

TP 3 - Metnumball

Ciro Román Zar Abad, Lucas Rafael Romero, Patricio López Valiente

*Departamento de Computación
UBA
Buenos Aires, Argentina*

Abstract

La liga de básquet de Estados Unidos (NBA) cuenta con una extensa base de datos de estadísticas ¹. Estos indicadores son utilizados frecuentemente para analizar el rendimiento de los equipos y el desempeño individual de los jugadores.

En este informe se presentan distintas métricas para evaluar dichas estadísticas, basándose tanto en opiniones de expertos en la materia, como también en métodos numéricos.

Keywords: NBA, APBRmetrics, Cuadrados mínimos, PER, Four Factors.

1 Introducción

El uso de estadísticas en los deportes es fundamental para medir el rendimiento de los equipos así como también el desempeño individual de los jugadores. Dichos indicadores se usan con el fin de comparar pero también para saber qué aspectos del juego son fortalezas o debilidades con el fin de aprovecharlas o mejorar.

¹ <http://stats.nba.com/>

El objetivo de esta investigación es analizar la extensa base de datos de estadísticas, tanto ofensivas como defensivas de temporadas de la NBA, con el fin de predecir el winrate (cantidad de partidos ganados sobre partidos jugados) de un equipo determinado.

1.1 Métodos utilizados

1.1.1 Cuadrados Mínimos Lineales

Este método fue utilizado para la resolución de los sistemas planteados, en el cual buscamos minimizar la relación entre nuestras métricas con respecto al winrate. Para ellos se utilizó 'CML'² ya implementado en MATLAB.

1.1.2 Validación por Cross Validation

Uno de los métodos que utilizamos para evaluar la efectividad de una métrica dada, es el de K-fold cross validation³, con parámetro $K = 10$. Esto nos deja con un set de test de 85 equipos, que al estar equitativamente distribuidos a lo largo de las temporadas, y repetirse 10 veces, consideramos suficiente.

1.1.3 Validación por Temporadas

Para tener en cuenta el cambio de las dinámicas de juego a lo largo de las temporadas, se desarrolló esta validación.

Para validar por temporadas, elegimos realizar la etapa de training sobre todos los equipos de 5 temporadas consecutivas, y la de testing sobre todos los equipos de la 6ta concurrentemente.

1.1.4 Error Cuadrático Medio

A la hora de realizar la validación utilizamos el método de Error Cuadrático Medio⁴ (MSE) en cada caso de test para comparar el resultado con los verdaderos winrates. El resultado expuesto es el promedio de los MSE de todos los tests.

² <http://www.utdallas.edu/~herve/Abdi-LeastSquares-pretty.pdf>

³ <https://www.cs.cmu.edu/~schneide/tut5/node42.html>

⁴ <https://www.vernier.com/til/1014/>

2 Métricas de jugadores

2.0.1 Desarrollo

Para el caso de las métricas de jugadores tomamos como base la presentada por Wayne Winston en su blog⁵. En el asegura que la métrica presentada aproxima PER con un 99% de precisión. Las estadísticas que se observan para generar los resultados son:

PTS, TRB, AST, STL, TOV, eFG%, FT%, BLK, PF.

Tanto la métrica original como las que desarrollamos, ponderan los valores por la cantidad de minutos jugados. La métrica original consideraba tiros de campo y tiros libres errados, para este caso consideramos que el porcentaje de estos valores analiza la misma información.

La primera métrica que consideramos fue una lineal, en la que no modificamos ninguno de los parámetros.

Con el objetivo de enfatizar ciertos aspectos del juego, aplicamos funciones a las estadísticas de nuestra métrica lineal buscando obtener mejores resultados. Las siguientes combinaciones fueron las más efectivas⁶:

Métrica 1: $RB^2, \sqrt{TOV}, FT\%^2, \sqrt{\log(PF)}$. Toma más importancia a los rebotes y a la efectividad de tiros libres y le resta a faltas personales y turnovers.

Métrica 2: $\log(TOV)$. Saca importancia a los turnovers.

Métrica 3: $STL^2, BLK^2, \log(PF)$. Le pone más importancia a estadísticas defensivas y le resta prioridad a las faltas personales porque pueden ser usadas de manera táctica

Estos experimentos fueron corridos utilizando toda la base de datos disponible para validar.

