

LA REVOLUCIÓN ANALÍTICA: EL BIG DATA EN LA PRÁCTICA

Los 'data driven business' (empresas como Google, Amazon, Facebook, Yahoo, Twitter, Netflix y el resto de la comunidad 'open source', que basan su éxito en el análisis de la información que reciben de sus usuarios) han generado con éxito un nuevo estándar de servicio consistente en proporcionar a sus clientes una experiencia única basada en una personalización absoluta. El Big Data ya ha dejado de ser un fenómeno para convertirse en una realidad. Pero ¿cómo lo emplean las empresas y qué ejemplos encontramos en cada sector? Aplican los viejos algoritmos a nuevos datos y les confieren usos que antes ni siquiera se nos habrían ocurrido

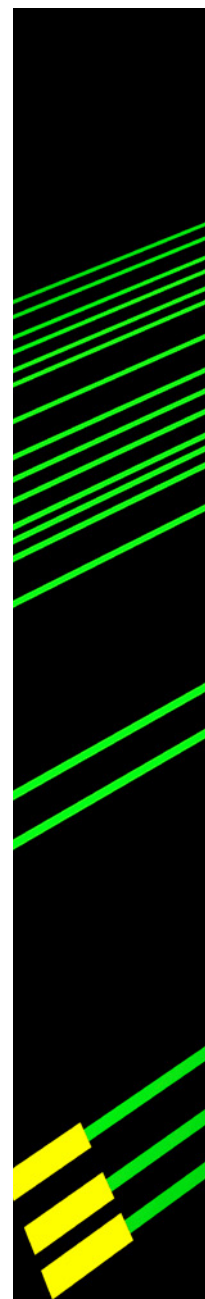
Juan José Casado Quintero

Director académico y profesor del Máster en *Business Analytics & Big Data* en el IE Business School

La realidad es que, a día de hoy, el Big Data no solamente impacta en el mundo de los negocios, sino que, verdaderamente, y sin que la mayoría nos demos prácticamente cuenta, lo está cambiando todo a nuestro alrededor, transformando nuestro mundo y hacien-

do realidad incluso algunos de nuestros sueños de futuro. Veamos algunos pocos ejemplos de ello:

- Barack Obama ganó sus primeras elecciones a la presidencia norteamericana gracias al desarrollo de un sistema que le permitía monitorizar las redes sociales y





adaptar su discurso según las necesidades individuales de cada localidad que visitaba.

- Existen sistemas de inteligencia artificial que ayudan a profesionales médicos a diagnosticar enfermedades y a detectar epidemias anticipadamente cada día.

- Equipos modestos de béisbol logran resultados espectaculares gracias al análisis masivo de las estadísticas de los jugadores.

- Aplicaciones musicales analizan nuestros gustos para saber qué canción nos apetecería escuchar en cada momento. >>>

- Empresas de moda usan las redes sociales como YouTube para detectar tendencias en el momento en que aparecen.
- Compañías de *video on demand* como Netflix aseguran el éxito de sus series de TV (*House of cards*) en función de la extracción, transformación y lectura de los datos procedentes de las recomendaciones de los usuarios.
- Durante la semana en que se celebró el Festival de Eurovisión, en mayo de 2015, un grupo de ingenieros de Microsoft especializados en matemáticas predictivas

LOS 'DATA DRIVEN BUSINESS' DESARROLLAN MODELOS PREDICTIVOS MEDIANTE UN COMPLEJO PROCESO, CASI ARTESANAL, LIDERADO POR LA FIGURA DEL 'DATA SCIENTIST'

usó el Big Data para tratar de predecir cuál sería el resultado: combinando las búsquedas realizadas en los buscadores con el entusiasmo y sentimiento que los distintos cantantes despertaban en las redes sociales, confeccionó una lista de los países con más probabilidades de llevarse a casa la victoria.

