

Projet IAS – Rapport d'étape

Groupe :

Ilyes Tebourski
Julien Maille—Paez
Maxime Parizot
Mohamed Hamdouni

Données : <https://www.kaggle.com/datasets/drgilermo/nba-players-stats-20142015>

Pour ce projet, nous avons décidé de choisir un dataset représentant les statistiques et les informations personnelles des joueurs NBA durant la saison 2014-2015 (source : <https://www.basketball-reference.com/teams/>)

L'objectif de ce projet est de prédire la position à laquelle le joueur évolue (Pivot, Ailier Fort, Ailier, Arrière, Meneur) en fonction des autres paramètres.

Ce dataset contient 34 paramètres pour 490 joueurs. Ces paramètres ne sont à priori pas tous utiles pour cette analyse, comme par exemple la date de naissance, l'âge ou le lieu de naissance (bien que certains pays sont connus pour mieux former certaines positions que d'autres).

Certains paramètres vont être simple à analyser et exploiter afin de déterminer la position du joueur, comme la taille ou le poids. En effet, un joueur grand et lourd sera très souvent ailier fort ou pivot, tandis qu'un joueur plus léger sera plus souvent meneur.

D'autres paramètres n'ont pas de valeur individuellement et devront être analysés ensembles : comme les points, les rebonds, les passes, les contres et les interceptions.

Un autre exemple : le ratio entre le nombre de tir à 2 points effectués et le nombre de tir à 3 points, car un meneur ou un arrière tente généralement beaucoup de tirs à longue distance, alors qu'un pivot tire très peu de loin.

Un joueur titulaire et un joueur de banc auront également des lignes statistiques très différentes donc ce sera une difficulté à prendre en compte dans l'analyse (certains joueurs de banc n'ont pour consigne que de tenter des tirs à 3 points, peu importe leur poste).

Enfin, certains postes sont plus dur à différencier : en regardant la ligne statistique d'un joueur il est facile de voir si ce joueur est plutôt un meneur ou un pivot, mais un ailier fort et un pivot auront souvent des statistiques très similaires et il est souvent très dur de différencier ces deux postes.

Ainsi, si cette analyse s'avère être trop dure, on pourra peut-être dans un premier temps séparer les joueurs en 2 classes : poste intérieur (ailier, pivot, ailier fort) et extérieur (arrière, meneur) puis affiner ensuite.

Certains joueurs ont également certaines informations manquantes donc il faudra les compléter ou retirer des joueurs du dataset.

