

Trabajo VETIC: El Big Data y la Economía Digital

Constantino Román Gómez
c.romang@alumnos.upm.es

Uno de los conceptos del momento en el mundo de las tecnologías de la información y las comunicaciones es el conocido como Big Data. Hay un consenso general en torno a la idea del Big Data como clave para el futuro y cada vez más el presente de numerosas industrias pero, ¿qué es realmente el Big Data? Como se puede deducir de su nombre, es un concepto relacionado con enormes volúmenes de datos, concretamente, con su captura y análisis, entre otros.

A lo largo de los años, la capacidad de procesamiento de los ordenadores se ha ido incrementando siguiendo la ley de Moore: aproximadamente cada dos años se duplica el número de transistores en un microprocesador, y por lo tanto su capacidad de procesamiento. Si añadimos este hecho a la disponibilidad de volúmenes cada vez mayores de datos, altas velocidades de acceso y gran variedad de datos disponibles, obtenemos las claves que explican el surgimiento de esta tecnología a gran escala.

Un proyecto de Big Data tiene una serie de fases elementales entre las que podemos identificar en primer lugar la captura y almacenamiento de los datos. Estos pueden provenir de fuentes muy diversas, desde información publicada en redes sociales y datos relacionados con nuestro comportamiento en webs y tiendas online hasta los datos biométricos almacenados por las fuerzas de seguridad o los procedentes de nuevos paradigmas como el Internet of Things, consistente en sensores que registran datos muy diversos y los comparten con otros dispositivos.

Otros pasos importantes consisten en recoger los datos en un mismo lugar y transformarlos a un formato común, para a continuación cargarlos en una base de datos. Una de las tecnologías más importantes que se aplican al Big Data son las bases de datos NoSQL, que permiten un almacenamiento más adecuado para grandes volúmenes de datos debido a su gran flexibilidad y velocidad de acceso y manipulación.

A continuación encontramos una de las fases más importantes, el análisis de los datos. Existen numerosas técnicas diferentes, entre las que destacan el data mining, el clustering y el análisis de textos. En esta fase es donde se extrae el valor de los datos, lo que permite conocer los patrones de comportamiento de clientes de una empresa, predecir variaciones en el precio de un bien o en el desarrollo de una epidemia, entre otros ejemplos.

Por último pero no menos importante, está la cuestión de cómo se visualizan los datos. Esto, que en un principio puede parecer irrelevante, es sin embargo fundamental para que se comprendan los resultados extraídos, especialmente si se trata de material destinado al gran público.

Clásicamente, se ha usado el llamado modelo de las 3 Vs de Gartner para describir el Big Data: grandes volúmenes de datos, altas velocidades y amplia variedad de información. Recientemente, algunos autores han expandido este modelo, añadiendo la variabilidad que pueden mostrar ocasionalmente los datos; la veracidad de estos, de la que depende la precisión del análisis, y la complejidad inherente a los grandes volúmenes de datos procedentes de fuentes distintas.

El valor del Big Data reside en la información que proporciona y que permite predecir comportamientos y mejorar la precisión en el proceso de toma de decisiones en las empresas, además de reducir costes y riesgos. Un ejemplo de la utilidad de los datos es el caso de Amazon. Esta tienda online usa la minería de datos de forma intensiva para mostrar anuncios personalizados a cada usuario en función de sus patrones de comportamiento a la hora de comprar y los de usuarios con patrones similares. Estos anuncios y recomendaciones personalizadas permiten dirigirse de forma más efectiva a los intereses del usuario, y en consecuencia, aumentar las ventas de Amazon.

Otro ámbito donde el Big Data ha sido aplicado de forma efectiva es el de la medicina. El análisis de datos permite detectar patrones de extensión de una enfermedad contagiosa y predecir las epidemias, proveer medicinas personalizadas por paciente o predecir los riesgos de un tratamiento o intervención, entre otros.

El tratamiento masivo de datos proporciona enormes beneficios también en el área de las manufacturas. La capacidad de predecir la demanda de productos en función de tendencias registradas anteriormente permite adaptar la producción a los momentos de pico de consumo y mejorar la eficiencia de la empresa. Además, el análisis de datos procedentes de los procesos productivos, permite mejorar la eficiencia de estos procesos y aumentar la productividad al detectar patrones en productos defectuosos o en la producción de una fábrica en función de la hora o el día, por ejemplo.