2.0.2 Resultados

	Métrica 1	Métrica 2	Métrica 3	Lineal
10-Fold Cross-Validation	0.00982	0.01000	0.01041	0.01180
Validación por temporadas	0.01041	0.01099	0.01110	0.01366

Table 1
MSE por métrica para ambas validaciones

⁵ <http://waynewinston.com/wordpress/?p=1326>

⁶ Si una estadística no es mencionada, se asume que se mantiene lineal

En la Tabla 1 figuran los promedios del MSE de todas las temporadas para cada métrica, obtenidas a través de los distintos métodos de validación. Bajo este criterio, la métrica 1 fue la que obtuvo mejores resultados.

En la primer figura presentamos el MSE entre el winrate predicho y el real, para cada métrica desarrollada. Lo primero que podemos observar es que todas las métricas presentan un error similar a lo largo de las distintas temporadas, exceptuando la temporada 1998-1999, la cual presenta anomalías de las que hablaremos ms adelante.

A continuación realizamos una comparación entre las predicciones de PER y la métrica 1 de cada equipo en la temporada 2016. Aquí podemos ver que nuestra métrica, al igual que PER concentra la mayor parte de los equipos alrededor de la media.

Por otro lado, si bien predice de mejor forma para ciertos equipos como CLE, OKC y BRK, como contrapunto, se ve que con los aspectos tenidos en cuenta genera predicciones desacertadas con respecto a PER en muchos otros.

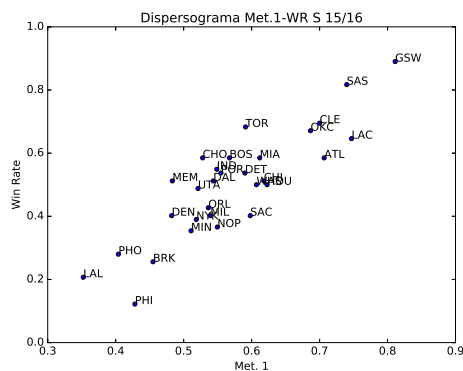
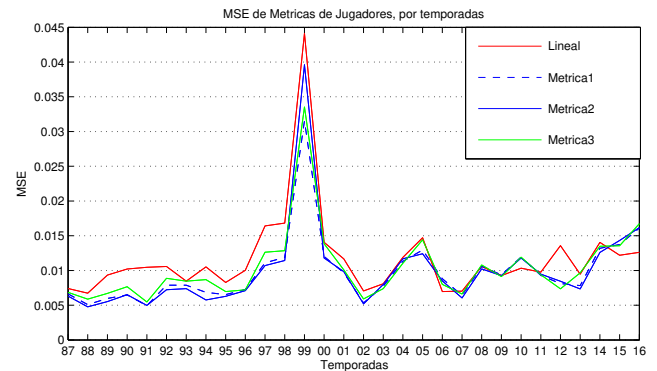


Fig. 1. Dispersograma de la métrica 1

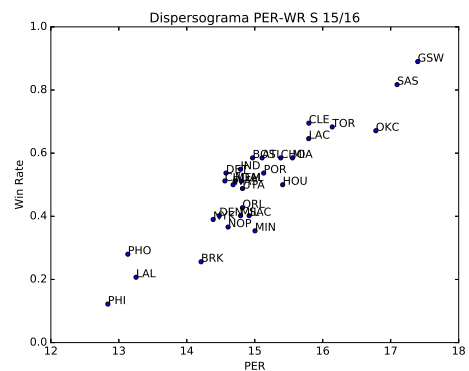


Fig. 2. Dispersograma de PER

3 Métricas de Equipos

3.0.1 Desarrollo

En esta métrica buscamos utilizar los aspectos más relevantes del juego en su forma mas simple, es decir aquellos que permiten a un equipo anotar y obtener/perder posesión del balón ⁷ ⁸, ya que a mayor cantidad de posesiones, mayor es la posibilidad de anotar. Además se buscó generar una métrica para evaluar el basketball moderno, por lo que se evaluaron sólo las últimas 10 temporadas en las cuales se juega más en equipo y ha disminuido la importancia de los pivotes en favor de los tiradores. Luego a ello se le calculaba el MSE para cada métrica.

Las estadísticas usadas fueron las siguientes: $3P\%$, $2P\%$, $FT\%$, TRB , STL , TOV .⁹

Elegimos estas estadísticas tomando como base Four Factors, la cual fue utilizada durante un largo tiempo, ya que releva de forma efectiva las anotaciones y los aspectos relevantes a la posesión de balón.

