

# La Minería de Datos: ¿Sueño o Realidad?

Manuel Barberena y Viterbo Berberena

Pearson

Comencemos por definir el concepto de Minería de Datos.

**La Minería de Datos** —extracción de patrones y reglas significativas de una gran cantidad de información— es obviamente útil en cualquier campo, donde exista una gran cantidad de datos y algo valioso que aprender. No nos podríamos sorprender de enterarnos, por ejemplo, que las organizaciones de inteligencia militar usan las técnicas de minería de datos para procesar grandes cantidades de imágenes satelitales con el propósito de clasificar las cosas en la tierra como tanques o tractores.

En el contexto de negocios, la misma regla se aplica: la minería de datos es útil donde sea que existan grandes cantidades de datos y algo valioso que conocer o aprender. En los negocios, hay una definición explícita de qué es lo que se considera valioso conocer, es decir, “algo se considera **valioso** si el resultado de su conocimiento aporta más dinero que lo que costó su descubrimiento”. En realidad la definición es aún más estricta que el enunciado anterior, en otras palabras, “**el conocimiento de algo es valioso si el retorno de la inversión requerida para su aprendizaje es mayor que el retorno de utilizar estos fondos en otro tipo de inversión**”.

En la academia, se considera como “**conocimiento**”, tener los valores intrínsecos de algo, independientemente de cualquier aplicación práctica. En el contexto de negocios, sin embargo, el conocimiento puede ser valioso de dos formas: si aumenta la ganancia por medio de la disminución de los costos o por incremento de los ingresos.

En verdad hay una tercera forma en la que el conocimiento puede tener valor, y es que este puede **incrementar el precio en bodega** por medio de la promesa de un incremento futuro de la ganancia por vía de alguno de los dos mecanismos anteriores.

En este momento hay cuatro formas fundamentales en las que se está empleando la minería de datos en los negocios:

- ◆ Como herramienta de investigación.
- ◆ En el proceso de mejora del negocio.
- ◆ En la mercadotecnia.
- ◆ Para la administración de las relaciones con los clientes.

## La Minería de Datos como Herramienta de Investigación

Una vía en la que la minería de datos puede disminuir los costos es en el mismo inicio del ciclo de vida del producto, durante la etapa de investigación y desarrollo. La industria farmacéutica es un buen ejemplo. Esta industria se caracteriza por muy altos costos de investigación y desarrollo. A nivel global esta industria dedica 13 mil millones de dólares anuales a la investigación y desarrollo de drogas utilizadas en la medicina. Bajo el riesgo de una simplificación grosera de un negocio muy complejo, el proceso de desarrollo de una droga es semejante a un largo embudo con millones de productos químicos entrando a la parte ancha y solamente una pequeña cantidad de drogas útiles y seguras saliendo por el caño estrecho del mismo.

La droga potencial, debe primeramente estar enlazada con la molécula activa correcta. La que cumple esto se llaman iniciador. A continuación, el iniciador tiene que demostrar que posee el efecto deseado en un tubo de ensayo, las que no se descargan. Después el iniciador debe ser capaz de ser absorbido por un cuerpo y sobrevivir en el ambiente complejo y hostil de un organismo vivo. Las que no se descargan. Luego el iniciador debe ser no tóxico y tener algún efecto útil en los ensayos con animales. Los que son tóxicos y/o no útiles se descargan. Finalmente el iniciador debe pasar a través de una serie de ensayos clínicos para probar que es seguro y efectivo en el cuerpo humano. Una pequeña cantidad de moléculas candidatas pasan todo este proceso y se convierten en drogas (medicamentos) —alrededor de 1 en 10 mil— y el costo promedio es de 300 millones de dólares.

Hay desde luego abundancia de datos en los cuales se pueden basar las predicciones. Los laboratorios farmacéuticos modernos emplean una técnica llamada HTS para seleccionar los candidatos a drogas. Sistemas automáticos desarrollan la química combinatoria para crear una amplia variedad de moléculas orgánicas a partir de un pequeño conjunto de reactivos conocidos. Estas moléculas son entonces tamizadas por exóticos robots que manipulan platos especiales que contienen docenas de cavidades que contienen soluciones con la molécula activa receptora. Estos robots reconocen cuando una molécula está enlazada con el receptor. Las moléculas que están enlazadas químicamente con el receptor, pero que no son similares a éste, son probablemente iniciadores.

Todos esos ensayos automatizados generan datos que son perfectos para la minería de datos –muchas variables de entrada y una simple variable de salida de sí/no (sirve/no sirve)–.

Las empresas farmacéuticas usan técnicas de predicción sofisticadas para determinar qué productos químicos son probables o muy probables para producir una droga. Si se pudieran enfocar en la investigación de los productos químicos apropiados (reactivos) o simplemente no investigar a los muy poco probables, estas compañías pudieran ahorrar millones y millones de dólares.

