**GRID SEARCH (Recherche par quadrillage)**

Permet de régler avec finesse nos modèles (trouver les valeurs optimales pour un modèle donné).

Rappel :

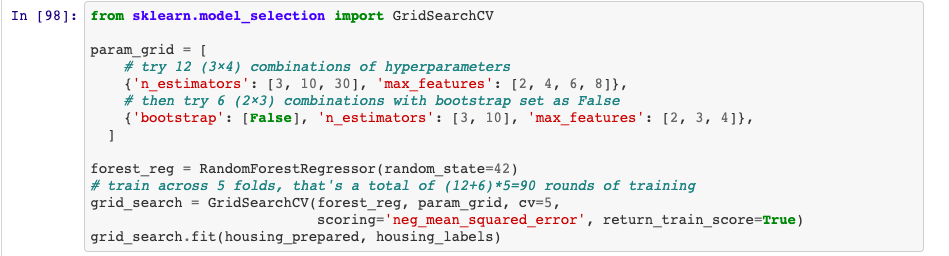
Un hyperparamètre est un paramètre de l’algorithme d’apprentissage (et non du modèle). De ce fait, il n’est pas affecté par l’algorithme d’apprentissage lui-même.

Il doit être défini avant l’algorithme d’apprentissage et reste constant au cours de celui-ci.

On veut trouver la meilleure combinaison de valeurs des hyperparamètres afin que le modèle soit précis. GridSearchCV de Scikit Learn cherche cette combinaison pour nous (ce serai fastidieux de jouer manuellement avec les hyperparamètres jusqu’à qu’on trouve une excellente combinaison de valeurs).

On doit uniquement indiquer à GridSearchCV quels sont les hyperparamètres que l’on souhaite faire varier et quelles valeurs utiliser. GridSearchCV évaluera pour nous toutes les combinaisons possibles des valeurs des hyperparamètres par validation croisée.

EXEMPLE :



Ici, param\_grid indique à Scikit-Learn d’évaluer d’abord les 3\*4 = 12 combinaisons des valeurs des hyperparamètres n\_estimators et max\_features spécifiées dans le premier dictionnnaire. Puis il va essayer les 2 \* 3 = 6 combinaisons des valeurs des hyperparamètres du second dictionnaire.

Dans cet exemple, au total, la recherche par quadrillage va explorer 12 + 6 = 18 combinaisons de valeurs des hyperparamètres de RandomForestRegressor et entraînera 5 fois chaque modèle (car cv=5 c’est une validation croisée à 5 passes).  A [cross validation process](https://medium.com/datadriveninvestor/k-fold-cross-validation-6b8518070833) is performed in order to determine the hyper parameter value set which provides the best accuracy levels.

Il y aura au total 18 \* 5 = 90 tours d’entraînement. Une fois terminé on peut obtenir la meilleure combinaison de paramètres.

Article explicatif : <https://medium.com/datadriveninvestor/an-introduction-to-grid-search-ff57adcc0998>