

Fundamentos de ciencia de datos con R

Palabras clave: libro, modelos, limpieza de datos, reportes, casos de usos

Abstract

El siglo XXI está siendo testigo de cambios vertiginosos en el contexto social y tecnológico, entre otros. Los tiempos han cambiado, la sociedad se ha globalizado y “exige” respuestas inmediatas a problemas muy complejos.

Vivimos en el mundo de la **información**, de los **datos**, o mejor, de las **bases de datos masivas**, y los ciudadanos y, sobre todo, las empresas y los gobiernos, dirigen su mirada hacia el mundo científico para que les ayude a “**oír las historias**” que cuentan esos datos acerca de la realidad de la que han sido extraídos.

Y dado su enorme volumen y sofisticación (en el nuevo renacimiento las imágenes y los textos, por ejemplo, también son datos), exigen algoritmos de nueva generación en el campo del *machine learning*, o incluso del *deep learning*, para “oír las historias” que cuentan. No parecen mirar al “antiguo” investigador científico, sino al “nuevo” **científico de datos**.

Es por ello imprescindible disponer de un manual con las características del que aquí se presenta: rigor, completitud, amplitud temática y variedad de perspectivas (está elaborado por más de cuarenta autores), y todo ello implementado de principio a fin con el **software estadístico R**.

En esta charla presentaremos el libro que está escrito en español y se puede acceder de forma libre y gratuita en la web: <https://cdr-book.github.io/index.html>