Pero, sin duda, al frente de todas estas iniciativas se encuentran empresas como Google, Amazon, Facebook, Yahoo y Twitter, que están dominando el mundo de los negocios en la actualidad. Todas han descubierto un nuevo modelo de negocio basado en desarrollar las capacidades necesarias para analizar los datos de sus clientes y adquirir un conocimiento que les permite ofrecerles una experiencia realmente diferencial y personalizada, con lo que obtienen una ventaja competitiva sostenible frente al resto de competidores. Lo hace Google cada vez que es capaz de personalizar la publicidad que nos muestra en las páginas web basándose en nuestros gustos, inferidos a través de las búsquedas que reali-

zamos. Lo hizo Amazon cuando cambió completamente la experiencia de comprar en Internet con su aclamado servicio "Clientes como tú...", que era capaz de analizar las compras realizadas y recomendar los libros que habían comprado usuarios con gustos afines. Y lo hace Facebook cada vez que selecciona y prioriza para nosotros, en base a su famoso algoritmo Edge Rank, qué contenido, de todo el que generan aquellos a quienes seguimos (amigos, conocidos y empresas), debe mostrarnos en nuestro *timeline* para que continuemos cada día "enganchados" a su red.

¿CÓMO LO HACEN?

Estos *data driven business* han generado un nuevo estándar de servicio basado en la personalización absoluta de la experiencia del cliente, al que el resto de industrias se tendrán que adaptar si quieren sobrevivir en la nueva economía de servicios digitales que gobernará nuestro siglo. ¿Cómo lo hacen? Generan algoritmos: desarrollan modelos predictivos mediante un complejo proceso, casi artesanal, liderado por la figura del *data scientist*, una especie de "alquimista" moderno que es capaz de extraer y destilar conocimiento de los datos. Para ello, el método que sigue el científico de datos consiste en:

- 1. Simplificar** el problema de negocio hasta formular una pregunta a la que se quiere dar respuesta.
- 2. Limpiar** las impurezas que se acumulan en los datos.
- 3. Seleccionar** el o los algoritmos matemáticos y computacionales más adecuados.
- 4. Programar** dichos algoritmos matemáticos con ayuda de algún software estadístico que, al procesar su juego de datos, le dará una respuesta y el grado de confianza en esa respuesta en función de la bondad estadística del modelo que ha realizado.
- 5. Recomenzar** el proceso si la respuesta es equivocada o la bondad del modelo no es suficiente. Para ello, añadirá nuevas variables y/o suprimirá otras, basándose en su intuición, hasta dar con una fórmula que le satisfaga, y también a sus usuarios de negocio.

POR SECTORES Y BUENOS EJEMPLOS

No es de extrañar, por tanto, que, en el último Global C-Suite Study, encuesta realizada por IBM a más de cuatro mil ciento ochenta y tres directivos de setenta países, se revelara que el 71% de los miembros de comités de Dirección consultados había declarado que desarrollar capacidades analíticas era su principal preocupación y la principal necesidad para generar ventajas competitivas. Por eso, no son pocas las industrias que, de una manera u otra, ya se han puesto manos a la obra, consiguiendo muy buenos resultados. Veamos qué sectores concretos están reorientando sus estrategias hacia la “revolución analítica” y qué ejemplos de éxito se dan en cada uno:

► **PRECIOS DINÁMICOS EN EL SECTOR TURÍSTICO.** Sin duda, el sector turístico ha sido uno de los primeros en entender cómo podía sacar partido del análisis masivo de información en tiempo real. En un sector en el que la mayoría de los costes son fijos y los costes variables son mínimos, los precios se convierten en algo tremendamente flexible. Tanto es así que las compañías hoteleras o las líneas aéreas nos han acostumbrado a que veamos variar dinámicamente en cada momento los precios que publican en Internet. Los cambios son controlados por algoritmos inteligentes que deciden en milisegundos qué precio mostrar en nuestra pantalla, en función de cuántas plazas quedan disponibles, cuántas personas están viendo ese mismo producto en ese mismo instante o cuán flexibles al precio somos según nuestro historial de compras. Esto significa “teras” de información analizados en tiempo real para mostrar exactamente el precio justo que hará que decidamos hacer clic e irnos de escapada ese fin de semana.