A la métrica planteada le aplicamos diversas funciones según el peso específico que se le quería otorgar a los stats utilizados. Luego de estas pruebas se tomaron las de mejor resultado.

Las métricas obtenidas fueron:

Métrica Lineal: Los exponentes usados fueron los triviales, osea los unarios.

Métrica 1: $3P\%^2$, $2P\%^2$, $FT\%^2$, TRB , STL , \sqrt{TOV} , donde se enfoco en nivelar el peso de los puntos sobre las otras estadísticas, al quitarles peso a estos.

Métrica 2: $\sqrt{3P\%}$, $2P\%$, $FT\%$, TRB^2 , STL^2 , TOV , aquí se priorizaron aquellas estadísticas que permiten obtener el balón. Además se siguió considerando los puntos, aunque en menor medida. El objetivo de esta métrica fue resaltar los aspectos defensivos ya que esta es una de las principales críticas a la métrica Four Factors.

⁷ <http://www.basketball-reference.com/about/factors.html>

⁸ <http://hangtime.blogs.nba.com/2013/02/15/the-new-nba-comstats-advanced-stats-all-start-with-pace-and-efficiency/>

⁹ <http://www.basketball-reference.com/about/glossary.html>

3.0.2 Resultados

	Lineal	Métrica 1	Métrica 2
10-Fold Cross-Validation	0.0195	0.0192	0.0191
Validación por temporadas	0.0222	0.0250	0.0245

Table 2

MSE por métrica para ambas validaciones

En la Tabla 2 figuran los promedios de los MSE de todas las temporadas para cada métrica obtenidas a través de los distintos métodos de validación. Como podemos observar, la métrica 2 fué la que obtuvo mejores resultados.

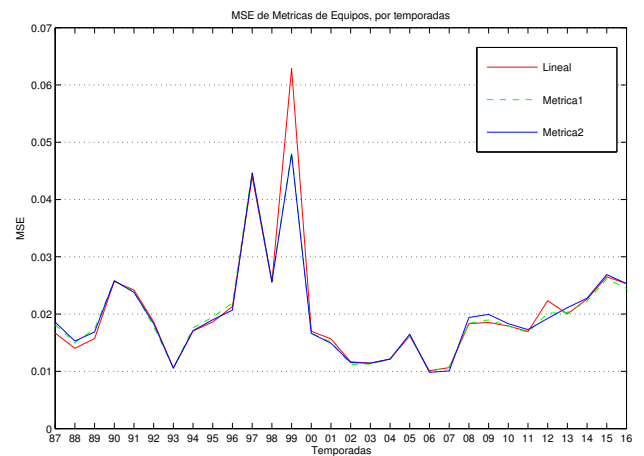
Hay que considerar que en este caso, Cross-Validation es más confiable que el método de validación por temporadas, porque promedia el resultado de 10 pruebas con una base de test de 30 equipos aleatoriamente distribuidos, contra solo 5 pruebas con base de test del mismo tamaño.

Si bien las métricas varían bastante en cuanto a qué estadísticas le ponen o quitan énfasis, los resultados obtenidos en función del 'MSE', no varían en tal magnitud:

En la primer figura presentamos el MSE entre el winrate predicho y el real, para cada métrica desarrollada. En el caso de éstas métricas, la función del MSE oscila más que en el experimento de jugadores, de hecho presentan más cantidad de picos con mayor MSE. Esto es algo esperable, ya que la métrica esta basada en la de Four Factors, la cual es inferior a PER.

Además, se vuelve a repetir la temporada 1998-1999 como la que presenta un notable pico.

A continuación realizamos una comparacin entre las predicciones de Four Factors y la métrica 2 de cada equipo en la temporada 2016. Elegimos ésta métrica porque, no solo presenta el menor MSE en función de Cross-Validation, sino que también porque es la que pondera los aspectos defensivos, que se encuentran relegados en Four Factors.



Analizando la 3, que presenta las predicciones de la métrica 2 en las temporadas 15/16, observamos en primera medida que los equipos con winrate predicho significativamente mayor al real como BOS, CHO y MEM, tienen similitudes en el hecho de que pierden poco la bola. Además en el caso de MEM presenta buenas posiciones en anotaciones y robos, mientras que en el caso de CHO presenta mejores posiciones en anotaciones y si bien no destaca en robos lo hace en rebotes. En el caso de BOS, el mejor rankeado, destaca entre los mejores 6 en 4 de nuestras estadísticas y los aspectos en los que no destaca, como 3P%, no son un factor de peso en nuestra métrica.

