance équivalente, voir la leçon 13.

**NE VOUS EN FAITES PAS SI
VOUS NE COMPRENEZ PAS.
- ANDREW NG**



Note

Certains algorithmes seront plus approfondis, mais on ne va pas approfondir le fonctionnement de tous les algorithmes qui seront vu dans le cadre théorique de ce cours.