Caso de éxito. Si hay una compañía turística que está sabiendo captar todo el potencial del Big Data, esa es, sin duda, Uber. Esta compañía internacional proporciona transporte a sus clientes mediante una conexión entre pasajeros y conductores de vehículos basada en una aplicación móvil que organiza los contactos en ciudades de todo el mundo. Aparte de lo controvertido del mecanismo, que ha sido puesto en tela de juicio por empresas de taxis,

resulta interesante analizar la aplicación de la tecnología Big Data en este negocio. El registro de los datos de usuarios pasajeros y conductores es la clave de un sistema fuertemente apoyado en la minería de datos, la gestión de la información y la tecnología GPS. ¿Cómo funciona? El precio de los trayectos se fija de antemano por la propia app en función de la dis-



tancia a recorrer, gracias a un algoritmo de cálculo de precio.

Patentado bajo el nombre Surge Pricing, este sistema se apoya en herramientas analíticas que no solo tienen en cuenta la oferta y la demanda para un trayecto en concreto, sino que, además, son capaces de calcular, en tiempo real y usando datos de GPS, la duración estimada de ese trayecto en función de la monitorización de las condiciones del tráfico ►►►

➤➤➤ existente en cada punto de la ciudad a cada instante, lo que permite al sistema establecer las rutas más eficaces. La clave de su éxito radica en que, de esta manera, el usuario puede saber de antemano cuánto le va a costar cruzar la ciudad ese día en hora punta, pero evitando verse atrapado en un taxi parado en medio de un embotellamiento y con el taxímetro subiendo despiadadamente.

► **LA BANCA DEL CONOCIMIENTO.**

El proceso que vive la banca hoy en día es el mismo que viven la mayoría de las industrias ya maduras: los productos se “comoditizan”, es decir, no consiguen diferenciarse unos de otros más allá del precio, por lo que, en muchos sectores, como la banca, hasta los precios acaban siendo semicontrolados por las auto-

ra compensarle por este incidente, como sabemos que es un gran aficionado al golf, tendrá una bonificación del 50% en cualquier compra que realice durante el día de hoy en el Golf Store que se encuentra en el centro comercial de esa misma calle. Podrá abonar sus compras a través de nuestro servicio de pago móvil, que usted ya activado en su iPhone 6. Aprovechamos para recordarle, si nos lo permite, que su seguro de automóvil vencerá próximamente y que, tras realizar una proyección de su saldo este mes, podría usted quedarse en descubierto. Pero no se preocupe; por si eso llega a ocurrir, ya le hemos preconcedido un crédito que le evitaría el pago de comisiones de descubierto, ahorrándole 27,36 euros. Nos hemos permitido analizar también su póliza, y este año, setenta y cuatro clientes

EN EL SECTOR TURÍSTICO, LOS CAMBIOS SON CONTROLADOS POR ALGORITMOS INTELIGENTES QUE DECIDEN EN MILLISEGUNDOS QUÉ PRECIO MOSTRAR EN NUESTRA PANTALLA, EN FUNCIÓN DE CUÁNTAS PLAZAS QUEDAN DISPONIBLES O CUÁNTAS PERSONAS ESTÁN VIENDO ESE MISMO PRODUCTO EN ESE MISMO INSTANTE

ridades (el Banco de España hace tiempo que ha puesto techo a la retribución de los depósitos). En este escenario, las entidades solo pueden competir y diferenciarse unas de otras a través de la experiencia de los clientes. Pero, para poder ofrecerles una experiencia diferencial y completamente personalizada, lo primero, e indispensable, es conocerles hasta el más mínimo detalle, para darles lo que quieren en cada momento.

