

# Oracle Data Mining

*La minería de datos (Data Mining) ayuda a las organizaciones a encontrar información que no es perceptible de forma directa, como por ejemplo patrones de comportamiento, relaciones, asociaciones, etc... que nos permitan tomar mejores decisiones. A través del análisis del pasado, y aplicando algoritmos, se construyen predicciones que nos permiten mejorar nuestra eficiencia y conseguir así una mayor rentabilidad de la actividad de negocio.*



A través de este proceso de minería se consigue:

- Sacar perfiles de los clientes y entender su comportamiento
- Fidelizar a los clientes ofreciendo lo que ellos esperan
- Mejorar los beneficios y márgenes
- Aumentar nuestra eficacia incrementando nuestra competitividad

Las herramientas de consulta y generación de informes nos permiten sacar información de nuestra base de datos o datawarehouse y son buenas para responder a preguntas tales como: ¿quién ha comprado el producto "A" durante los últimos 6 meses?

Las herramientas OLAP van más allá y permiten de forma interactiva profundizar en detalle en los datos y hacer resúmenes, comparaciones y análisis. Las herramientas OLAP son buenas para responder a preguntas tales como: ¿cuál es el salario medio de los clientes que compraron el producto "A", clasificados por año y región geográfica?

Las herramientas de consulta y OLAP son buenas para analizar el pasado; si queremos llegar más allá necesitamos la minería de datos.

La minería de datos analiza la información en profundidad, a través de algoritmos de aprendizaje, recorre los datos para descubrir patrones e información encubiertos, y basándose en estas conclusiones construye predicciones de futuro. La minería de datos es buena para responder a preguntas tales como: ¿quién es posible que compre el producto A en los próximos 6 meses y porqué?

## Oracle Data Mining

La base de datos Oracle incluye funcionalidad para la minería de datos en la edición Enterprise. Esta funcionalidad está totalmente integrada y bajo el mismo motor que la parte relacional de la misma. Se puede acceder a toda la funcionalidad Data Mining a través de la API Java que incluye la base de datos, de manera que las aplicaciones puedan sacar el máximo partido de las funciones disponibles.

Al estar integrado en la base de datos, Oracle Data Mining simplifica el proceso de extracción de conclusiones basadas en grandes cantidades de datos, ya que se elimina la necesidad de movimientos de datos para el proceso de análisis. Todas las operaciones de preparación, creación de modelos y análisis permanecen en la base de datos lo que resulta en una mejora de la productividad, automatización e integración.

Oracle Data Mining acepta tablas transaccionales y no transaccionales (resúmenes, registros únicos). Oracle Data Mining hace todas las transformaciones necesarias automáticamente de forma interna, liberando así de este trabajo a los usuarios o desarrolladores.

Oracle Data Mining soporta la clasificación de valores dentro de un campo en grupos que tengan sentido, por ejemplo, el campo edad puede ser clasificado en rangos como: 0-16,17-21,etc...

## Oracle Data Mining



### Algoritmos Oracle Data Mining

Oracle proporciona dos algoritmos para dos tipos diferentes de análisis:

- 1) *Naive Bayes para clasificaciones y predicciones*
- 2) *Reglas de asociación para encontrar patrones.*

#### Naive Bayes:

Naive Bayes es una técnica de clasificación y predicción que construye modelos que predice la probabilidad de posibles resultados. Naive Bayes utiliza datos históricos para encontrar asociaciones y relaciones y hacer predicciones.

Este algoritmo predice resultados binarios o multiclase. En los problemas binarios, cada registro cumplirá o no el comportamiento modelado. Por ejemplo, se puede construir un modelo para averiguar si un cliente será fiel o cambiará de proveedor.

Naive Bayes puede hacer predicciones para problemas multiclase, en los cuales hay varios resultados posibles. Por ejemplo, se puede construir un modelo para predecir qué clase de servicio prefiere cada cliente.

#### Ejemplo de modelo binario:

- ¿Es probable que el cliente sea un cliente altamente rentable en el futuro?
- Si, con un 85% de probabilidad

#### Ejemplo de un modelo multiclase:

- ¿En cuál de los tipos de cliente (en crecimiento, estable, en decrecimiento) encaja mejor este cliente?
- Crecimiento, con el 55% de probabilidad

Las predicciones binarias y multiclase conjuntamente cubren un gran rango de problemas de negocio, incluyendo respuesta a campañas, ofertas de ventas, detección de fraude, predicción de rentabilidad, perfilado de clientes, etc...

#### Reglas de Asociación:

Las reglas de asociación detectan eventos asociados que se ocultan en las bases de datos. Este tipo de análisis a menudo se utiliza para encontrar combinaciones populares de productos, tales como cereales y leche asociados con plátanos. Las reglas de asociación generan un conjunto de pares A-B con una probabilidad n%.

Una vez creados los modelos Naive Bayes, los registros de datos pueden ser puntuados. La puntuación es el proceso de predicción de resultados, y puede hacerse en modo batch o bajo demanda. En modo batch el algoritmo recorre una tabla y va almacenando las predicciones en otra tabla, bajo demanda el algoritmo puntúa un solo registro y devuelve la predicción, que puede utilizarse directamente en la aplicación que haya pedido esta puntuación.

#### Para más información:

[http://www.oracle.com/solutions/business\\_intelligence](http://www.oracle.com/solutions/business_intelligence)  
**900 952 900**

ORACLE®