



Actividad práctica

Leer detenidamente cada uno de los siguientes casos.

Caso 1: Walmart y el Big Data

6 marzo, 2020, en [Datos y Marketing](#) por [Joels Hidalgo](#)



Uno de los casos reales que más impacto ha generado, debido a la practicidad y adelanto, es el de Walmart. Este gigante del retail estadounidense tiene, probablemente, uno de los sistemas de Big data y de clasificación más avanzado y vanguardistas a la fecha. Desde hace tiempo atrás, antes de la llegada del Big data, Walmart estaba preocupado por entender, tanto los patrones de compra y comportamiento

de sus clientes, cómo el hecho de generar una gestión logística eficiente que le permitiera tener ventajas competitivas palpables además de una visión holística de su propio ecosistema de trabajo, fundamentado en la distribución, trato con proveedores y clasificación junto al seguimiento de toda su mercadería.

La punta de lanza la lleva la tecnología RFID. Este sistema, basado en etiquetas electrónicas radio receptoras, es capaz de guardar y transmitir gran cantidad de datos de cada producto, permitiendo la localización en tiempo real, además de cientos de datos referentes al tipo de producto, peso, calorías, fabricante, etc. El tener un sistema de clasificación centralizado, directamente desde el proveedor, le ha permitido a esta multinacional concentrar un gran poder de decisión y previsión sobre la gestión, almacenamiento y control de sus miles de productos en sus más de 20.000 tiendas. Walmart es consciente del valor de los datos y tanto más para su negocio basado en artículos minoristas de bajo precio a gran volumen.

La aplicación del método de clasificación comienza con la estrategia de integración empresarial. En principio la idea es almacenar desde el mismo proveedor todos los detalles, características e información relevante de un producto basándose en las etiquetas RFID y un sistema de ID único. Este a su vez se mantiene monitorizado de manera constante a través de lectores en los centros de distribución al igual que en los puntos de venta. De esa forma se mantiene un control en tiempo real de la mercadería saliente y la entrante. La integración permite, además de clasificar de manera eficiente, la generación de pedidos únicos a los proveedores de ciertas zonas, mismos que al llegar al centro de distribución son repartidos según las necesidades de consumo requeridas por cada tienda.

Esta agrupación y control permite entender muy bien los patrones de consumo por regiones geográficas y ante eventos particulares, lo que permite el aprovisionamiento previsorio basado en datos anteriores o históricos. La creación de estos clusters, por ciudad, región, gustos, estacionalidad, venta cruzada, etc., y los posteriores análisis resulta en una estrategia revolucionaria y altamente eficiente para el modelo de negocio de la compañía, además de aportar un control casi total sobre el inventario. Las ventajas



para los departamentos de marketing consiste fundamentalmente en un conocimiento casi omnipresente y omnicanal de las ventas regulares, cruzadas, periódicas, estacionales y excepcionales en cada canal de distribución.

Fuente: <https://comunidad.iebschool.com/bigdatamarkfans/2020/03/06/walmart-y-el-big-data/>

Lectura Complementaria

<https://www.eleconomista.es/tecnologia/noticias/9279795/07/18/Walmart-y-Microsoft-unen-sus-fuerzas-para-plantar-cara-a-Amazon.html>

Caso 2: Cómo Twitter usa Big Data e inteligencia artificial (IA)

- Publicada el 14 de mayo de 2021



Twitter es una plataforma de redes sociales donde 328 millones de usuarios activos mensuales realizan microblogs (comparten actualizaciones de 280 caracteres) con sus seguidores. Es una mezcla entre la mensajería instantánea y los blogs, o la mensajería social, pero también ha sido crucial para la presentación de noticias, la promoción de eventos, el marketing y los negocios. Tanto si eres fanático como si no, Twitter se ha convertido en la novena red social más grande del mundo y Cortex, el equipo de ingeniería interno de Twitter, ha recurrido al poder de la inteligencia artificial (IA) para ayudar

a mejorar la experiencia del usuario de la plataforma.

