Generación de bases de datos a través de los sistemas de información. Este es el paradigma que gobierna la modernidad. Transformar datos, realizar un cambio evolutivo de la era industrial a la era digital es en ultimas, el resultado esperado de los últimos años.

La llegada de la computación a principios de siglo y junto a mentes prodigios como Turing, fueron las que se encargaron de llevar la era digital al estado actual.

el fenómeno de la revolución 4.0 ha generado una interconexión directa entre el individuo y a la forma de interactuar y consumir. Es pues, como un mantra que se sigue a raja tabla hoy día.

“Te digitalizas o desapareces”

Cada vez más será más común que los datos que generamos, sean utilizados para con nosotros, véase la domótica, las redes sociales, la publicidad enfocada, los servicios hospitalarios como cirugías remotas, etc.

La era digital ha traído consigo un sinfín de nuevas reglas que no pueden ser ignoradas.

El computador personal, laptop, dispositivos móviles, tabletas, junto con todos los procesos de automatización vinieron para quedarse. Grandes mentes como Turing o Steve Jobs fueron los encargados de revolucionar y estandarizar el consumo de datos. Los datos se generan al existir en la sociedad actual y es por eso que un profesional enfocado en el manejo, interpretación e inferencia de datos es clave en la sociedad moderna.

Pero no todo son flores en el jardín. Debido a la automatización, se pierden puestos de trabajos que no volverán. Los carteros, los repartidores o los conductores son un ejemplo claro de esto. Efectivo desaparecerá. Cada vez más hay una migración hacia las divisas digitales, véase BTC, véase PayPal.

¿Solución? Capacitación continua y movilidad laboral. El paradigma de la preparación universitaria para la vida (con un pregrado) y estabilidad de una empresa con la que se jubilará el individuo, es una idea producto de la era industrial que ya no es vigente. El campo laboral valora un individuo en constante preparación.

INTELIGENCIA SE TRADUCE EN TRANSFORMAR DATOS EN INFORMACION. Este es otro mantra con el cual los análisis de datos a través de las técnicas modernas son llevadas a cabo. La materia prima, el petróleo moderno son los datos. Quien posea la información, tiene diamantes en bruto.

El combinar excelentes técnicas de Analitics con un experto debidamente preparado es la mejor herramienta para afrontar la revolución actual.

Los procesos de analítica hoy en día siguen una secuencia que, en su base, en su unidad primordial, en el átomo de todo, lleva siempre la misma idea.

ETL: Extract, Transform and load.

Este proceso es la base fundamental de cualquier análisis de datos en la actualidad. Los datos en si mismo han de ser trabajados, extraídos (extract) de la base de datos donde se encuentre. Sean servidores, bases de datos etc.

Luego de acceder a esta información, ha de filtrarse acorde al problema de interés. La depuración de la base de datos suele ser el proceso mas exhaustivo pues re quiere entre otras cosas, fuerza computacional y un equipo de trabajo preparado para afrontar el problema.

Con esto, se busca preparar los datos para que de ellos se pueda extraer información y dejarlos estructurados para ser almacenador y finalmente cargados en un DataWarehouse. Será el ambiente donde los datos serán cargados para proceder a ser analizados de mano de un experto y los modelos utilizados.

DATA SCIENCE y su relación con la IA:

Estadística según Víctor Leiva está diseñada para manejar datos organizados, cuando hablamos de big data, las reglas son diferentes. La estadística sigue modelos y reglas, hay inferencias, medidas y se apega al marco teórico matemático.

“la ciencia de datos (dará science) de Tukey es el sucesor de la estadística sin distribuciones paramétricas ni inferencia formal” premio nobel 2020.

La evolución de la estadística es la ciencia de datos. Las técnicas de aprendizaje que se utilizan para el tratado de ese oro en polvo denominado datos. Esto, va de la mano con los profesionales en el campo, generalmente llamados científicos de datos.

“científico de datos quien usa métodos para crear información a partir de datos”

La ciencia de datos es un campo interdisciplinario sobre los procesos de la extracción de datos que se manifiestan de diferentes maneras. Está vinculado a la data mining, estadística y análisis predictivo.

Debido a su alta demanda, los profesionales son muy cotizados en la actualidad y representa una de las profesiones más atractivas en la modernidad, pero, en general, todo profesional debe de conocer y manejar los tipos de modelos que a grandes rasgos se puede categorizar de la siguiente manera:

TRCNICAS SUPERVISADAS. Usualmente tienen una variable respuesta o variable de interés. Y esta variable es afectada por otras que explican a esta variable que nos interesa. Regresión lineal o múltiple, por ejemplo.

TECNICAS NO SUPERVISADAS. Cuando las variables son dependientes entre sí, dada una fuerte multicolinealidad por ejemplo y donde todas las variables usadas son explicativas y significativas, se utiliza técnicas no supervisadas. Clustering es el mas usado. Agrupación o segregación de datos.

TECNICAS DE CLASIFICAION: técnicas como arboles de decisión, reglas de asociación, algoritmos genéricos, Redes neuronales por mencionar algunas, son las diversas técnicas que un profesional en el área ha de entender y manejar.

Sin embargo, la inteligencia artificial (IA) es el uso de datos par aun proceso de aprendizaje de la máquina. Es usar la inteligencia (datos transformados) para su correcta extracción, realizando los procesos ya mencionados, basados en el poder de las computadoras.

El objetivo general y de manera simple, es que la maquina aprenda para que estas gestionen enormes volúmenes de datos para sacar provecho de los datos y así acceder a los metadatos que en ultimas, es en general una de las grandes ventajas comparativas.

La diferencia entre un dato y un metadato puede verse en el siguiente ejemplo:

Si usted se dirige a una tienda y compra un dispositivo con una tarjeta de crédito, esto genera información (un dato) acerca de la compra.

Pero el dónde se realizó la compra, la hora, el lugar, el tipo de dispositivo, el dinero gastado, la cantidad de cuotas, la selección del lugar de compra y el dispositivo escogido de uno vs otros, son, por ejemplo, metadatos que se pueden extraer gracias a un único dato. La compra realizada.

Machine learning vs Deep learning

Esencialmente tiene que ver con el uso de redes neuronales y de big data, el Ml no es suficiente cuando los datos son demasiados, necesitamos Deep learning que se especializa en técnicas de aprendizaje para trabajar con datos de grandes magnitudes. Debemos ser capaces de ser capaces de gestionar esas “cajas negras”. Hace parte del proceso de formación del profesional empaparse acerca del conocimiento sobre el cual está ejerciendo poder, las denominadas cajas negras, son los algoritmos utilizados para su predicción, pero el como funcionan es parte importante del proceso para entender su funcionamiento y correcto uso. De allí surgen técnicas para ya sea el Machine learnign o el Deep leargning.

El big data se puede definir como, gran volumen de datos generados a gran velocidad, pero también de gran VARIEDAD. Una vez el profesional logra esto, se le puede dar valor a esos datos.

Este es el objetivo final de todo profesional, tener la mayor capacidad de técnicas que le permitan, a partir de cualquier tipo de conjunto de datos, si importar la magnitud de ellos, extraer información valiosa acorde a las necesidades o problemas planteados.

Jhonatan Smith garcia muñoz

1039705595