

BIG DATA:

EL NUEVO PETRÓLEO DE LA ECONOMÍA

**JAVIER ZAMORA**Colaborador científico,
Sistemas de Información,
IESE

Hoy en día tenemos acceso a más datos de los que nuestros sistemas son capaces de absorber. Podemos usar esta información para mejorar nuestra toma de decisiones, pero solo resultará eficaz si sabemos lo que buscamos y distinguimos lo importante de lo superfluo.

La información es poder, pero solo si se sabe qué hacer con ella. En estos momentos estamos entrando en una nueva era, la del *big data*, la de la información que va más allá de los típicos datos transaccionales que manejan las empresas. Cada día nos enfrentamos a una ingente cantidad de información procedente de las redes sociales y de los dispositivos, como el móvil, que también transmiten un gran volumen de datos. Sin embargo, esta información debe cribarse del mismo modo que se criba el oro, ya que si no se filtra y traduce en datos que tengan sentido y permitan tomar buenas decisiones, carecerá de todo valor.

EL VALOR DE LOS DATOS

Existen tres fuentes principales de *big data*. La primera hace referencia a los tradicionales datos estructurados que las empresas poseen sobre sus clientes o que recaban con cada transacción. En segundo lugar, existen múltiples datos no estructurados que engloban desde la conducta de los clientes al navegar por una página web, hasta los *blogs* o el tono y los sentimientos que se expresan en un correo electrónico, sin olvidar los comentarios en las redes sociales acerca de distintos productos y servicios. La

tercera fuente de información son los dispositivos, como el móvil, que aportan mucha información interesante sobre el usuario, empezando por su ubicación y continuando con las transacciones y llamadas realizadas o las aplicaciones usadas. La cuestión es: ¿cómo puede usarse toda esta información para mejorar el proceso de toma de decisiones?

El *big data* se caracteriza por las tres “uves”: volumen, velocidad y variedad. Y una cuarta “uve” se refiere al valor que puede obtenerse a partir de estos datos. Las empresas deben entender que manejar el *big data* no es una mera cuestión tecnológica, sino que requiere un cambio en la cultura y los procesos corporativos.

Muchas decisiones empresariales se basan en la intuición y, aunque el *big data* no pretenda reemplazarla, sí puede aportar más información. No se trata de elegir entre las corazonadas y los datos, sino de combinar el instinto con el pensamiento analítico. Buen ejemplo de ello es la película *Moneyball: Rompiendo las reglas*. El filme relata la historia real de **Billy Beane**, propietario de los Oakland Athletics, y su lucha por crear un equipo de





EL *BIG DATA* REQUIERE UNA NUEVA MANERA DE PENSAR, A LA VEZ QUE PLANTEA IMPORTANTES PROBLEMAS COMO LA PRIVACIDAD, LA SEGURIDAD Y EL RIESGO DEL USO DE LA INFORMACIÓN

béisbol ganador con una mínima parte del presupuesto de sus competidores. Según la intuición –personificada por los ojeadores del club–, la única manera de ganar consiste en fichar a estrellas consolidadas o a jóvenes promesas de este deporte, pero el Oakland no puede permitírselo.

El béisbol es un deporte basado en los números, en el que cualquier aficionado conoce el promedio de bolas bateadas y los *strikes* conseguidos por un determinado jugador. En esta película, un genio de la informática que trabaja en las oficinas del club realiza un análisis estadístico de varios jugadores que podrían ser asequibles para el equipo, bien porque se encuentran al final de su vida deportiva, bien porque su carrera se ha visto afectada por diversos problemas sociales. Este analista convence a **Beane** de que la suma de sus talentos les permitiría crear, desde un punto de vista estadístico, un equipo ganador. Tanto los ojeadores como los directivos del equipo encuentran la idea ridícula, pero tras un inicio de temporada desastroso, el Oakland consigue ganar veinte partidos seguidos y batir el récord de victorias consecutivas.

Esto no significa que a los ojeadores del club les hubiera fallado el instinto. De hecho, muchos grandes equipos de béisbol se han forjado alrededor de unas pocas estrellas apoyadas por un grupo de jugadores complementarios. Lo que demuestra esta historia es que los datos pueden ser un buen complemento de la intuición y la experiencia.

BUSCAR LAS CONEXIONES ENTRE LOS DATOS

- El volumen de datos que es capaz
- de manejar un ordenador permite establecer correlaciones que no serían identificables de otro modo. Por ejemplo, en la área metropolitana de Atlanta (EE.UU.), se decidió realizar un estudio en las escuelas públicas del condado de Gwinnett, el decimocuarto más grande el país, que siempre obtenía malos resultados académicos, pese a haberse implantado varias medidas de mejora, como aumentar el número de profesores o reducir la cantidad de alumnos por clase. Al analizar los datos de dicho distrito se observó que los

alumnos que se habían graduado con las mejores notas destacaban en Álgebra 1 y, a continuación, se descubrió que los alumnos que sacaban buenas notas en Álgebra 1 también obtenían buenos resultados en la asignatura de Escritura creativa. Esta correlación no es evidente a simple vista, pero cuando se dispone de un gran volumen de datos, aparecen nuevas conexiones que permiten abordar y resolver un problema desde nuevos enfoques.

