
Projet 4 de Scoring par Machine Learning

Cours : Big Data et Marketing Prédictif

Chargé de cours : Aboubacar BAGAYOGO et Boubacar TRAORE

Date : Mai, 2023

Thème : Sport

Le projet est une application d'un modèle de Scoring dans le domaine du sport pouvant aider à évaluer les revenus d'un sportif et la rentabilité d'un club.

Contexte

Les basketteurs les plus connus ont des revenus (salaires) faramineux et très variables durant une période incertaine. Plus cette durée augmente pour ces sportifs, plus leurs revenus acquis augmentent. Par conséquent, une longue carrière procure à ces derniers des revenus importants.

En outre, une longue carrière procure également aux clubs des revenus importants à travers la vente des maillots, des tickets de match, etc.

Ainsi, les basketteurs ont été divisés en deux catégories selon leurs longévités de carrière. La première catégorie concerne les basketteurs ayant au moins cinq ans de carrière, et la seconde ceux ayant moins de cinq ans.

Le but du projet est de construire un modèle de Scoring qui permet d'expliquer et de prédire la longévité de carrière des basketteurs en fonction de leurs performances athlétiques sur le terrain.

Variable à expliquer

La variable à prédire Y est une variable qualitative binaire. Elle est définie par :

$$Y = \begin{cases} 1 & = \text{Oui, i.e., le basketteur a au moins 5 ans de carrière.} \\ 0 & = \text{Non, i.e., le basketteur a moins de 5 ans de carrière.} \end{cases}$$

Base de données

La base de données du projet se nomme "DATABASE_PROJECT_4_VF.xlsx". et contient Fourth Database Project» et contient 19 variables explicatives sur plusieurs basketteurs (Pour plus de détails

sur les variables explicatives, cliquer ici sur ce [lien 1](#) et le [lien 2](#)).

Les variables descriptives (explicatives) d'un basketteur se trouve sur la page suivante.

Variables	Description
X1	Pourcentage de 3 Points.
X2	Minutes joués.
X3	Points par match.
X4	Tentative de lancers francs.
X5	Tirs réalisés.
X6	Pourcentage de tirs réalisés.
X7	3 Points réalisés.
X8	Tentative de 3 Points.
X9	Match joué.
X10	Tentatives de tirs.
X11	Lancers francs réalisés.
X12	Pourcentage de Lancers francs.
X13	Interceptions.
X14	Rebonds.
X15	Rebonds offensif.
X16	Rebonds défensif.
X17	Passes décisives1.
X18	La perte de ballon.
X19	Les blocks (contres).

NB : La variable Id désigne l'identifiant d'un basketteur.