

Definición de Minería de Datos

Se puede definir la Minería de Datos como *El proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos*. Es decir, la tarea fundamental de la Minería de Datos es encontrar modelos inteligibles a partir de los datos.

Para que este proceso sea efectivo, debería ser automático o semiautomático y el uso de los patrones descubiertos debería ayudar a la toma de decisiones, y por tanto, un beneficio a la organización.

Entonces se puede deducir que la Minería de Datos posee dos retos bien diferenciados:

1. Trabajar con grandes volúmenes de datos
2. Usar las técnicas adecuadas para analizar dichos datos y extraer el conocimiento novedoso

De esta forma se puede decir que el objetivo de la Minería de Datos es convertir datos en conocimiento

Así, las ventajas que podemos encontrar en la Minería de Datos son las siguientes:

1. A largo plazo, ahorra dinero a la empresa
2. Contribuye a la toma de decisiones de forma estratégica
3. Mide los resultados en la forma de mejora
4. Genera modelos descriptivos, es decir, qué datos influyen en los resultados finales
5. Genera modelos predictivos

Sin embargo, los inconvenientes que podemos encontrar en la Minería de Datos son las siguientes:

1. En un principio, la implantación de los productos es costosa
2. Requiere mucha experiencia en el uso de herramientas tecnológicas

Técnicas de Minería de Datos

Las Técnicas de Minería de Datos son las siguientes:

1. **Análisis de datos:** lo primero que se realiza es una consulta a los datos para conocer la forma en que éstos se encuentran antes de aplicar las Técnicas de Minería de Datos
2. **Técnicas de visualización:** se realizan para saber la calidad de los datos
3. **Redes neuronales artificiales:** son modelos predecibles, no lineales
4. **Reglas de asociación:** se realizan en base al perfil de los clientes que van a utilizar la Minería de Datos
5. **Algoritmos genéticos:** son técnicas de optimización de Minería de Datos
6. **Redes Bayesianas:** se dan para conocer las relaciones casuales que explican un fenómeno y se realizan en base a los datos de la base de datos
7. **Árbol de decisión:** a través de su estructura aparecen las reglas de clasificación de un conjunto de datos