Si nos dedicamos a este negocio, la “banca del conocimiento” significa que, cuando nos llame un cliente, podamos anticiparnos y decirle: “Sr. Martínez, sabemos que nos llama porque el cajero en el que usted se encuentra ahora se acaba de estropear y se ha quedado con su tarjeta. Disculpe las molestias y no se preocupe, porque ya le hemos enviado una nueva a su domicilio. Además, le facilitamos un código SMS con el que podrá sacar gratuitamente dinero en otro cajero que está a doscientos metros de su localización actual. Pa-

con su mismo vehículo y compañía acaban de cambiar su seguro a la compañía ‘X’, con lo que han conseguido ahorrar un 28% anual. Esta compañía trabaja con nuestra entidad, por lo que, si lo desea, con solo decir la palabra ‘autorizo’, nosotros mismos llevaremos a cabo todos los trámites para gestionar su cambio de póliza. El ahorro calculado para usted es de 254 euros al año. ¿Desea algo más?”.

Todo este largo párrafo, que puede parecer hoy ciencia ficción, no lo es, realmente. Las entidades financieras trabajan ya en desarrollar lo que llaman “visión 360 grados del cliente”, algo que les permitirá no solo conocer al dedillo todos nuestros datos financieros, sino también nuestros gustos, aficiones, hábitos y estilos de vida, inferidos a través de los pagos que realizamos por tarjeta, nuestras domiciliaciones, recibos e, incluso, nuestra cuenta de Twitter. Podrán saber, por tanto, si llevamos un estilo de vida lujoso o ahorrador, si nos gusta salir a cenar o a co-

mer los fines de semana, si viajamos de manera frecuente, si somos deportistas... Todo ello, con el objetivo de desarrollar potentes algoritmos, motores de decisión que en tiempo real nos ofrecerán productos y servicios adaptados completamente a nuestras necesidades y que nos tienen que garantizar, *a priori*, una experiencia verdaderamente diferencial en cada uno de nuestros contactos con la entidad.

Caso de éxito. Francisco González, presidente de BBVA, declaraba en diciembre de 2013, en una entrevista titulada *Banks need to take on Amazon and Google or die*, publicada en *Financial Times*, que "la gran competencia" de la banca del futuro la constituirán Google, Amazon, PayPal y Facebook, y que todos aquellos bancos que no sean capaces de desarrollar las mismas capacidades que tienen estas compañías para dar a sus clientes exactamente lo que necesitan, gracias al análisis de sus datos, están destinadas a desaparecer. Pues bien, si algo hay que "admirar" en el presidente de BBVA no es solo su visión de futuro, sino la capacidad de hacer realidad lo que dice y de predicar con el ejemplo: en 2014, el banco anunciaba la creación, no de un departamento, sino de toda una empresa dedicada al Big Data, un *spin-off* bautizado como BBVA Data Analytics.

El objetivo de esta empresa es maximizar el enorme valor que tienen los datos de los millones de clientes del banco (movimientos de cuentas, contratos, e-mails, llamadas telefónicas...). Pero, sobre todo, donde ha encontrado un auténtico filón es en los millones de transacciones al día que suponen los pagos realizados con las tarjetas de crédito, perfectamente geolocalizados a través de los terminales de punto de venta y debidamente agregados y anonimizados para salvaguardar la privacidad de sus clientes.

