



Predicción de resultados de baloncesto: ML y simulación por cadenas de Markov homogéneas.

Abstract

El objetivo de este proyecto es el desarrollo de métodos estocásticos para la predicción (*forecasting*) de partidos de baloncesto. Para ello, por un lado, se modela el problema como uno de clasificación y se implementan técnicas de regresión logística y de árboles de decisión. Por el otro, se modelan partidos de basquetbol como cadenas de Markov homogéneas que transicionan posesión a posesión. El entrenamiento, en el caso de ML, se realiza de la manera tradicional. En el Markoviano, en tanto, se utilizan datos de posesión a posesión para estimar empíricamente probabilidades de transición. A la hora de predecir, para el primer modelo, consideramos conocido el resultado hasta el primer cuarto, y a partir de eso, predecimos el ganador. Para el caso Markoviano, utilizamos la data *play-by-play* del primer cuarto, y la combinamos convexamente con las matriz de transición estimada previamente.

Habiendo optimizado los parámetros de los modelos, se comparan estos usando data jugada-a-jugada de la temporada 2009/10 de la NBA. Se observa, en general, que la predicción por ML es menos costosa y más certera que la por Markov, que resulta inconsistente. Se conjetura que esto último se debe a la naturaleza del básquetbol como deporte y por el alcance de especificidad de este proyecto, al no particularizar por equipos.

Bibliografía

1. Shirley, K. (2007). A Markov model for basketball. Presentación en el *New England Symposium for Statistics in Sports*, Boston, MA, Septiembre 2007.
http://www.nesis.org/nessis07/Kenny_Shirley.pdf.
2. Štrumbelj, E., amp; Vračar, P. (2011, May 28). Simulating a basketball match with a homogeneous Markov model and forecasting the outcome. *International Journal of Forecasting*. <https://www.sciencedirect.com/science/article/abs/pii/S0169207011000458>.