

# Advanced Machine Learning

---

## Classification

---

Le but d'une classification est d'obtenir, à l'aide d'un ensemble plus ou moins grand de variables comprise dans  $X$ , une fonction nous permettant par la suite de calculer une variable résultat notée  $Y$ . Pour se faire il existe plusieurs méthodes détaillées ci-dessous.

## Classification des moindres carrés

---

Cette méthode, qui peut être utilisée avec des valeurs pour  $X$  et  $Y$  continues et distinctes, considère que pour toutes suite d'individus  $X$  (individu positionné en ligne avec leur caractéristiques en colonne) ayant leur résultat respectif dans  $Y$ , il y a une matrice  $A$  et une matrice  $E$  tel que l'on peut écrire :

- $Y = XA + E$

où  $Y$  est de la forme :  $Y = (y_1, y_2, y_3...)$ ,  $X = [(x_{11}, x_{12}, x_{13}...), (x_{21}, x_{22}, x_{23}), .....]$ ,  $A = (a_1, a_2, a_3 ...)$ ,  $E = (e_1, e_2, e_3)$

Ainsi  $E$  représente l'erreur entre le résultat calculé et réel, son objectif est de le réduire mais le réduire au minimum reviendrait à trouver une concordance parfaite. Dans la pratique, cela est presque impossible. De même, n'ayant que  $X$  et  $Y$ , afin de trouver  $A$ , on l'approchera au mieux et on notera  $\hat{A}$  son estimateur. Afin d'estimer ce dernier, on utilisera la méthode des moindres carrés, c'est-à-dire :  $\hat{A} = (X^tX)^{-1}X^tY$  où  $X^t$  représente la transposée de  $X$ .

## Classification logistique

---

Le principe de cette classification est relativement similaire à une classification des moindres carrés au détail près que celle-ci se penche tout particulièrement sur le principe de classification entre deux classes bien séparées,  $Y$  doit alors être compris dans  $\{0;1\}$ . Afin de réaliser ce tour de force, cette dernière applique une fonction sigmoïde sur son résultat obtenu. La fonction appliquée à la particularité de retourner une valeur dans  $]0;1[$  ce qui pourrait être compris comme la possibilité de "n'appartient pas" ou d' "appartient". On pourrait définir naturellement le seuil de changement de classe à 0.5 en revanche dans certains cas il est de meilleur augure d'augmenter ou diminuer ce seuil. Par exemple, pour des patients atteints de cancers on pourrait avoir tendance à vouloir éviter l'usage de chimiothérapie (on suppose ici que l'utilisation est décidée via notre fonction) et ainsi en positionnant le seuil à 0.8 alors un patient ne se dirigera vers une chimio uniquement s'il a plus de 80% de risque d'avoir un cancer nécessitant un tel traitement. On utilisera ici une estimation à l'aide du maximum de vraisemblance.

## Classification multi-classe

---

Alors qu'avec la classification des moindres carrés, le résultat pouvait être continu, qu'avec la classification logistique, il se retrouvait entre 2 classes, la classification multi-classe permet de classer les individus entre différentes classes distinctes. Bien que possibles à l'aide de manipulations avec les classifications des moindres carrés et logistique, la classification multi-classe peut utiliser d'autres méthodes de classification telle que le K-nn (K nearest neighbour) ou le SVM (Support Vector Machine).