La investigación científica es otro campo que se está beneficiando enormemente de las posibilidades que ofrece el Big Data. Éste permite trabajar con gigantescas cantidades de datos. Por ejemplo, los experimentos realizados en el Gran Colisionador de Hadrones del CERN a partir de sus 150 millones de sensores generan 25 petabytes al año. Sin las tecnologías usadas en Big Data, el análisis de esa cantidad de datos sería sencillamente imposible.

Un campo donde ha existido polémica con el Big Data es el de la seguridad. La policía y otras fuerzas de seguridad tienen a su disposición gran variedad de datos sobre delitos, como áreas donde se producen, información demográfica o historial de los criminales, etc. Surge entonces la posibilidad de aprovechar los datos para intentar predecir el crimen. Sin embargo, hay un debate en este asunto: el análisis de los datos puede mostrar qué grupos demográficos y qué lugares están más relacionados con el crimen. Esta información puede ser utilizada por la policía para prestar más atención a estos grupos, pero esto puede suponer discriminación y prejuicios hacia personas que no han cometido delitos pero se encuentran dentro de dichos grupos demográficos, lo que puede generar más crimen y descontento social. Existen también críticas que razonan que dado que los crímenes que

más interés tenemos en predecir son aquellos más dañinos como terrorismo, fraude a gran escala y otros que se caracterizan por ser menos habituales, hay menos datos disponibles para su análisis y la predicción es mucho más difícil.

También se alzan voces que critican el Big Data por sus consecuencias en la privacidad. ¿Dónde empieza y dónde acaba la línea que separa el uso legítimo de los datos y las agresiones contra la privacidad? Un criterio claro a tener en cuenta sería el permiso de las personas para usar sus datos. Sin embargo, el número de aplicaciones y servicios online que usamos crece sin descanso, y podemos llegar a perder la cuenta de cuántos servicios tienen nuestros datos privados, que pueden ser objeto de un ataque informático en cualquier momento.

Además, el uso del Big Data por parte del gobierno y de agencias de inteligencia ha provocado el temor de que podamos estar avanzando hacia una sociedad vigilada, un Estado policial hasta ahora sólo visto en ficciones distópicas. Superar esta visión será un reto fundamental para el desarrollo de Big Data, buscando un equilibrio que reduzca estos temores, asegure un nivel adecuado de privacidad y garantice un uso adecuado de los datos.

Otro problema que surge con el gran desarrollo que está viviendo el Big Data en la actualidad es su posible uso inadecuado al no tener en cuenta elementos clave para el análisis como puede ser elegir muestras representativas, escoger datos útiles para su análisis y que proporcionen conclusiones precisas. Es muy importante por lo tanto tener en cuenta que el uso de enormes volúmenes de datos no es garantía de que los resultados sean relevantes.

Como hemos visto, el Big Data tiene enormes beneficios que proporcionar a numerosas industrias, así como inconvenientes que deben ser resueltos pero, ¿qué es del ámbito de la economía digital en concreto? Este término se refiere a la parte de la economía que se basa en tecnologías de computación digitales, y que se está convirtiendo en clave para el crecimiento económico. Aunque la economía digital está cada vez más entrelazada con la economía tradicional, podemos hacer algunas consideraciones respecto a esta área concreta.

La economía digital está basada principalmente en tres componentes: la infraestructura que da el soporte, el e-business y el e-commerce, además de nuevas aplicaciones que están surgiendo, como las redes sociales, y el Big Data tiene mucho valor que aportar a esta industria.

Como hemos visto anteriormente con el caso de Amazon, las aplicaciones del Big Data al comercio electrónico suponen grandes beneficios. Según un estudio realizado por IBM y Accenture, las empresas que usan el análisis predictivo de datos consiguen ventas un 73% mayores que aquellas que no lo hacen. Como se ha señalado previamente, las tiendas online pueden usar datos sobre patrones de comportamiento de usuarios para predecir en qué productos pueden estar interesados y mostrarlos como productos recomendados. En el caso de Amazon, otra aplicación que le da al Big Data consiste en la monitorización y

seguimiento de sus productos a lo largo de sus centenares de almacenes. El análisis de datos a gran escala también ha sido usado para optimizar los precios de los productos en función de la fecha, para proporcionar una atención al cliente personalizada o para detectar el fraude, por ejemplo.

