

El Síndrome de los Datos Ricos e Información Pobre en Deportes de Competición

-Perspectiva desde las Ciencias Computacionales y Ciencia de Datos-

Nelson Fernández ^{1, 2, 3} Martín Zumaya ³ Carlos Gershenson ^{3, 4} .

¹ Laboratorio de Investigaciones en Hidroinformática, Universidad de Pamplona, Colombia.

² Grupo de Investigación en Ecología y Biogeografía, Universidad de Pamplona, Colombia.

³ Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México

⁴Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México

nfernandez@unipamplona.edu.co

martinzh@icf.unam.mx

cgg@unam.mx

Resumen

La gran capacidad existente de capturar datos, conlleva la subsecuente responsabilidad de producir información confiable, verificable y auditable para la toma de decisiones. En el fútbol, la existencia de compañías y plataformas con capacidad de medir un sinnúmero de variables de desempeño, ha generado una explosión de datos de difícil interpretación. En este sentido, las dificultades relativas al análisis y visualización de estos datos, ha derivado en el “Síndrome de los datos ricos e información pobre”. En este contexto, esta plática se centra en evaluar las lecciones aprendidas y las perspectivas futuras en el manejo de datos en el fútbol, desde una perspectiva computacional y de ciencia de datos. Nuestro enfoque metodológico, parte de la evaluación de los formatos en que se produce los datos y los tipos de reportes generados para distintos tipos de usuarios. Planteamos una forma adecuada de manejar e interpretar múltiples variables con soporte en técnicas de aprendizaje automático, con técnicas de ordenación y clasificación para discriminar los factores y variables que tienen mayor contribución en el juego. Finalmente, brindamos información sobre perspectivas novedosas para el modelado de los eventos espacio-temporales, que tienen lugar en los partidos, como la aplicación desde la ciencia de redes, redes de latencia y modelos de gravitación para el modelado. Nuestra perspectiva computacional y de ciencia de datos brinda la posibilidad de mejores visualizaciones, con el propósito de simplificar el gran número de dimensiones y categorías que se inspeccionan en el fútbol. De esta forma, nos enfocamos en las interacciones relevantes del juego, que darían soporte a una mejor toma de decisiones por parte de distintos tipos de usuarios, como jugadores, entrenadores y directivos.

Caja 1 Resultados de los Análisis Multivariados de Componentes Principales y de Análisis Jerárquico al Caso de la Liga Mexicana

