



FACULTAD DE CIENCIAS FÍSICO
MATEMÁTICAS



MINERÍA DE DATOS

PATRONES SECUENCIALES

Alumnos:

Saúl Arath Hernández Hernández 1815642

Miguel Alejandro Noriega Ortega 1798528

Alison L. Roldán Sánchez 1806488

Aranza Sifuentes Soto 1887979

Valeria Solis Agundis 1815413

Tadeo Alejandro Suárez Martínez 1806069

Docente:

Mayra Cristina Berrones
Reyes



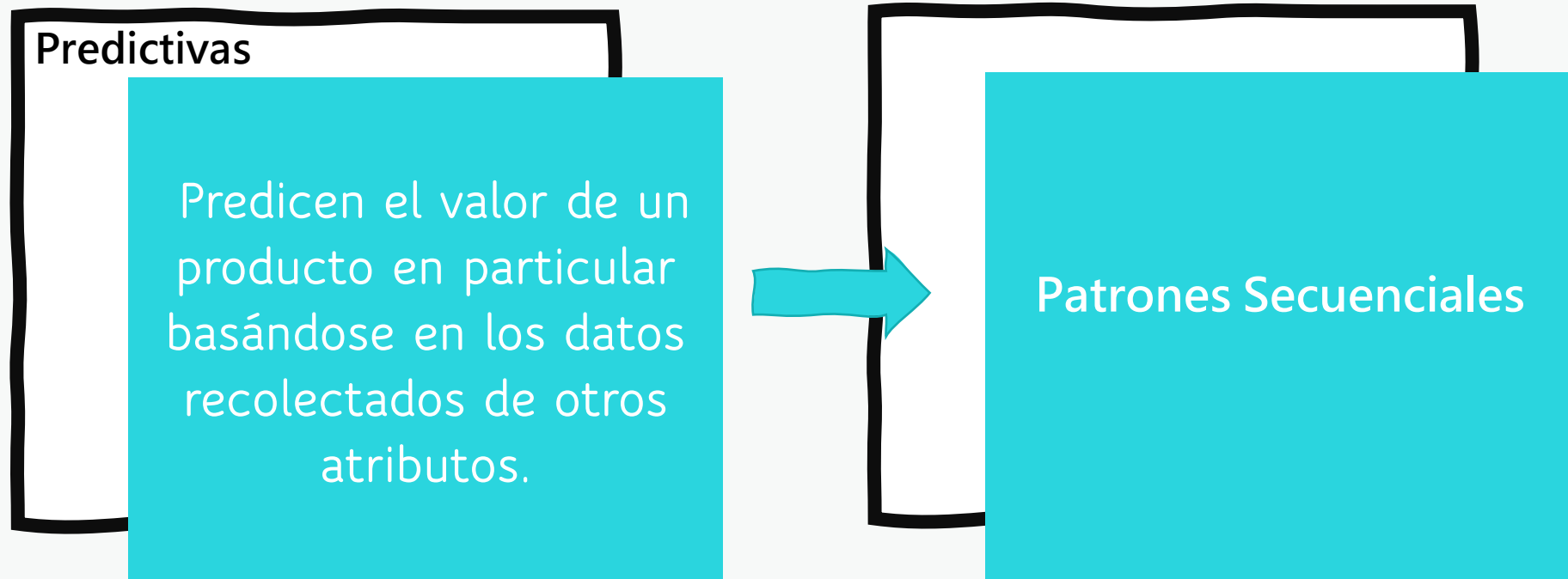
Minería de Datos

La minería de datos puede definirse como la extracción no trivial de información implícita, previamente desconocida y potencialmente útil a partir de los datos, es decir, es el descubrimiento eficiente de información valiosa (no obvia) de una gran colección de datos.



