

Trabajo de Final de Grado

NBAStatR

Machine Learning para la predicción de
eventos en la NBA

Albert Villar Ortiz

Universidad Autónoma de Barcelona

ÍNDICE GENERAL

Introducción	3
Dataset partidos	3
Bibliografía	4

Introducción

Este documento tiene la finalidad de especificar el uso de la librería **NBAStatR** para la obtención de los datos estadísticos. Dicha librería basada en R fue creada por Alex Bresler, y compartida de forma pública mediante *GitHub* [1].

Dicho paquete se encuentra actualmente en la versión 0.1.12020301, a partir de la cual tenemos a nuestra disposición un total de 77 funciones que nos proporcionaran datos sobre la NBA mediante el acceso a diferentes portales oficiales como puede ser *NBA Stats* o *Basketball Reference*. Así pues, este documento resume el análisis realizado para determinar cuales son las funcionalidades que nos interesan en todo momento y, en algunos casos, qué datos decidimos obtener.

Dataset partidos

El primer *dataset* que decidimos crear fue el relacionado con datos de los partidos jugados cada día durante 5 años consecutivos. Para ello, el análisis determinó que debíamos utilizar la funcionalidad de *box_scores()* [2], pues no proporcionaría un amplio abanico de datos estadísticos de cada uno de los partidos jugados; des de el resultado, como los porcentajes de acierto en cada uno de los tipos de tiro, entre otros tipo de estadísticas basadas en cada uno de los equipos en el partido.

En el momento del desarrollo, la función generaba un seguido de errores internos que no nos permitían obtener los datos esperados. Es por ello por lo que, finalmente, se decidió recopilar los datos de otras funciones algo menos extensas, con el objetivo de recopilar lo básico y mínimo necesario para poder predecir correctamente.

En este segundo análisis se decidió utilizar la función *game_logs()* [3], así como la función *win_probability()* [4]. Mediante estas dos vías podíamos obtener los resultados básicos de un partido; datos como la victoria, si los equipos vienen de *back to back*, la probabilidad de victoria de cada uno de los equipos y datos individuales de los jugadores.

El uso de estas funciones nos provocaba realizar un tratamiento posterior de los datos para ordenarlos y representarlos tal y como deseamos.

Bibliografía

- [1] A. Bresler, «GitHub,» [En línea]. Available: <https://github.com/abresler/nbastatR>. [Último acceso: 11 03 2019].
- [2] A. Bresler, «asbcllc,» [En línea]. Available: http://asbcllc.com/nbastatR/reference/box_scores.html. [Último acceso: 12 03 2019].
- [3] A. Bresler, «asbcllc,» [En línea]. Available: http://asbcllc.com/nbastatR/reference/game_logs.html. [Último acceso: 15 03 2019].
- [4] A. Bresler, «asbcllc,» [En línea]. Available: http://asbcllc.com/nbastatR/reference/win_probability.html. [Último acceso: 16 03 2019].