En otros negocios digitales innovadores del ámbito del entretenimiento también se pueden apreciar las ventajas del Big Data. Por ejemplo, Netflix, que proporciona un servicio de streaming de películas y series online, lo ha usado de cara a sus producciones propias. Cuando la empresa decidió crear una versión estadounidense de la serie británica House of Cards, que estaría dirigida por David Fincher y protagonizada por Kevin Spacey, estudiaron datos acerca de sus usuarios. Gracias a este análisis, Netflix pudo saber que muchos usuarios habían visto otras películas de David Fincher de principio a fin, que la serie británica fue muy vista y sobre todo, que los usuarios que habían visto la serie también veían películas de Kevin Spacey y de David Fincher. Estos datos tuvieron gran importancia a la hora de tomar la decisión de invertir 100 millones de dólares en la serie, que fue finalmente un gran éxito.

Además, en el área de las infraestructuras que dan soporte a la economía digital también hay importantes aplicaciones. Los operadores de telecomunicaciones pueden obtener fácilmente enormes cantidades de datos de sus usuarios, que pueden analizar para numerosos propósitos. Entre ellos, encontramos usos que hemos visto en otras industrias, como en el área del marketing o en la de la seguridad, para evitar el fraude, por ejemplo. Sin embargo, también existen otras interesantes aplicaciones, como la monitorización masiva de los sistemas de la red para averiguar las causas de las caídas de red o el análisis de los datos de desempeño de la red, como tasas de pérdida de paquetes o de retardo, para conocer los elementos de la red que necesitan cambios o mejoras. Estas aplicaciones permiten a los operadores mejorar la calidad de sus servicios e incrementar su productividad.

Como hemos visto a lo largo de este ensayo, el Big Data es una tecnología que se está abriendo paso en los últimos tiempos y que tiene el potencial para introducir grandes cambios en numerosas industrias. Cada año que pasa tiene más impacto sobre la economía digital en todos sus aspectos: e-commerce, e-business, infraestructuras de soporte, redes sociales, etc. Introducir el análisis masivo de datos tiene indudablemente enormes beneficios para la industria digital, aumentando la productividad, eficiencia y calidad del servicio, entre muchas otras ventajas.

Sin embargo, como también se ha expuesto anteriormente, la introducción del Big Data trae consigo algunos posibles inconvenientes. Es fundamental superarlos y ofrecer soluciones y garantías en sus áreas más polémicas. Para ello es necesario un esfuerzo conjunto de los gobiernos, empresas y operadores para el adecuado desarrollo del Big Data. Se necesita por parte de los gobiernos una regulación que proteja los derechos de los ciudadanos y su privacidad sin limitar los beneficios del Big Data y que sienta las bases para la innovación en el sector. Es importante también que las empresas y los operadores de la red ofrezcan garantías del cumplimiento de estas normas para despejar posibles dudas, además de extender las buenas prácticas en el Big Data, usando la metodología adecuada para

obtener resultados relevantes que mejoren la vida de las personas e impulsen el crecimiento de la economía.

Como todas las tecnologías innovadoras que han supuesto una disrupción en las industrias, la implantación del Big Data en la economía trae consigo grandes retos en cuanto a su adecuada aplicación y desarrollo. Sin embargo, no es posible ignorar los enormes beneficios potenciales de esta tecnología, de ahí la importancia de sentar las bases adecuadas para el futuro. El Big Data es imparable y es responsabilidad de los principales actores (gobiernos, empresas, operadores) asegurar su correcta implantación.

Como algunos autores han expresado, los datos son el combustible de la economía digital, clave para el crecimiento económico. El futuro de la economía digital está incuestionablemente ligado al del Big Data, así como a nuevas tecnologías innovadoras como el Internet of Things. En los próximos años surgirán aplicaciones del Big Data que probablemente no podamos imaginar en la actualidad y que transformarán esta y otras industrias. Nuestro futuro pasa ineludiblemente por el Big Data.