Son múltiples los procesos del banco que pueden beneficiarse de incluir nuevos algoritmos matemáticos o modificar los tradicionales incorporando esta nueva información, como el proyecto Risk360, que busca predecir la probabilidad de *default* de un negocio a partir de la información contextual de sus pagos. Si el nuevo algoritmo detecta, por ejemplo, que un negocio, ya sea una peluquería o una panadería, al compararlo con establecimientos similares de su zona, presenta

más volumen de transacciones y de mayor importe medio, y, además, sus clientes acuden desde mayor distancia a comprar allí, obtendrá un mejor *scoring* de riesgo, teniendo más probabilidades de obtener un crédito, y,

LAS ENTIDADES FINANCIERAS TRABAJAN YA EN DESARROLLAR LO QUE LLAMAN "VISIÓN 360 GRADOS DEL CLIENTE", ALGO QUE LES PERMITIRÁ, NO SOLO CONOCER AL DEDILLO TODOS NUESTROS DATOS FINANCIEROS, SINO TAMBIÉN NUESTROS GUSTOS, AFICIONES, HÁBITOS Y ESTILOS DE VIDA

el banco, mayor seguridad de que lo va a recuperar en el futuro.

Pero las posibilidades de explotar esta información exceden más allá de la propia entidad financiera, porque son muchos los actores que podrían beneficiarse de su correcta explotación. Como el Ayuntamiento de Madrid, que, en una colaboración conjunta realizada recientemente para estudiar el turismo de la capital, tras analizar los datos de los pagos generados con tarjeta durante el desfile del Orgullo Gay de los años 2011 y 2012, decidió rectificar el recorrido de 2014 para beneficiar a decenas de pequeños comercios y potenciar las oportunidades económicas del evento.

► **MONETIZACIÓN DE LOS DATOS EN LAS COMPAÑÍAS DE TELECOMUNICACIONES.** Los nuevos competidores digitales, como WhatsApp o Skype, están "obligando" a las compañías de telecomunicaciones a buscar nuevas formas de ingresos que hagan sostenible su negocio en el futuro. Las más innovadoras están apostando por la diversificación y la apertura de nuevas líneas de negocio, como sería, por ejemplo, la monetización del Big Data. ¿En qué consiste? ►►►

➤➤➤ Pues se trata de algo tan sencillo como aprovechar todos los datos que tienen estas compañías de sus clientes para identificar a aquellos que puedan ser de interés a otras empresas, y, una vez anonimizados y agregados debidamente para salvaguardar la privacidad, comercializarlos y obtener ingresos adicionales.

Caso de éxito. Esta monetización es precisamente lo que está llevando a cabo Telefónica a través de Smart Steps, un producto que se encarga de procesar de manera masiva todos los eventos de red, señales que emiten nuestros móviles constantemente a la red de antenas con el objetivo de estar geolocalizados y poder realizar llamadas. El objetivo es saber, en cada punto de la ciudad, cuántas personas transitan en cada momento del día. Pero no solo eso, sino poder analizar también qué edad tienen, si son mayoría de hombres o mujeres, qué estilo de vida llevan, qué modo de transporte utilizan, la frecuencia con la que realizan esa ruta, o incluso saber de dónde proceden y a qué distancia queda su lugar de trabajo o el domicilio habitual. Todo ello debidamente presentado a través de un sistema de información geográfica que representa esta información

OTRO CLARO EJEMPLO DE CÓMO EL BIG DATA VA A CAMBIAR RADICALMENTE LA FORMA TRADICIONAL DE HACER LAS COSAS EN UN SECTOR ES EL DE LOS SEGUROS DE AUTOMÓVIL

de manera visual y fácilmente analizable. Esta información permite hacer estudios estadísticos geotemporales que ayudan a analizar áreas de consumo, hacer comparaciones entre zonas para saber cuál tiene mayor potencial para la captación de clientes, adaptar promociones de *street marketing* en ubicaciones con mayor movilidad de personas, determinar cuál es la mejor zona para abrir un nuevo local... Pensemos, por ejemplo, que somos una gran

cadena de moda que quiere abrir una nueva tienda y que se debate entre varios locales comerciales disponibles. ¿Cuánto estaríamos dispuestos a pagar por obtener una información que puede suponer el éxito o el fracaso de nuestro futuro negocio?