Cómo Twitter usa la IA en la práctica

El multimillonario Mark Cuban dijo en 2017 que compró acciones de Twitter porque “finalmente consiguieron actuar junto con la inteligencia artificial”. En 2016, Twitter compró la tecnología Magic Pony, una **consultoría de análisis de datos**, para aumentar el volumen de Cortex, que quiere “construir la plataforma de inteligencia artificial más avanzada del mundo, a escala de Twitter, para aplicar los algoritmos de inteligencia artificial más complejos a nuestros conjuntos de datos más desafiantes, sin problemas”.

Una de las formas en que Twitter utiliza la inteligencia artificial es para determinar qué recomendaciones de tweets sugerir en las líneas de tiempo de los usuarios con el objetivo de resaltar los tweets más relevantes para cada individuo. Antes de este cambio, Twitter mostraba a sus usuarios los tweets



en orden cronológico inverso. En la actualidad, los algoritmos escanean y puntúan miles de tweets por segundo para clasificarlos en el feed de cada usuario.

La plataforma de redes sociales también implementó inteligencia artificial para combatir el contenido inapropiado y racista en su plataforma, ya que las autoridades en el Reino Unido y Alemania aumentan las medidas y multas para prevenir el discurso de odio, las noticias falsas y el contenido ilegal en las redes sociales. En los primeros seis meses de 2017, Twitter eliminó casi 300.000 cuentas de terroristas que fueron identificadas por sus herramientas de inteligencia artificial; de hecho, el 95 por ciento de las cuentas suspendidas relacionadas con el terrorismo fueron identificadas por algoritmos en lugar de usuarios humanos.

Las herramientas de inteligencia artificial también admiten pequeños ajustes en Twitter que mejoran la experiencia general del usuario. Por ejemplo, sus herramientas de recorte de imágenes utilizan IA para recortar imágenes automáticamente de una manera mucho más atractiva y monitorean las transmisiones de video en vivo y las categorizan según el tema para mejorar su capacidad de búsqueda y ayudar a los algoritmos a identificar videos que los usuarios podrían estar interesados en ver. en sus feeds.

Detalles técnicos

De acuerdo con Agencia Big Data, una **empresa de inteligencia artificial**, el algoritmo de clasificación de Twitter ha tomado una gran cantidad de datos, los ha procesado a través de redes neuronales profundas y ha aprendido con el tiempo qué contenido sería relevante para cada usuario individual. Todos los tweets se puntúan en un modelo de clasificación que se utiliza para determinar la probabilidad de que un usuario valore ese contenido en su feed. El modelo de ranking considera el contenido del tweet en sí, y aplica **business intelligence**, incluyendo si va acompañado de una imagen o video y cuántos retweets o me gusta ha recibido; el autor del tweet para ver si tuvo interacciones pasadas con el autor y la fuerza de su conexión con el autor; y considera el tipo y el tono de los tweets que le han gustado en el pasado y cómo este tweet se parece a otros que parecía apreciar. Cuanto más alta sea la puntuación de relevancia, más alto será el tuit en tu feed y la probabilidad de que aparezca en el módulo “En caso de que te lo hayas perdido”.

A medida que Twitter continúa perfeccionando sus algoritmos, debe equilibrar la velocidad y la calidad del proceso de revisión del algoritmo para cumplir con los requisitos de la plataforma para las actualizaciones en tiempo real. Además de la velocidad y la calidad de las predicciones, los algoritmos se evalúan en función de sus demandas de recursos y su capacidad de mantenimiento a lo largo del tiempo.

La compañía utiliza IBM Watson y su conjunto de habilidades de procesamiento de lenguaje natural para rastrear y eliminar mensajes abusivos, ya que la tecnología de inteligencia artificial de Watson no solo comprende el lenguaje natural, sino que puede inferir la entonación y extraer el significado de las imágenes rápidamente, puede analizar millones de tweets en un segundo. .

