

Classifications (Machine Learning)

-Classification Binaire (supervisée) :

On dispose d'un échantillon (X_1, \dots, X_p) que l'on représente par des points dans un espace de dimension M . (M = nombres de variables communes qui décrivent chaque $X_i = (x_{i1}, x_{i2}, \dots, x_{iM})$)

Cas binaire : Chaque X_i possèdent deux types de labels Y_1, Y_2 déjà définis (classification supervisée). Le but est de pouvoir décider (prédire) le label à attribuer à un point X selon ces deux catégories à l'aide de paramètres estimés sur un échantillon test.

Méthodes Classiques : Régression Linéaire, Régression Logistique, SVM, Boosting.

1) Régression linéaire :

On sépare l'espace dans lequel évoluent les individus (caractérisés par leurs variables/coordonnées) en deux à l'aide d'un hyperplan

2) Régression logistique:

3) SVM

4) Boosting (Adaboost)