Pero no solo el sector *retail* es un potencial usuario de este tipo de servicios. El sector turístico, el transporte y toda la Administración Pública en general necesitan hacer uso intensivo de estas nuevas fuentes de información. Por ejemplo, si fuésemos los responsables de



movilidad de una gran ciudad y tuviésemos que programar la frecuencia de una línea de metro, elegir el recorrido y paradas de los autobuses o decidir cuándo es el mejor momento para acometer una nueva obra de infraestructura, ¿no nos sería útil poder disponer de los flujos de movimiento entre los distintos barrios de la ciudad por tramo horario? ¿Cómo afectaría a la ciudadanía disponer de una información que permitiese mejorar la toma de

este tipo de decisiones, que tanto nos impactan en el día a día?

► **SEGUROS BASADOS EN NUESTROS COMPORTAMIENTOS.** Otro claro ejemplo de cómo el Big Data va a cambiar radicalmente la forma tradicional de hacer las cosas en un sector es el de los seguros de automóvil. Como todos sabemos, hasta ahora, el precio de nuestro seguro venía influido por las características particulares de nuestro vehículo, más allá de lo bueno o malo que fuera cada uno

frenamos bruscamente, el número de accidentes de cada tramo de vía que transitamos o la frecuencia con la que realizamos un mismo trayecto son solo algunos de los datos que, una vez analizados por un potente algoritmo matemático, determinarán el precio más adecuado para cada cliente en función de su comportamiento como conductor y, por tanto, del riesgo calculado que tenga de estar involucrados en un accidente.

Caso de éxito. Aunque en Europa apenas se han dado a conocer hasta hace unos meses, en Estados Unidos, los seguros de automóviles *Pay as you Drive* surgieron ya hace unos cuantos años, y más de treinta compañías aseguradoras ofrecen esta modalidad. De entre todas ellas, destaca Progressive Insurance, la cuarta aseguradora del país, pionera en el uso de tecnologías telemáticas y que, además, cuenta con *Snapshot*, el programa *Pay as you Drive* más popular entre los conductores americanos. Consiste en la instalación de un dispositivo que, durante treinta días, monitoriza los hábitos de conducción y analiza parámetros como el tiempo que le dedicamos a conducir, a qué horas del día o con qué frecuencia frenamos violentamente o precipitadamente. Progressive puede aplicar descuentos hasta del 30% a los mejores conductores, es decir, a aquellos que eviten conducir a las horas en que ocurren más accidentes de tráfico o a los que conduzcan menos tiempo.

Snapshot fue lanzado en enero de 2008, y desde entonces ha logrado unas espectaculares cifras solo comprensibles desde la óptica del Big Data: su base de datos, de más de ciento diez *terabytes* (unas siete veces la cantidad de información almacenada en la librería del Congreso de Estados Unidos), ha sido generada por los dos millones de clientes del programa, que han realizado un total de 1.500 millones de desplazamientos de 10,6 kilómetros de media cada uno, lo que significa tener almacenados y analizados unos 16.000 millones de kilómetros de conducción, o, lo que es lo mismo, el equivalente a 50 viajes de ida y vuelta desde el Sol a la Tierra.

En lo económico, *Snapshot* no ha parado de crecer desde su lanzamiento. Actualmente genera unos dos mil seiscientos millones de euros en pólizas anuales, lo que representa ya un 15% del negocio total de la compa- ►►►

al volante. La capacidad de procesar millones de registros que facilitan las herramientas de Big Data va a permitir a las compañías de seguros cambiar esta forma tradicional, e injusta, de fijar los precios.

Para ello, algunas compañías se han lanzado a instalar sensores en los vehículos de sus asegurados, que recogen vía satélite gigas de información cada vez que suben a ellos. La velocidad a la que conducimos, las veces que



►►► ña, que, además, no cesa de probar nuevos modelos de negocio: desde 2015, no solo los buenos conductores se verán beneficiados con un descuento de hasta el 30% en su póliza, si no que Progressive empezará también a penalizar a los malos conductores con mayor riesgo de accidente con hasta un 10% adicional en su póliza.

