

En XX SIMMAC

¿Cuál es la relación entre Big Data, minería de datos y Estadística?

Las redes sociales son la fuente más grande de datos en el mundo

25 FEB 2016 Ciencia y Tecnología



Las cuatro características de la *Big Data* (fuente: IBM, ilustración tomada de Internet).

Datos masivos, macrodatos o grandes volúmenes de datos. Así se puede definir el significado de *Big Data*, un término que ha tomado mucho auge en la actualidad debido al crecimiento exponencial de datos.

Pero, ¿es una moda o un tema que surgió en la academia? A criterio del Dr. Oldemar Rodríguez Rojas, profesor e investigador de la [Escuela de Matemática](#) de la Universidad de Costa Rica (UCR) y especialista en el tema, el concepto *Big Data* fue popularizado por los grandes motores de búsqueda en Internet, debido a los **enormes volúmenes de datos que se manejan en la red**.

“Hay dos términos que caminan muy de la mano: *Big Data* y *Cloud Computing* (datos en la nube)”, expresó el matemático. Esto ha llevado a que las computadoras se estén convirtiendo en simples terminales porque la información se almacena en espacios virtualizados.

Las redes sociales son la fuente de datos más grande en el mundo. El auge de los dispositivos móviles ha permitido el aumento de la velocidad y el volumen de datos almacenados en las redes, tales como fotografías y videos, explicó el experto.



Las sesiones de la vigésima edición del Simposio Internacional de Métodos Matemáticos Aplicados a las Ciencias se realiza en la Facultad de Educación, en la Sede Rodrigo Facio de la UCR (foto Laura Rodríguez).

Rodríguez ofreció la conferencia de apertura del XX Simposio Internacional de Métodos Matemáticos Aplicados a las Ciencias (**Simmac**) sobre la **relación entre *Big Data*, minería de datos y Estadística**.

El Simmac es organizado por el Centro de Investigaciones en Matemática Pura y Aplicada ([Cimpa](#)) y su vigésima edición se realiza del 23 al 26 de febrero en la Sede Rodrigo Facio de la UCR, con la participación de más de **100 matemáticos de 20 países**.

El Simposio está organizado en esta oportunidad en tres conferencias, nueve minicursos prácticos y 93 ponencias. Entre los **temas** que se abordan están: análisis de datos, Estadística multivariada, clasificación, minería de datos, optimización, investigación de operaciones, probabilidad, procesos estocásticos, Matemática financiera, análisis numérico, modelación, biomatemáticas y ecuaciones diferenciales.

Minería de datos

El otro concepto relacionado con *Big Data* es minería de datos, que surgió en el año 2000 asociado a la inteligencia de negocios y que ha cobrado importancia en la actualidad.



El Dr. Oldemar Rodríguez, profesor e investigador de la Escuela de Matemática de la UCR, ofreció la conferencia inaugural del XX Simmac sobre la relación entre *Big Data*, minería de datos y Estadística (foto Laura Rodríguez).

“La minería de datos consiste en la transformación de datos en conocimiento para la toma de decisiones”, lo cual se logra mediante técnicas de análisis simbólico; es decir, “no se trabaja sobre los datos, sino sobre conceptos”, explicó el académico.

A diferencia de la Estadística, en la minería de datos no se trabaja con muestras de la información, sino que se buscan patrones ocultos en los datos mediante métodos algorítmicos y utilizando un sistema automatizado.

De ahí que en la minería de datos confluyen varias ciencias. A criterio de Rodríguez, el desarrollo de estas áreas es tan rápido que ya en este momento la *Data Science* se perfila como una nueva profesión, en la que los profesionales deben prepararse en una serie de áreas, entre estas Estadística, Matemática, Computación, Ingeniería de datos, Negocios y hasta Hacking.

El Dr. Javier Trejos Zelaya se refiere al Simposio Internacional de Métodos Matemáticos Aplicados a las Ciencias (SIMMAC), que se realiza cada dos años

en la UCR.



[Patricia Blanco Picado](#)

Periodista Oficina de Divulgación e Información

patricia.blancopicado@ucr.ac.cr

Etiquetas: [big data](#), [minería de datos](#), [simmac](#), [simposio](#), [matematica](#), [estadística](#), [escuela de matematica](#), [centro de investigaciones en matematica pura y aplicada](#).