El pionero en las conexiones basadas en grandes volúmenes de datos fue Google, que contrató como director económico a **Hal Varian**, profesor emérito de Economía de Berkeley. Google aprovecha la enorme cantidad de datos a la que tiene acceso para mejorar sus modelos predictivos. Por ejemplo, la herramienta Google Flu permite trazar un mapa de las personas que buscan información sobre los síntomas de la gripe. Si alguien consulta los síntomas de la gripe en Google es bastante probable que lo haga porque los padece él o alguien próximo, por lo que estas búsquedas permiten estudiar la propagación de la enfermedad con mayor precisión que los centros de vigilancia epidemiológica y control de enfermedades. Este proceso también es aplicable al precio de la vivienda: en lugar de analizar las tendencias de los precios en el último trimestre, es cuestión de observar la manera en las personas buscan casa en Google en distintas regiones o través de diferentes agencias inmobiliarias. La previsión resultante de estas búsquedas será más exacta que cualquier predicción basada en otros modelos económicos que emplean datos del pasado en lugar de información actual.

Este método, denominado *now casting* (predicción inmediata), permite generar previsiones muy exactas gracias a la enorme cantidad de datos que maneja. El *now casting* analiza datos de diversas fuentes para prever las tendencias económicas. Por ejemplo, si se extrapolan los datos de desempleo y de la balanza comercial, entre otros, puede obtenerse una previsión bastante precisa del Producto Interior Bruto (PIB) de este trimestre antes de que se publiquen las cifras oficiales.

Otro ejemplo de la enorme escala del *big data* es Wal-Mart, que realiza un millón de transacciones por hora, lo cual se traduce en 2,5 petabytes de información, o lo que es lo mismo, en todo el contenido de la biblioteca del Congreso de Estados Unidos reproducido 167 veces por hora. Si Wal-Mart fuera capaz de analizar toda esta información, podría optimizar su cadena de suministro y decidir cuándo y dónde debe colocar cada producto en las distintas tiendas. Este es justamente el principal problema del *big data*: hay que saber lo que se está buscando. De lo contrario, es fácil caer en la trampa de confundir la correlación con la causalidad.

La explosión de los datos es un fenómeno exponencial. Si una ciudad de 300.000 habitantes creciera al ritmo propuesto en la ley de Moore, en 20 años tendría la población actual del planeta. No solo son los ordenadores los que evolucionan a gran velocidad, sino también los algoritmos. Esta nueva masa de información generará más riqueza a medio y largo plazo, pero en estos momentos nos hallamos en el año cero del *big data*, en una especie de “salvaje oeste”, en lo que a los derechos de la propiedad intelectual y la privacidad de los datos se refiere.

NUEVAS FORMAS DE PENSAR

- En la última reunión del Foro Económico Mundial de Davos, el *big data* se presentó como un nuevo activo, como el nuevo petróleo de la economía. Pero si el *big data* es un activo para el cual desea crearse un mercado, lo primero que hay que hacer es establecer su titularidad. Para ello se requiere una legislación que defina los datos que pueden emplearse y los que pueden recabar y usar las empresas. Además, a pesar de que es necesario crear un mercado para el *big data*, la cuestión de la propiedad de la información será un tema complicado y controvertido.

El *big data* cambiará la arquitectura del mundo de los negocios del mismo modo que la dínamo eléctrica cambió el diseño de las fábricas cuando sustituyó al motor de vapor como fuente de energía. Con la dínamo,

las fábricas pasaron de ser edificios altos y estrechos a edificios bajos y horizontales con unos nuevos procesos de trabajo. El *big data* cambiará los modelos de gobernanza y los procesos, pero para ello es necesario ir más allá de los esquemas tradicionales. Es como tener un reloj analógico y querer que se convierta en uno digital: no es posible porque se trata de algo totalmente diferente. Con el *big data* pasa lo mismo: seguimos pensando en términos analógicos e intentando reproducir experiencias analógicas en el nuevo mundo digital. El *big data* requiere una nueva manera de pensar, a la vez que plantea importantes problemas como la privacidad, la seguridad y el riesgo del uso de la información.

UN CAMBIO DE MODELO

- Sony Music se enfrentó a la revolución digital con un ejército de abogados, pero la tecnología ha venido para quedarse y, por lo tanto, las cosas deben cambiar. Por ello, lo mejor sería pensar en un nuevo modelo de negocio que pudiera aportar valor, en vez de luchar por defender el modelo antiguo. El problema es que los viejos modelos se han visto amenazados tan rápido que no ha habido tiempo de generar unos nuevos. Ahí radica la tensión entre ambos. Las cosas están avanzando a una velocidad tan vertiginosa que se están desmoronando las capas de la economía sin que haya tiempo para adaptarlas. Esto es lo que está sucediendo en Estados Unidos, donde la economía se está recuperando de la crisis del 2008, pero el empleo, no.

Einstein dijo que la información no es conocimiento. Esto es cierto, ahora más que nunca, no solo por el hecho de que exista tanta información, sino porque se está convirtiendo en una “commodity”. La cuestión es cómo pueden transformarse estos datos en conocimientos y decisiones. La situación actual es opuesta a la que había hasta ahora, cuando la información era escasa y estaba filtrada por expertos. Hoy en día, el reto consiste en descartar datos para conseguir tomar buenas decisiones.

HOY EN DÍA, EL RETO CONSISTE EN DESCARTAR DATOS PARA CONSEGUIR TOMAR BUENAS DECISIONES