Para quienes se ven perjudicados por esta métrica, es decir con winrate predicho significativamente menor al real, como por ejemplo GSW y OKC, se da el caso que están entre los líderes en pérdidas de la liga. Además para OKC mientras que destaca en robos y FT% sólo alcanza para que promedien nuestro estadístico. En el caso de GSW están poco por encima del promedio en todos los demás estadísticos, excepto en dos y tres puntos, que como ya se mencionó no aportan gran peso.

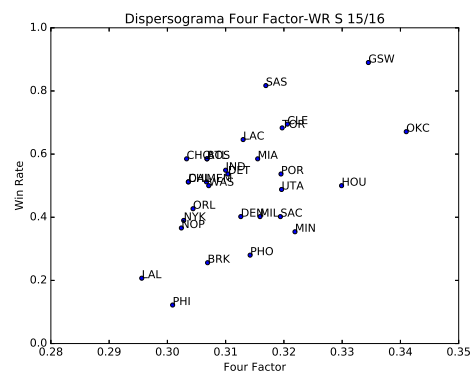
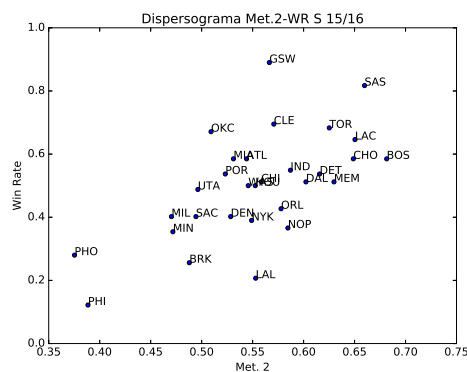


Fig. 3. Dispersograma de la métrica 2 Fig. 4. Dispersograma de Four Factors

Como se observa en las 3 y 4, se ve que el priorizar los aspectos defensivos y la posesión no soluciona la deficiencias originales de Four Factor. Se puede observar que mientras nuestra métrica beneficia a los equipos defensivos en demérito de los ofensivos, inversamente a lo que ocurre en Four Factors. Esto genera que en muchos casos nuestra métrica acierte en los equipos en los que Four Factors comete un error alto y viceversa.

Con otros equipos, que presentan una forma de juego notablemente ofensiva o defensiva, como OKC, CHO, PHO, ambas métricas fallan mucho en la predicción pero de forma inversa, es decir donde una lo beneficia de más, la otra lo perjudica.

4 Basket is like War

4.0.1 Desarrollo

La motivación de este experimento es la de comparar como afectan a las predicciones considerar puramente aspectos ofensivos vs. aspectos defensivos. Dado que observamos que la mayoría de las métricas conocidas son mayormente ofensivas. Para esto confeccionamos dos métricas de jugadores, las cuales al igual que en el experimento anterior de jugadores, ponderamos las estadísticas usadas por los minutos jugados. Para ambas métricas tomamos todas las estadísticas disponibles básicas relevantes a los dos aspectos antes mencionados, y además tomamos los winshares ya que son una estadística reconocida para medir los aportes tanto ofensivos como defensivos que realizan los jugadores.

Las siguientes métricas fueron las elegidas:

Métrica defensiva: *DRB, STL, BLK, PF, DWS*

Métrica ofensiva: *eFG%, FT%, ORB, TOV, PTS, OWS*

Ambas métricas fueron desarrolladas de manera lineal, para no afectar en la comparación realizando mejores optimizaciones sobre ninguna.

Estos experimentos fueron corridos utilizando toda la base de datos disponible para validar.

4.0.2 Resultados

	Métrica Def.	Métrica Off.
10-Fold CV	0.0108	0.0093
Validación por Temp.	0.0118	0.0101

Table 3
MSE por métrica para ambas validaciones

En la Tabla 3 figuran los promedios de los MSE de todas las temporadas para cada métrica obtenidas a través de los distintos métodos de validación. Como podemos observar, la métrica ofensiva fue la que obtuvo mejores resultados.

