

Contexte

La maîtrise de l'attrition des clients, c'est-à-dire du taux de départ (appelé "churn" par les Anglo-saxons) est devenue fondamentale pour la survie des fournisseurs d'accès à internet !

La volatilité des clients oblige les entreprises à accentuer les actions de fidélisation. En effet, c'est le meilleur investissement marketing possible, car le coût pour fidéliser un client existant est bien moindre que pour en acquérir un nouveau. Toutefois, pour être réellement efficaces, ces actions doivent être finement ciblées.

Description

Les candidats ont accès à un jeu de données comprenant 15 attributs et 5986 instances.

La description des variables est la suivante :

ID : identification de la ligne (ce n'est donc pas une variable ! Elle se trouve en première colonne du fichier)

La variable à prédire ($Y = \text{« Désabonnement »}$) est une variable binaire qui correspond à un désabonnement (Oui, Non) ; elle se trouve en dernière colonne du fichier

Les variables explicatives ($X = \dots$) sont les suivantes :

Informations démographiques sur les clients

Genre : le sexe du client (Homme, Femme)

Senior : indique si le client a 65 ans ou plus : Oui, Non

Enfant : indique si le client a des enfants à charge.

Services auxquels chaque client s'est inscrit

Ancienneté : depuis combien de temps (en mois) il est client

Lignes multiples : si le client a plusieurs lignes ou non

Services internet : type de fourniture d'accès Internet du client (DSL, Fibre optique, Non).

Autre service : si le client dispose ou non d'une sécurité en ligne.

Partenaire : si le client a un partenaire.

Informations sur le compte client

Type de contrat,

Mode de paiement,

Facturation sans papier,

Frais mensuels,

Frais totaux.