

Minería de datos

Como ya sabemos estamos viviendo una época en la que la información es la clave para obtener el éxito en el mundo de los negocios, para que una empresa se mantenga constante en el aspecto competitivo, los gerentes y los tomadores de decisiones necesitan un acceso fácil y rápido a información útil y valiosa que tiene la empresa y para esto les sería muy útil una base de datos.

No puede haber algo más crítico para una empresa que tener tanta información disponible y no saber que hacer con esa información, ya que con la misma puede obtener muchos beneficios, por ejemplo con esa información las empresas pueden crear pronósticos y/o reportes que apoyen a la toma de decisiones y todo esto a su vez lo podríamos considerar como una ventaja competitiva.

Dicha información se puede concentrar en una base de datos que posteriormente puede ser vendida como conocimiento empresarial. En la actualidad existe una gran variedad de software especializado en bases de datos, la mayoría tienen características similares entre ellas, pero pueden ser usadas para diferentes fines de las empresas, ya que pueden ser utilizadas para obtener beneficios, en su área de ventas, marketing, finanzas, etc.

Algunas de las metas que tienen las bases de datos son, por ejemplo:

1. Procesar automáticamente grandes cantidades de datos crudos.
2. Identificar los patrones más significativos y relevantes.

3. Presentarlos como conocimiento apropiado para satisfacer las metas del usuario

Las bases de datos pueden ser utilizadas para hacer un análisis de la cesta de la compra mediante reglas de asociación, Modelos para análisis de riesgos (seguros, Créditos...), evaluación de campañas publicitarias, análisis de la fidelidad de clientes, análisis de valores de bolsa, detección y prevención de fraude en comercio electrónico, modelos de tráfico a partir de datos, perfiles de usuarios de redes, detección de intrusos en redes, etc.

Actualmente, la minería de datos es una actividad en expansión aplicada cada vez en más disciplinas, que han visto la utilidad del estudio de datos para apoyar la toma de decisiones. Especial énfasis está teniendo en aquellas áreas relacionadas con la economía y dentro del campo de inteligencia de los negocios

Uno de los elementos problemáticos de la minería de datos está en cómo se invade la privacidad de los usuarios. Esta situación hace que muchos usuarios puedan recelar del hecho de ceder sus datos (propios o generados) para su explotación y beneficio de las empresas.

En el siguiente artículo hay información sobre la minería de datos

Empresa Navarra

"La minería de datos permite a las empresas explicar el pasado, entender el presente y predecir resultados de futuro"

(18/05/2005) - *"El volumen y variedad de información que se encuentra informatizada en bases de datos ha crecido espectacularmente en las últimas décadas. Para que esta información nos pueda ser útil es necesario servirnos de la minería de datos, un novedoso proceso que permite extraer conocimiento útil a partir de grandes cantidades de datos almacenados en distintos formatos".* "Esta técnica, que se puede aplicar en cualquier ámbito, permite a

las empresas disponer de información muy valiosa para tomar decisiones y proyectarse en el futuro".

De esta manera se expresaron en Pamplona el profesor César Ferri Ramírez, del Departamento de Sistemas Informáticos y Computación de la Universidad Politécnica de Valencia, que impartió en la Universidad Pública de Navarra una conferencia sobre "Técnicas del aprendizaje automático para la asistencia en la toma de decisiones", organizada por el Departamento de Automática y Computación de la universidad navarra.

Según explica, la mayoría de decisiones de empresas, organizaciones e instituciones se basan en la información almacenada sobre experiencias pasadas. "En muchas situaciones, el método tradicional de convertir esos datos en conocimiento consiste en un análisis e interpretación que se realiza de modo manual". Una forma de actuar que, en opinión de César Ferri, es "lenta, cara, muy subjetiva e impracticable en dominios donde el volumen de datos es alto".

Sin embargo, existe un proceso relativamente moderno llamado minería de datos que engloba varias técnicas de aprendizaje automático y que permite extraer conocimiento de esos datos. "Se trata de un área muy amplia que bebe de muchas fuentes como las bases de datos, la inteligencia artificial, la estadística y las matemáticas, y lo que hace es sacar a la luz, rastrear y comparar información almacenada de bases de datos de gran tamaño. En definitiva, se trata de extraer y descubrir información que existe en las bases de datos pero que es desconocida para el usuario".

Para ello, se sirve de varias técnicas que permiten recopilar los datos, seleccionarlos, corregir los que son erróneos y completarlos. Posteriormente se combinan los patrones y relaciones encontrados con información complementaria y, finalmente, se interpretan estos datos para que el cliente disponga de la información que necesita y que le ayudará a adoptar decisiones de cara al futuro.

Aplicaciones en todos los campos

Mayoritariamente, las técnicas de minería de datos se emplean para mejorar el rendimiento de procesos de negocio o industriales en los que se manejan enormes volúmenes de información almacenada en bases de datos, pero su aplicación es viable en todos los campos.

Así, por ejemplo, se utiliza con éxito en aplicaciones de control de procesos productivos; finanzas y banca; análisis de mercado, distribución y comercio; seguros y salud privada; agencias de viaje; marketing; propaganda; medicina; educación; procesos industriales; telecomunicaciones; y en biología, bioingeniería y otras ciencias. Además, es fundamental en la investigación científica y técnica porque permite analizar y descubrir conocimiento a partir de datos de observación o de resultados experimentales.

Como ejemplo, el profesor de la Universidad Politécnica de Valencia indica que, a través de esta técnica, “los bancos obtienen información que les permite saber qué productos deben ofertar a sus clientes, evitar utilidades fraudulentas de las tarjetas de crédito o prever de antemano la fiabilidad de un nuevo cliente.

A las compañías de seguros les ayuda a detectar fraudes; en los supermercados posibilita conocer el comportamiento de los clientes; y en medicina se aplica en los diagnósticos y en la gestión de los hospitales. En Estados Unidos, por ejemplo, también se está utilizando para la detección de mails o llamadas telefónicas sospechosas”.

Regulación legal e implantación

César Ferri reconoce que trabajar con este tipo de información puede plantear problemas éticos y legales. “A veces, los datos con los que se trabaja contienen información privada, por lo que hay que ser muy respetuoso y eliminar aquellas informaciones que no aportan nada y que, además, identifican a la persona. Así por ejemplo, en medicina se eliminan los nombres de los pacientes, lo cual no influye en las conclusiones”. Aunque no existe una ley específica que regule esta técnica, la minería de datos debe cumplir la Ley de Protección de Datos que, en su opinión, es “bastante restrictiva”.

Por otro lado, explica que esta técnica se utiliza con éxito en Estados Unidos desde hace 10 años, pero en Europa su implantación todavía no se ha generalizado. Por su parte, en España está siendo aplicada por grandes empresas de telefonía, bancos, grandes supermercados, compañías de seguros y hospitales.

La razón principal de esta lenta implantación es que “las empresas desconocen lo que les puede aportar. Además, se trata de un sistema caro, exige disponer de un alto volumen de datos para que se puedan predecir resultados de futuro, y los resultados se obtienen a medio plazo”. No obstante, “resulta barato si se compara con el beneficio que puede reportar”. Otro de los problemas es que “existe poca gente formada en este campo, aunque ya se están empezando a impartir estas técnicas en las Universidades españolas”.

Bibliografía:

- http://www.cajanavarra.es/PortalCAN/es-ES/CanalEmpresas/EmpresaNavarra/Noticias/_OtrosAnnos/2005/tratamientoddatos.htm