

CONCEPTOS A INVESTIGAR	OLAP	MINERÍA DE DATOS
CONCEPTO	<p>* Es el procesamiento analítico en línea . Es una solución utilizada en el campo de la llamada Inteligencia de negocios cuyo objetivo es agilizar la consulta de grandes cantidades de datos.</p> <p>Es el tratamiento informático que permite a los usuarios extraer fácilmente y de forma selectiva y ver <u>los datos</u> de diferentes puntos de vista.</p> <p>Es un método para buscar en los datos de diferentes maneras. Con OLAP los datos son clasificados en diferentes dimensiones las que pueden ser vistas unas con otras en cualquier combinación para obtener diferentes análisis de los datos que contienen.</p> <p>OLAP es el acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Es una solución utilizada en el campo de la llamada inteligencia empresarial (o Business Intelligence) cuyo objetivo es agilizar la consulta de grandes cantidades de datos.</p> <p>Para ello utiliza estructuras multidimensionales (o</p>	<p>Es el análisis de archivos y bitácoras de transacciones, trabaja a nivel del conocimiento con el fin de descubrir patrones, relaciones, reglas, asociaciones o incluso excepciones útiles para la toma de decisiones.</p> <p>Realiza dos operaciones básicas:</p> <ul style="list-style-type: none"> • Predecir tendencias y comportamientos. • Identificación de patrones desconocidos con anterioridad. Las aplicaciones normales de BI normalmente proporcionan a los usuarios una visión de lo que ha sucedido, la minería de datos ayuda a explicar qué está sucediendo y predice lo que sucederá en el futuro. <p>Proceso que tiene como propósito descubrir, extraer y almacenar información relevante de amplias bases de datos, a través de programas de búsqueda e identificación de patrones.</p>

	<p>cubos OLAP) que contienen datos resumidos de grandes bases de datos o Sistemas Transaccionales (OLTP). Se usa en informes de negocios de ventas, marketing, informes de dirección, minería de datos y áreas similares. Es una solución utilizada en el campo de la llamada <u>Inteligencia de negocios</u> cuyo objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras multidimensionales (o <u>Cubos OLAP</u>) que contienen datos resumidos de grandes <u>Bases de datos</u> o Sistemas Transaccionales (OLTP). Análisis multidimensional. permite a los usuarios la visualización de los datos de diferentes formas utilizando dimensiones múltiples. Se apoya en cubos multidimensionales o base de datos multidimensionales donde cada aspectos representa una dimensión.</p>	
CARACTERÍSTICAS	<p>*Capacidad de análisis: Para que los usuarios sean capaces de explotar el potencial de la información en forma efectiva</p> <p>*Accesibilidad: es de gran valor para los niveles ejecutivos y</p>	<p>*Explorar los datos se encuentran en las profundidades de las bases de datos, que son almacenados por varios años.</p> <p>* descubrimiento de resultados valiosos e inesperados</p>

gerenciales debido a la importancia que reviste el contar con información oportuna y facilidades para su análisis en el momento de la toma de decisiones.

***Rapidez**.- se apoya de cubos multidimensionales, donde cada aspecto representa una dimensión.

*Los Cubos OLAP son vectores en los cuales se dispone la información, y gracias a esta ordenada jerarquía es posible llevar a cabo un análisis rápido de los datos.

*Manera más efectiva y tecnológica más avanzada para integrar, transformar y combinar los datos para facilitar al usuario o a otros sistemas el análisis de la información.

*La tecnología OLAP generalmente se asocia a los almacenes de datos, aunque se puede tener almacenes de datos sin OLAP y viceversa.

*Habitualmente se utilizan herramientas OLAP (On-line Analytical Processing) como herramientas frontales para el acceso a los datos.

*Las herramientas OLAP, como los almacenes de datos y

*La minería de datos produce cinco tipos de información:

- Asociaciones.
- Secuencias.
- Clasificaciones.
- Agrupamientos.
- Pronósticos.

-Método para buscar en los datos de diferentes maneras.

- predecir tendencias y comportamientos

- Explorar los datos se encuentran en las profundidades de las bases de datos, como los almacenes de datos, que algunas veces contienen información almacenada durante varios años.
- En algunos casos, los datos se consolidan en un almacén de datos y en mercados de datos; en otros, se mantienen en servidores de Internet e Intranet. El entorno de la minería de datos suele tener una arquitectura cliente/servidor.

bases de datos multidimensionales, están basadas en el modelo multidimensional.

*Se destacan la utilización de avanzadas interfaces gráficas, capaces de presentar la información en forma numérica o gráfica y de permitir el diálogo con el usuario directamente sobre la información que es visualizada.