En la primer figura presentamos el MSE entre el winrate predicho y el real, para cada métrica desarrollada. Lo primero que podemos observar es que los gráficos son principalmente homogéneos y en la mayoría de las temporadas la métrica ofensiva presenta menor error pero con una diferencia muy leve. Sin embargo, al igual que en los experimentos anteriores, en el año 1999 se observa un pico muy significativo.

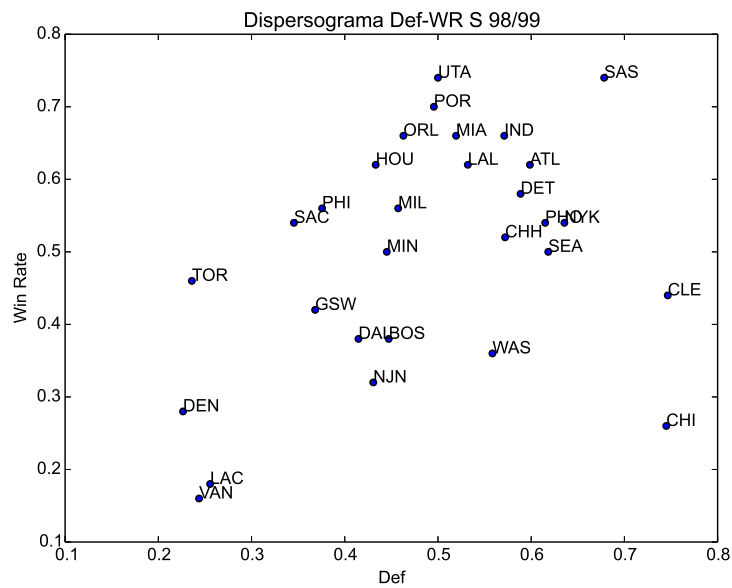
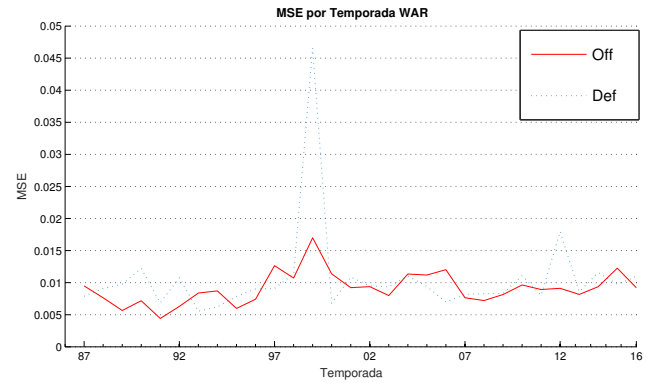


Fig. 5. Dispersograma de la métrica defensiva en temporada '99

En la figura 5 analizamos en detalle la predicción de la métrica defensiva de la temporada 98/99 ya que es la que presenta un MSE anormal. En esta se puede concluir que los si bien al ser una métrica puramente defensiva sobrevalora los equipos que destacan en este aspecto, esta provee información sustanciosa sobre lo ocurrido en este año.

Esta temporada es particular y merece un análisis detallado, ya que coincide con un cambio importante que se produjo en la dinámica de la liga. Cobraron más importancia la estrategia defensiva y el juego en equipo, tomando

a SAS, equipo que saldría campeón de la temporada, como principal representante de este nuevo paradigma.

También en esta figura se nota que los equipos con buena defensa pero con ofensivas defectuosas como es el caso de CHI y CLE presentan los mayores errores. Esto también deja en evidencia que a pesar de lo mencionado anteriormente, no alcanza solo con esto. Puntualmente en el ejemplo de CHI, el que presentó un mayor error por amplio margen, estamos analizando la temporada posterior al retiro de Michael Jordan y Dennis Rodman, dos de sus mayores estrellas.

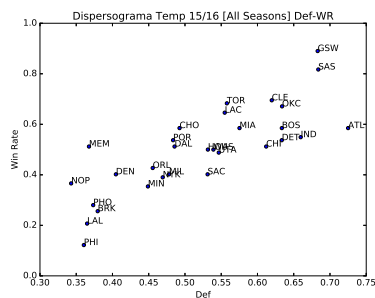


Fig. 6. Dispersograma de la métrica de-
fensiva.

Esto presumimos se debe a que en la liga si bien desde hace tiempo la defensa es un aspecto muy relevante, todavía predominan los equipos ofensivos, por lo que quienes poseen un juego de golpe por golpe en la métrica defensiva son perjudicados, aunque finalmente pueda ser una estrategia efectiva.