Las herramientas de inteligencia artificial de Twitter “recortan utilizando la prominencia” para mostrar el aspecto más interesante de las imágenes, sean caras o no. Para entrenar la herramienta, el equipo de Twitter definió lo que es más destacado utilizando datos de académicos que estudiaron el seguimiento ocular. Luego, para optimizar la herramienta en tiempo real en el sitio, usaron IA para entrenar una versión más rápida de la herramienta para acelerar los cultivos. Primero, entrenaron a una red más pequeña usando el primer programa que era bueno, pero lento. Luego, los ingenieros de software simplificaron el proceso al eliminar las señales visuales menos importantes en las imágenes.

Según la **agencia de comunicacion corporativa**, para entrenar su algoritmo para reconocer lo que está sucediendo en una transmisión en vivo, Cortex utilizó el aprendizaje profundo. Entrenaron una gran red neuronal para reconocer el contenido en video a partir de una gran cantidad de ejemplos. Los humanos vieron videos y los etiquetaron con varias palabras clave para identificar lo que vieron. Entonces, un video de un perro no solo fue etiquetado con la palabra clave perro, sino también animal, canino, mamífero y más. Luego, esos datos se usaron para entrenar el algoritmo para que luego pudiera identificar el contenido en video.

Ideas y conocimientos que puedes robar



Como ilustra el equipo de Twitter, las herramientas impulsadas por inteligencia artificial se pueden utilizar para mejorar o mejorar sus productos y servicios. La colaboración entre humanos e inteligencia artificial ayuda a acentuar los mejores atributos de cada uno.

Extraído de: <https://www.linkedin.com/pulse/c%C3%B3mo-twitter-usa-big-data-e-inteligencia-artificial-ia-felix-garcia/?originalSubdomain=es>

Caso 3: La NBA y la NFL se suben a la data y la inteligencia artificial



Las ligas, asociaciones deportivas, equipos y corporativos tienen un aliado en la tecnología para mejorar las prácticas deportivas y las experiencias de los aficionados.

La NBA utiliza el big data para personalizar las experiencias de los aficionados.

La tecnología está revolucionando el deporte. Esta tendencia estaba en la cancha desde antes de la pandemia, pero con el confinamiento aceleró la adopción de big data, realidad virtual, inteligencia artificial e internet de las cosas desde los jugadores hasta los espacios deportivos y las experiencias para los aficionados.

Jugadores de equipos de fútbol soccer, americano o basquetbol y atletas de alto rendimiento suelen utilizar la data para desarrollar sus estrategias de competencia y también para monitorear el rendimiento que tienen durante sus entrenamientos, que les ayuda a reducir lesiones.

Estas estrategias de innovación se impulsan, en algunos casos, desde instituciones deportivas con alcance mundial, como es la Asociación Nacional de Basquetbol (NBA, por sus siglas en inglés), que ha desplegado una serie de proyectos para conectar con jóvenes que tiene interés en el deporte, hasta otros que permiten personalizar la publicidad de sus patrocinadores en sus partidos que se transmiten en línea.

Matt Wolf, director de Estrategia e Innovación Global de la NBA, dijo que el análisis de datos ha permitido a la organización enfocarse en el futuro, con la mira puesta en el desarrollo de una estrategia corporativa enfocada en la experiencia de los usuarios fuera y dentro de los estadios, pensando en la vida después del confinamiento.

Con los juegos sin la presencia de los seguidores, la NBA lanzó diversas iniciativas para mantener el contacto con los usuarios y, con la data, crea contenidos personalizados de acuerdo al perfil de cada uno de los aficionados. Estos proyectos nacen en un hub virtual, en donde los colaboradores pueden pelotear ideas, revisar avances hasta ver los proyectos materializados.