En España, Generali, con una modalidad de seguro bajo el nombre “Pago como conduzco”, y Mapfre, con el modelo llamado “YCAR”, son las primeras aseguradoras que han introducido en el mercado español este tipo de pólizas.

► **LA PROMESA DE UNA MEDICINA DE PRECISION.** Pero no todo es maximizar beneficios en el mundo del Big Data. Si hay un sector en el que la capacidad de analizar información para la toma de decisiones más ciertas cobra especial relevancia es el de la salud, donde una decisión correcta puede suponer la diferencia entre la vida y la muerte para un paciente. Pero el Big Data no solo está presente en el mundo de la salud con el uso de técnicas de inteligencia artificial para apoyar o mejorar las capacidades diagnósticas de los médicos. Las herramientas de Big Data han sido claves también para acelerar y abaratar los costes de la secuenciación del genoma humano y para facilitar el uso de dispositivos *wearables*, que recogen mediante sensores millones de datos generados por nuestro organismo. Son estos

**POCAS PERSONAS
DEBIERON CREER, CUANDO
VIERON EN 2002 LA
PELÍCULA ‘MINORITY
REPORT’, QUE VIVIRÍAN
PARA VER UN ORDENADOR
AYUDANDO A PREDECIR
DÓNDE Y CUÁNDO SE VA A
PRODUCIR UN CRIMEN**

dos de los cambios tecnológicos que nos permitirán hacer realidad la promesa de una verdadera medicina de precisión. De esta forma, los médicos serán capaces de personalizar paciente a paciente el tratamiento no solo en fun-

ción de sus pruebas clínicas, sino que sofisticados algoritmos de decisión analizarán todas y cada una de nuestras características individuales: genética, hábitos alimenticios, forma física, estilo de vida e, incluso, estado de ánimo. Variables todas ellas que serán tenidas en cuenta para asegurar que se da a cada paciente el tratamiento adecuado en el momento adecuado, para garantizarle así los mejores resultados de salud.

Caso de éxito. Que una decisión correcta puede suponer la diferencia entre la vida y la muerte para un paciente es lo que quizá debió de pensar IBM cuando, tras desarrollar Watson, un ordenador capaz de entender cualquier pregunta formulada por una persona y responder con certeza en tiempo real, que se hizo famoso tras ganar el concurso de preguntas y respuestas más popular de la TV americana (*Jeopardy*), decidió afrontar el reto de enseñarle medicina. Un año después, en colaboración con el prestigioso Centro Oncológico Memorial Sloan-Kettering, Watson ha desarrollado el conocimiento necesario para realizar diagnósticos y ofrecer recomendaciones de tratamiento a personas diagnosticadas con cáncer.

Tras incorporar más de medio millón de pruebas médicas, alrededor de dos millones de páginas de texto de investigación sobre cáncer de mama y de pulmón y los registros de 1,5 millones de pacientes diagnosticados, los expertos enseñaron a Watson cómo procesar, analizar e interpretar el significado de toda esta información durante meses, en un esfuerzo que, sin duda, mereció la pena y que ha cambiado el rumbo de la medicina para siempre. Actualmente, Watson ha demostrado ser capaz de diagnosticar correctamente el 90% de los casos de cáncer de pulmón, mientras que, de media, un especialista en oncología solo es capaz de diagnosticar correctamente el 50%.

