



XXXI SEMANA NACIONAL DE INVESTIGACIÓN Y DOCENCIA EN MATEMÁTICAS

Ciencia de Datos

Rodríguez Martínez Eliud Gilberto, Soto Sánchez Oscar René, Juan Carlos Rojas Coffey
Dra. María de Guadalupe Cota Ortiz

Introducción

En nuestro presente la comunicación por medio de internet se ha vuelto algo más que necesaria y por ende nuestra producción de información (datos) ha, y seguirá aumentando de manera que se volverá imprescindible el análisis de estas gigantescas cantidades de información, para eso se requiere de equipo y habilidades, desgraciadamente en un país donde el desarrollo de equipo útil se está extinguiendo, un claro ejemplo de esto es el hecho de pasar del lugar 255 en 2015 a ya ni figurar en la lista de las top 500 supercomputadoras del mundo, y contamos con una inexistente cultura cibernética, por esto hemos decidido investigar con respecto a la ciencia de datos, pues con ella podemos, no solo podemos analizar los datos previamente mencionados, sino que podemos fomentar la cultura de la protección de nuestros datos (parte primordial de la cultura cibernética) y demostrar que nuestro país debería seguir fomentando el desarrollo de equipo para estos científicos.

¿Qué es la ciencia de datos?

La ciencia de los datos se refiere a un área emergente de trabajo relacionado con la recolección, preparación, análisis, visualización, administración y conservación de grandes colecciones de información.

Aunque el nombre de Datos Científicos parece conectar más fuertemente con áreas tales como bases de datos y la informática, incluye diferentes tipos de habilidades – incluyendo las habilidades matemáticas. Para algunos, el término “ciencia de datos” evoca imágenes de estadísticos con batas blancas de laboratorio que miran fijamente parpadear la pantalla del ordenador llenos de números en desplazamiento

Nada más lejos de la verdad. En primer lugar, los estadísticos no llevan batas de laboratorio: esta moda está reservada para los biólogos, médicos, y otros que tienen que mantener su ropa limpia en ambientes asépticos. En segundo lugar, muchos de los datos en el mundo no son numéricos ni estructurados. En este contexto, no estructurado significa que los datos no están dispuestos en filas y columnas ordenadas. Por ejemplo, una página web con fotografías y mensajes cortos entre amigos.

¿Qué es la Big Data?

Big Data es un término que describe el gran volumen de datos, tanto estructurados como no estructurados, que inundan los negocios cada día. Pero no es la cantidad de datos lo que es importante. Lo que importa con el Big Data es lo que las organizaciones hacen con los datos. Big Data se puede analizar para obtener ideas que conduzcan a mejores decisiones y movimientos de negocios estratégicos.

Cuando hablamos de Big Data nos referimos a conjuntos de datos o combinaciones de conjuntos de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles. Aunque el tamaño utilizado para determinar si un conjunto de datos determinado se considera Big Data no está firmemente definido y sigue cambiando con el tiempo, la mayoría de los analistas y profesionales actualmente se refieren a conjuntos de datos que van desde 30-50 terabytes a varios petabytes.

¿Dónde se aplica la ciencia de datos?

La ciencia de datos es uno de los campos más emocionantes que existen en la actualidad. Ahora que la tecnología moderna ha permitido la creación y el almacenamiento de cantidades cada vez mayores de información, el volumen de datos explotó y esto nos ha permitido la aplicación de la ciencia de datos en diversos campos, como:

- Detección de fraudes y riesgos
- Cuidado de la salud
- Búsqueda de internet
- Publicidad dirigida
- Recomendaciones de sitios web
- Reconocimiento de imagen avanzado
- Reconocimiento de voz
- Planificación de rutas de aerolíneas
- Juegos
- Realidad aumentada
- Entre muchos otros.

Ejemplos de aplicaciones

Ahora bien, es probable que la búsqueda de internet es lo primero que se le ocurra cuando piensa en aplicaciones de ciencia de datos. Hay muchos otros motores de búsqueda como Google, Yahoo, Bing, Ask, AOL, etc. todos estos utilizan algoritmos de ciencia de datos para ofrecer el mejor resultado para nuestra consulta de búsqueda en una fracción de segundos.

Así mismo, utilizando ese tipo de algoritmos, podemos hacer que estas búsquedas se hagan de manera imperceptibles y referentes a los gustos o comportamientos del usuario, esto lo podemos implementar en varios tipos de casos, desde los sitios web que nos aparecen en las búsquedas y recomendaciones de otras páginas, pasando por los carteles publicitarios en las mismas, hasta las vallas publicitarias digitales en los aeropuertos. Esta es la razón por la cual los anuncios digitales han tenido mucho más éxito que los anuncios tradicionales. Se pueden orientar según el comportamiento anterior de un usuario.



Referencias

1. <https://www.oracle.com/mx/data-science/what-is-data-science/>
2. <https://www.edureka.co/blog/data-science-applications/>
3. <https://www.powerdata.es/big-data>
4. <http://anabelforte.com/2020/06/14/ciencia-de-datos-que-y-por-que/>