“Buscamos diferentes iniciativas y vemos esto no solo como una oportunidad, sino como la construcción de una cultura. Es pensar en nuevas oportunidades que se crean con el análisis de datos para enfocarnos en definir las direcciones y objetivos a largo plazo para definir la escala de nuestro negocio”, comentó durante su participación en el NBA Innovation Summit.

Wolf declara que como parte de la innovación, la asociación pasó mucho tiempo observando al interior, es decir en los equipos, y ahora han mirado fuera de la caja para desarrollar experiencias que van más allá de la transmisión en tiempo real de los partidos. Ahora, durante las transmisiones desarrollaron publicidad virtual personalizada, que permite poner logos o nombres de marcas de acuerdo al comportamiento de cada usuario de Twitch o NBA Pass. Para evocar a los usuarios de los estadios durante



los juegos en vivo, han creado stickers o máscaras virtuales de temáticas, como superhéroes, que los usuarios pueden compartir.

Este análisis de data también permite a la NBA vender a los usuarios de sus diferentes plataformas en línea paquetes coleccionables de sus equipos, o regalan souvenirs virtuales de los partidos, como ocurre cuando las personas asisten a los estadios.

Otras iniciativas en la NFL

Las experiencias con realidad aumentada e inteligencia artificial también van al alza. Ana Michelle Concepción, vicepresidenta de soluciones empresariales de AT&T México, considera que el deporte como entretenimiento permite ofrecer experiencias personalizadas a los aficionados con el uso de esta tecnología desde los estadios, aún con las limitaciones de movilidad.

Como ejemplo, explicó que el AT&T Stadium, el hogar de los Dallas Cowboys, en Texas, cuenta con una red LTD de internet dedicado, infraestructura de 3,500 monitores que procesan la data de lo que ocurre durante los juegos, que permite a los jugadores del equipo tener elementos que requieren para evaluar sus jugadas.

Del lado de los aficionados, instalaron un sitio de experiencia en 3D en la que los asistentes tienen acceso a la sala virtual de los Dallas Cowboys para poder tener fotos con los jugadores, que pueden compartir a través de redes sociales.

Además, se han desarrollado aplicaciones móviles que permiten a los asistentes ubicar sus asientos, las rutas para llegar a ellos y mapea cada uno de las ubicaciones, como sanitarios o áreas de alimentos. Con estas también es posible monitorear los asientos para saber el número de asistentes, cuántas veces se mueven de su lugar.

Con Chicago Bulls, AT&T trabaja en el desarrollo de una aplicación que permita a los usuarios tener estadísticas de los jugadores en tiempo real. “En todo este entorno, la tecnología hace un match perfecto con los deportes y ayuda a que los aficionados tengan una experiencia positiva. Es identificar necesidades y personalizar soluciones”, dice la directiva.

Extraído de <https://expansion.mx/tecnologia/2021/09/29/data-inteligencia-artificial-realidad-virtual-nba-nfl>

Caso 4: Cómo el big data predijo la cantidad de pacientes de COVID-19 en una terapia intensiva

Un sanatorio porteño y una empresa de base tecnológica trabajaron en la creación de un algoritmo que permite tomar decisiones con dos semanas de anticipación sobre la base de la información de la terapia intensiva, enriquecida con las estadísticas nacionales. Cómo una política de datos abiertos puede mejorar la gestión sanitaria.

Big data es un término que hace referencia al conjunto de datos y sus combinaciones, cuyo volumen, complejidad y velocidad de crecimiento hacen que no sea posible su captura, gestión, procesamiento o análisis con tecnologías convencionales. La recopilación de esas grandes cantidades de datos y el estudio de tendencias se hace mediante algoritmos que permiten predecir con altos grados de certeza lo que ocurrirá y, a partir de eso, tomar decisiones. El análisis de big data es un recurso que se puede aplicar en cualquier tipo de empresas o instituciones. Por ejemplo, en un hospital en medio de una pandemia.