*Son entornos especialmente diseñados para la ejecución de análisis multidimensionales de los datos corporativos, que soportan amigablemente los análisis de cualquier usuario así como las posibilidades de navegación, seleccionando la información a obtener, permitiendo el análisis de datos segmentados y que permiten ir reduciendo el conjunto de datos reportados.

***Capacidad de análisis.**

*Para que los usuarios sean capaces de explotar el potencial de la información en forma efectiva es necesario brindarles una buena combinación de facilidades avanzadas para la manipulación y presentación de datos.

VENTAJAS	<ul style="list-style-type: none"> *Por su ordenada jerarquía es posible llevar a cabo un análisis rápido de los datos. *procesamiento de importantes volúmenes de información. *mayor rapidez en las consultas y en el procesamiento de la información, lo cierto es que su gran falla reside en la imposibilidad de realizar cambios en su estructura. * Facilidad de uso. el usuario puede consultarla en cualquier momento. Proporcionan a las compañías un sistema confiable para procesar datos que luego serán utilizados para llevar a cabo análisis e informes que permitan mejorar las operaciones productivas, tomar decisiones inteligentes y optimizar la competitividad en el mercado. *Mayor productividad de los responsables de la toma de decisiones. 	<p>Auxilia a los usuarios empresarios en el procesamiento de reserva de datos.</p> <p>Ayuda a elegir cursos de acción y definir estrategias competitivas.</p> <p>Mediante modelos avanzados y reglas de inducción, y puedan examinar una gran cantidad de datos.</p> <p>Hace uso de programas de búsqueda para detectar desviaciones, tendencias y patrones ocultos en los datos históricos.</p> <p>Se combinan fácilmente y pueden analizarse y procesarse rápidamente.</p>
DESVENTAJAS	<p>Su gran falla reside en la imposibilidad de realizar cambios en su estructura.</p>	<p>Dificultad de recopilación de los datos</p> <p>Dependiendo del tipo de datos que se quieran recopilar puede conllevar mucho trabajo o la necesidad de tecnología de elevado coste.</p>

	<p>forma de funcionamiento y almacenamiento de la información</p> <p>Tamaño limitado por la arquitectura del cubo.</p> <p>Requiere conocimientos y experiencias elevados para crear y gestionar las herramientas.</p>	<p>El preprocessamiento de datos puede llevar demasiado tiempo</p> <p>No está asegurada la obtención de un modelo válido</p> <p>Dificultad de recopilación de los datos.</p>
INSTITUCIONES PÚBLICAS, SOCIALES, EDUCACIÓN	<ul style="list-style-type: none"> *Se usa en informes de negocios de ventas. *Marketing. *Informes de dirección. *<u>Minería de datos</u> y áreas similares. 	<ul style="list-style-type: none"> *Algoritmos genéticos. *Inteligencia Artificial. *Sistemas Expertos. *Redes neuronales. *Análisis de varianza. *Análisis de agrupamiento o <i>clustering</i>.
SEMEJANZAS	<ul style="list-style-type: none"> *Almacena datos. *Dirigido por el usuario. 	<ul style="list-style-type: none"> *Lo pueden usar diferentes usuarios. *Dirigido por el usuario.
DIFERENCIAS	<p>El análisis que realizan las herramientas OLAP es dirigido por el usuario, son consultas lanzadas sobre cubos OLAP que tienen la información precalculada y almacenada.</p> <p>*Datos dinámicos</p> <p>*Dirigido por transacciones</p> <p>*Orientado a la aplicación</p>	<p>la minería de datos permite razonar de forma inductiva a partir de lo que se llaman vistas "minables" de datos para llegar a una hipótesis general que modele el problema.</p> <p>*Datos estáticos</p> <p>*Dirigido por análisis</p> <p>*Orientado a los temas</p>
HERRAMIENTAS	<p>Herramientas Libres para OLAP: Mondrian y JPivot</p> <p>Entre las herramientas disponibles se incluyen motores para</p>	<p>* Clementine / SPSS: Herramienta de data mining que permite desarrollar modelos predictivos y desplegarlos para mejorar la toma de decisiones.</p>

implementar bases de datos multidimensionales nativas (solución **MOLAP**) o para emularlas sobre una implementación relacional (solución **ROLAP**), diseñadores de cubos, navegadores, generadores de reportes, graficadores, programas para realizar la extracción, transformación y carga de datos (**ETL**) desde bases transaccionales.

***SAS Enterprise Miner / SAS:** Solución de minería de datos que proporciona gran cantidad de modelos y de alternativas. Permite determinar pautas y tendencias, explica resultados conocidos e identifica factores que permiten asegurar efectos deseados.

* **Keel:** Es un software para evaluar la evolución de los algoritmos de minería de datos y problemas de regresión, entre ellos: clasificación, agrupamiento y patrón de la minería.