► **ANALÍTICA CONTRA EL CRIMEN.** Pero si ha habido un entorno en el que el uso del Big Data siempre ha llevado aparejada polémica, este ha sido el de la seguridad, por las múltiples consideraciones legales, éticas y de privacidad que trae consigo. La capacidad de las nuevas tecnologías de Big Data para analizar en tiempo real datos no estructurados (texto, audio, imágenes o videos) ha convertido algoritmos y modelos predictivos en las nuevas

“armas” de la lucha contra el crimen. Tras los atentados del 11 de septiembre en Estados Unidos, los gobiernos americanos e israelí dieron un fuerte impulso al desarrollo de sistemas capaces de analizar millones de comunicaciones telefónicas y correos electrónicos para detectar redes de terroristas mediante complejos algoritmos de minería de texto y de análisis del lenguaje natural. La aplicación de algoritmos de redes neuronales para el tratamiento de imágenes permite hacer el *matching* entre las personas que aparecen en un video o una fotografía y en cualquier base de datos de criminales para identificar posibles sospechosos. Y mediante algoritmos de lógica difusa que ana-

variables que caracterizan cada uno de los sucesos delictivos acontecidos (hora del día, día de la semana, calle y número, tipo de crimen cometido, personas implicadas, su edad y sexo...), y se ha enriquecido con otras fuentes de información de la ciudad que aportan datos sociodemográficos y medioambientales de cada punto urbano (desde el número de viviendas a la temperatura, el ciclo lunar o la hora a la que amanece y anochece, pasando por la altura media de los habitantes de cada zona o su nivel de ingresos).

A toda esta gran base de datos se han aplicado los más modernos algoritmos predictivos para identificar patrones que permiten

SOLO NOS QUEDA TERMINAR DE ENTENDER CÓMO PODEMOS IMPULSAR EL USO DE TODAS ESTAS TECNOLOGÍAS SIN QUE EL HACERLO SUPONGA UN RIESGO PARA LA SEGURIDAD Y LA PRIVACIDAD DE LAS PERSONAS

lizan números de teléfono, nombres, *nicks*, etc. puede identificarse a un mismo individuo que utiliza identidades diferentes en Internet, lo que permite a los investigadores conseguir desarticular redes de fraude en la red.

Caso de éxito. Pocas personas debieron creer, cuando vieron en 2002 la película *Minority Report*, que vivirían para ver un ordenador ayudando a predecir dónde y cuándo se va a producir un crimen. De hecho, muchos todavía hoy lo considerarían ciencia ficción o algo que, al final, es meramente imposible. Excepto en Memphis, donde, si preguntásemos a cualquier policía, nos respondería que para ellos esto es algo tan normal como comerse un Donut para desayunar. Y es que, efectivamente, el Departamento de Policía de Memphis (Tennessee) ha implantado un innovador sistema de Big Data, en colaboración con IBM, que está ayudando a los agentes en su lucha contra el crimen.

El primer reto ha sido digitalizar la información en papel para, posteriormente, dotarla de una estructura que permitiera poderla analizar. Sobre toda esta información se han aplicado algoritmos de procesamiento del lenguaje natural que han posibilitado identificar las

desarrollar un modelo predictivo que, basándose en los cientos de variables analizadas, indique con un grado de fiabilidad suficiente cuál es, en cada punto de la ciudad y en cada momento del día, la probabilidad de que se cometa un acto criminal. Con esta información ya nos podemos imaginar que, cada día, se definen las rutas exactas que tienen que realizar las diferentes patrullas, permitiendo a los agentes estar en el sitio adecuado en el momento adecuado y persuadiendo a los criminales de hacer algo de lo que luego puedan arrepentirse, lo que les ha facilitado reducir la tasa de criminalidad en más de un 30% en los dos últimos años.

Y es que, como nos demuestra este último caso, no existen apenas límites en cuanto a lo que podemos hacer en el ámbito del Big Data para mejorar la sociedad en la que vivimos. Ahora solo nos queda terminar de entender cómo podemos impulsar el uso de todas estas tecnologías sin que el hacerlo suponga un riesgo para la seguridad y la privacidad de las personas, pero sin permitir tampoco que leyes del siglo pasado nos hagan evolucionar más despacio de lo que lo hace el pensamiento abierto de nuestra sociedad digital. ■