La empresa RockingData trabajó en conjunto con profesionales de la terapia intensiva de un sanatorio privado de Buenos Aires —el nombre se mantiene en reserva— para organizar el enorme volumen de datos disponible, establecer patrones y prever cosas como cuántas camas debían tener libres, cuántos insumos eran



necesarios, cómo ordenar y planificar los descansos de los médicos y —sobre todo— cuántas personas iban a ingresar con COVID-19 a terapia intensiva.

Fredi Vivas trabajó en un algoritmo para que un sanatorio pudiera prever la cantidad de pacientes de COVID-19 que ingresarán a terapia intensiva. (Imagen: gentileza)

En los modelos que estructuran este tipo de proyectos residen algoritmos inteligentes que son capaces de encontrar patrones y de hacer predicciones o de modificarse, en caso de que la tendencia cambie. “El análisis de los datos otorga previsibilidades y respuestas a las preguntas que nos hacemos en estas circunstancias”, dice Fredi Vivas, CEO de RockingData. La ciencia de datos se basa en saber qué pasó para predecir qué puede ocurrir.

“Son herramientas, modelos predictivos que analizan, predicen y aprenden mediante los ejemplos. Cuantos más casos de análisis tengan, más precisión se conseguirá”, agrega Vivas, quien también se desempeña como coordinador académico del Programa Big Data de la Universidad de San Andrés y este año publicará su primer libro: ¿Cómo piensan las máquinas?, que terminó de escribir en mayo pasado y que nació cuando él mismo tuvo COVID-19 en 2020.

La pandemia hace más urgente una regulación tecnológica global

“Con la incorporación de más y más datos, la máquina encuentra sola las reglas para analizarlos”, agrega. De este modo, a través del algoritmo de predicción de la demanda, el sanatorio logró anticiparse al alto volumen de pacientes que recibió el sector de terapia intensiva. Además, permitió tomar decisiones estratégicas basadas en datos para determinar los turnos y rotación de personal y, así, cubrir de manera más segura y efectiva la situación generada por la pandemia.

La predicción de la demanda está centrada en el machine learning, un concepto que en su traducción al español tiene tres acepciones: aprendizaje automático, aprendizaje automatizado o aprendizaje de máquinas. El objetivo de Vivas y su compañía es desarrollar técnicas que permitan que las computadoras aprendan. Una obsesión y una pasión en las que aplica y combina el periodismo, la narrativa, la ciencia ficción y su profesión de ingeniero en sistemas.

El algoritmo de predicción se basó en la información sobre los pacientes que todavía estaban en el sanatorio y los datos de alta, empleados y datos históricos del sanatorio

Un presagio

“No hay tanto material en español sobre big data. Usé un lenguaje llano para que todos lo pudieran leer y entender. Hacía tiempo que quería escribir un libro, pero no me enamoraba ninguna idea, hasta que en septiembre del año pasado tuve COVID-19. Un día que estuve con mucha fiebre, me quedé dormido y soñé que miraba la calle desde mi ventana y veía todo prendido fuego”, repasa Vivas los detalles de esta escena.

En el sueño, mientras contempla la destrucción total, Fredi ve que unas máquinas que emulan el aspecto de los humanos están recorriendo el barrio. “Estoy sorprendido, mirando lo que ocurre, cuando se me acerca un hombre que se parece al personaje de Terminator 4 y me dice: ‘Fredi, te trajimos acá porque te necesitamos, vos entendés cómo piensan las máquinas’”. En ese momento Vivas se despertó, entre los dolores del cuerpo, las altas temperaturas y la inspiración y empezó a tramar una historia que introduce en la ficción los detalles del trabajo que junto a su equipo realizan a diario en la empresa.

El desafío y la solución

Cuando los integrantes de RockingData fueron convocados por las autoridades de la terapia intensiva del sanatorio tuvieron un desafío claro: el algoritmo debía predecir cuántas personas ingresarían con la patología COVID-19 en los siguientes 15 días.