Una disciplina completa, la bioinformática, ha crecido para hacer minería de datos e interpretar los datos generados por la técnica de HTS y otras fronteras de la biología como el mapa completo de secuencia genética humana.

### **La Minería de Datos en el Proceso de Mejoramiento del Negocio**

Otra área donde la minería de datos puede ser rentablemente empleada para ahorrar dinero es en la empresa manufacturera. Muchos de los procesos modernos de manufactura, desde la fabricación de un chip hasta la industria cervecera se controlan usando procesos de control estadístico. Sensores mantienen la traza de la presión, temperatura, velocidad, color, humedad y cualquier otra variable que sea apropiada al proceso. Los programas de computación vigilan las salidas de estos sensores y ordenan pequeños ajustes para mantener las lecturas dentro de los límites de control. Pero, ¿qué son los

límites de control? Hay tantas interacciones complejas entre las variables que aun operadores humanos expertos frecuentemente no están seguros.

La consecuencia del problema en el proceso de manufactura es a veces la ruina de un lote de productos y la muy costosa parada y reinicio del proceso. Una vez más, los datos producidos por los sistemas automatizados de manufactura son perfectos para la minería de datos: un enorme volumen de entradas, consistentes en medidas precisas de los valores de los resultados de cientos de variables y pequeñas y simples variables de salida como bueno/malo.

### **La Minería de Datos en la Mercadotecnia**

Muchas de las más exitosas aplicaciones de la minería de datos son en el área de la mercadotecnia, especialmente en la parte conocida como mercadotecnia de bases de datos (database marketing). En este mundo, la minería de datos se usa en el término de costo y el de ingresos de la ecuación de beneficio. En la mercadotecnia de bases de datos, las bases de datos se refieren a una colección de datos sobre los objetivos posibles de la campaña de mercadotecnia. En dependencia de la situación, estos datos pueden ser datos del comportamiento en detalle de los clientes existentes, recogidos por los sistemas de operaciones de la compañía, tales como los sistemas de pedidos, pagos, puntos de ventas, etcétera; o puede ser información rudimentaria del tipo de, pago de cuentas, cuotas u honorarios, que está disponible en cada hogar de los E.U., y en menor extensión en otros países.

La minería de datos se usa para disminuir los costos de la mercadotecnia por medio de la eliminación de las llamadas y las cartas a gente, que son improbables que respondan a una oferta. Por ejemplo, la AARP (American Association of Retired People) ahorra millones de dólares excluyendo al 10% de los hogares elegibles que fueron considerados menos probables que se conviertan en miembros. Por el lado de los ingresos, la minería de datos se puede utilizar para determinar los prospectos (candidatos) más valiosos, tales como los que comprarían las coberturas de seguros más amplias o las más caras. A partir de la experiencia acumulada en la actualidad en esta área, se han publicado muchos ejemplos del uso de la minería de datos para mejorar los objetivos de las campañas de mercadotecnia.



## La Minería de Datos en la Administración de las Relaciones con los Clientes

En la actualidad, la frase “Administración de las Relaciones con los Clientes” (del inglés Customer Relationship Management), parece estar en los labios de cada jefe ejecutivo y consultor de la administración. Este término ha venido a englobar mucho de lo que se ha estado llamando *one-to-one marketing*, adelantándose con las ideas acerca de las fuerzas de ventas automatizadas y la personalización.

Una buena administración de las relaciones con los clientes significa:

Presentar una sola imagen de la compañía a lo largo de todos los muchos canales que el cliente puede usar para interactuar con la misma.

Mantener una sola imagen del cliente que es compartida en toda la empresa.

Una buena administración de las relaciones con los clientes requiere del entendimiento de quiénes son los clientes y qué es lo que prefieren. Esto significa

anticiparse a sus necesidades y abordarlas proactivamente. Es decir, reconocer cuándo los clientes están insatisfechos y hacer algo sobre esto antes de que se harten y se pasen a la competencia.

La minería de datos juega un rol destacado en cada faceta de la administración de las relaciones con los clientes. Solamente a través de la aplicación de las técnicas de minería de datos puede una gran empresa esperar cambiar los innumerables registros de sus bases de datos de consumidores en alguna clase de imagen coherente de sus clientes.

La minería de datos permite a la corporación aprender de las observaciones del comportamiento de sus clientes, que están en el almacén de datos (*data warehouse*) de información de clientes, que sirve como memoria corporativa.

En relación al ciclo de vida de los clientes (desde antes que los prospectos se conviertan en clientes hasta después que abandonan la firma), la minería de datos se aplica para mejorar la administración de las relaciones con los clientes en cada punto a lo largo de este camino.

野生のリスも時折姿を見せ見せる  
広大な日本庭園に囲まれた美しい  
直線美をいかした外観した外観  
アメリカの著名な建築家ミル・ヤ  
内装と庭園の設計はの設計は  
元芸術院会員故村野藤吾氏の  
日本の伝統  
「都ホテル」  
くろぎとや  
その緑に美  
野生のリス  
広大な日本  
直線美をい  
アメリカの著名な建築家ミル・ヤ  
内装と庭園の設計はの設計は  
元芸術院会員故村野藤吾氏の  
日本の伝統美を織りまぜた 落着き  
「都ホテル東京」はみなさまさまに  
くろぎとやすらぎをお約束いたしま

¿Estos fueron los resultados que le entregaron en su último estudio cualitativo?

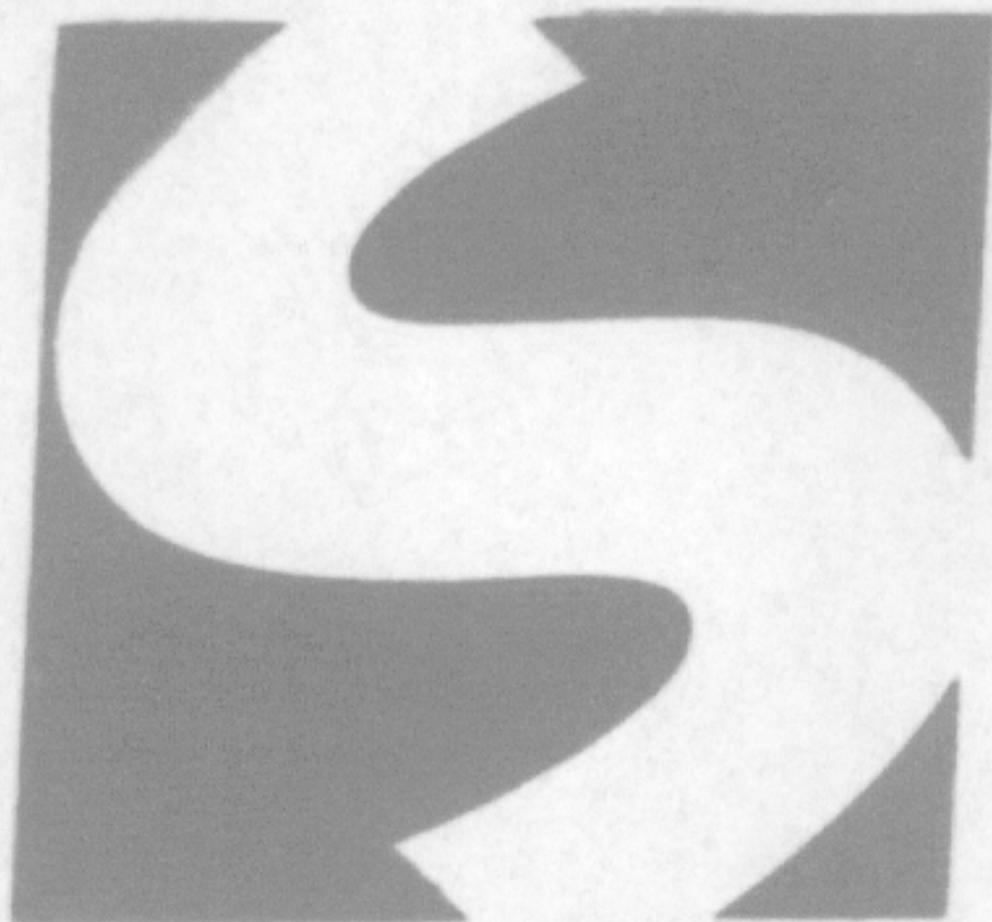
「都ホテル東  
と見せ見せる  
れた美しい  
観した外観は  
「都ホテル東  
なさまさまに  
約束いたしま  
れた 落着き  
なさまさまに  
約束いたしま

Liámenos

**iq** innerquest

Newton No. 285 - Col. Chapultepec Morales • C.P. 11670 • México D.F. • Tels. 52535211 • 52535212 • 52542856 • Fax. 55510137 • [innerquest@net-mex.com](mailto:innerquest@net-mex.com)

# EXPERIENCIA Y ATENCIÓN PERSONALIZADA



## **SERTA de México** INFORMACIÓN ESTRATÉGICA DE MERCADOS

[www.serta.com.mx](http://www.serta.com.mx)

**DIVISIÓN CUANTITATIVOS**  
Cruz de Valle Verde 16-2,  
Santa Cruz del Monte,  
Naucalpan Estado de México

**DIVISIÓN CUALITATIVOS**  
Vía Láctea No.10 PB esq. Av.Lomas  
Verdes, Jardines de Satélite,  
Naucalpan Estado de México

Tel. conmutador (5)5613264

Fax: ext. 102

[obalcazar@serta.com.mx](mailto:obalcazar@serta.com.mx)