El objetivo de Vivas y su compañía es desarrollar técnicas que permitan que las computadoras aprendan. (Imagen: gentileza)

En la Argentina se transitaba aún la primera ola de la pandemia. La herramienta debía ser de máxima utilidad para tomar decisiones estratégicas. El algoritmo de predicción se basó en la información sobre los pacientes que todavía estaban en el sanatorio y los datos de alta, empleados y datos históricos del sanatorio. A partir de esa base “pudo entrenarse y aprender de las curvas de todo el año”, describen en RockingData. De



este modo, se logró determinar que para fines de abril o comienzos de mayo de este año se iba a producir una situación parecida a la de junio o julio de 2020, cuando se generó un pico de ingresos.

Según Vivas, se obtuvo más de un 85 por ciento de precisión en la predicción de la demanda . “Con la implementación de este algoritmo el sanatorio pudo prever con dos semanas de anticipación el volumen de trabajo que tendría y, por lo tanto, qué recursos necesitaría —camas, refuerzo de personal, etc.— para afrontar el aumento de casos”, agregó.

El diseño del algoritmo se enriqueció con las estadísticas nacionales y, a fines de 2020, llegó a tener un ciento por ciento de precisión en las predicciones. “Los algoritmos necesitan un tiempo de aprendizaje. Nosotros trabajamos en el nuestro tres meses antes de que sirviera para la toma de decisiones sobre la gestión, organización y mejora en la capacidad de servicio del sanatorio”, agrega Vivas.

La calidad en los datos

Si bien desde comienzos de la pandemia muchas instituciones se pusieron en contacto con esta empresa, en un principio la calidad de los datos que se manejaban no era buena. “La data no era suficiente, ni veraz, ni existía calidad para tener predicción. No tenía sentido trabajar así, hubiera sido una irresponsabilidad de nuestra parte”, dice Vivas.

Cuando habla de la calidad de los datos se refiere a que en los albores de la pandemia existía demasiada información trunca, había registros sin completar, no se sabía cuáles eran los síntomas de los pacientes o faltaban las fechas de los ingresos en los hospitales. “Hay mecanismos para mejorar la calidad. Pero, por ejemplo, en muchos lugares la información estaba solo en papel y eso hace que no sea fiable, a lo que se suman los errores que puedan surgir en la carga de datos”, agrega.

Lo que Vivas denomina “huecos en la información” se puede cubrir con técnicas de mejora en la carga, aunque en algunas circunstancias es demasiado complejo. Para que se produzca un cambio real, sostiene, es necesario que toda la sociedad esté atravesada por una apertura de datos.

Extraído de: <https://www.redaccion.com.ar/como-el-analisis-de-big-data-predijo-la-cantidad-de-pacientes-de-covid-19-en-una-terapia-intensiva/>

Tarea:

Teniendo en cuenta que el punto de partida para organizar una infraestructura de Big Data es la selección de fuentes de datos y considerando el caso mencionado realice la siguiente actividad.

- 1- Describa brevemente el caso.
- 2- ¿Cuál es el objetivo cada implementación llevada a cabo?
- 3- ¿Qué significa información para cada caso?
- 4- ¿Tratan con datos históricos? ¿Cuál es su utilidad?
- 5- ¿Qué V de Big Data encuentra en cada solución? Describa y ejemplifique
- 6- Suponiéndose parte de la definición de la estrategia de Big data a diseñar, mencione al menos 6 fuentes de información, clasifique y justifique su elección. ¿Qué tipos de datos provienen de cada una?
- 7- ¿Utilizaron algún recurso específico para lograr su objetivo?
- 8- ¿Qué características deben reunir las fuentes para ser consideradas de Big Data?
- 9- Elabore una conclusión.

Para la presentación se solicita la construcción de Diapositiva, con gráficos, imágenes, y lo que considere necesario para exponer en clase, y un PDF con la información detallada. Ambos se deben compartir en el classroom.