Tareas de la minería de datos

Se dividen generalmente en dos categorías: Descriptivas y Predictivas





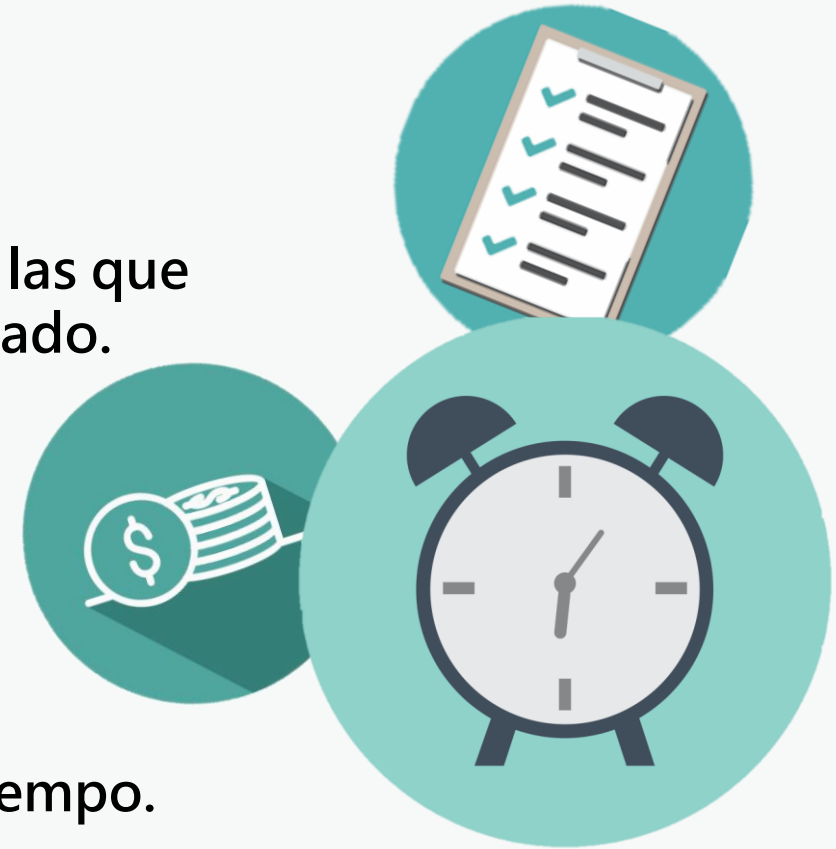
Patrones Secuenciales

Se especializan en analizar datos y encontrar subsecuencias interesantes dentro de un grupo de secuencias.

Es una clase especial de dependencia en las que el orden de acontecimientos es considerado.

El patrón secuencial describe el modelo de compras que hace un cliente particularmente o un grupo de clientes relacionando las distintas transacciones efectuadas por ellos a lo largo del tiempo.

Son eventos que se enlazan con el paso del tiempo.

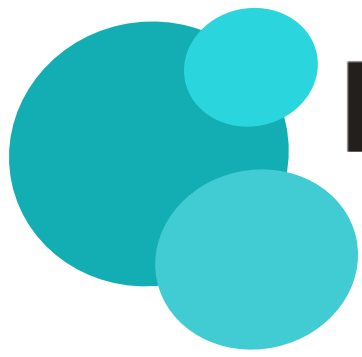


Para los Patrones Secuenciales

- Se trata de buscar asociaciones de la forma "si sucede el evento X en el instante de tiempo t entonces sucederá el evento Y en el instante $t+n$ ".
- El objetivo de la tarea es poder describir de forma concisa relaciones temporales que existen entre los valores de los atributos del conjunto de ejemplos.
- Utiliza reglas de asociación secuenciales.- reglas que expresan patrones de comportamiento secuencial, es decir, que se dan en instantes distintos en el tiempo.

Características:

- El orden importa
- Su objetivo es encontrar patrones en secuencia.
- Una secuencia es una lista ordenada de itemsets, donde cada itemset es un elemento de la secuencia.
- El tamaño de una secuencia es su cantidad de elementos (itemsets).
- La longitud de una secuencia es su cantidad de ítems.
- El soporte de una secuencia es el porcentaje de secuencias que la contienen en un conjunto de secuencias S .
- Las secuencias frecuentes (o patrones secuenciales) son las subsecuencias de una secuencia que tienen un soporte mínimo.



Entornos en los que se desarrolla

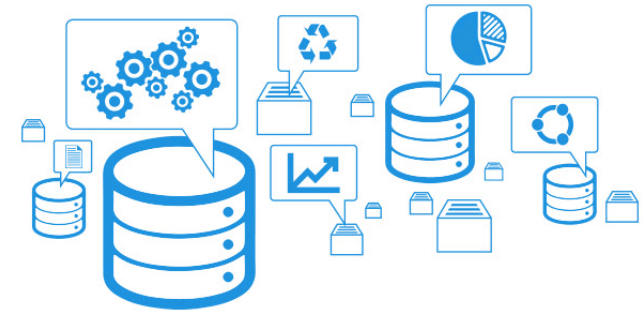
Áreas

Tipo de base de datos


- ☐ Medicina
- ☐ Biología, bioingeniería
- ☐ Web
- ☐ Análisis de mercado, distribución y comercio
- ☐ Aplicaciones financieras y banca
- ☐ Aplicaciones de seguro y salud privada
- ☐ Deportes



- ☐ Base de datos temporales
- ☐ Base de datos documentales
- ☐ Base de datos relacionales



Resolución de Problemas



AGRUPAMIENTO DE
PATRONES SECUENCIALES

CLASIFICACIÓN CON
DATOS SECUENCIALES

REGLAS DE ASOCIACIÓN
CON DATOS SECUENCIALES

Agrupación de Patrones Secuenciales

Se define como la tarea de separar en grupos a los datos, de manera que los miembros de un grupo sean muy similares entre sí, y al mismo tiempo sean diferentes a los objetivos de otros grupos.

Para la creación de agrupamientos:

- Se selecciona arbitrariamente el centro del primer agrupamiento.
- Posteriormente, se procesan secuencialmente los demás patrones mediante cálculos de distancia.

Cada M patrones se mezclan agrupamientos, estos pueden ser:

- Mezcla por cercanía.
- Mezcla por tamaño.
- Mezcla forzada.

REGLAS DE ASOCIACIÓN CON DATOS SECUENCIALES

Se presenta cuando los
datos contiguos
presentan algún tipo de
relación.

CLASIFICACIÓN CON DATOS SECUENCIALES

Éstos expresan patrones de
comportamiento
secuenciales, es decir que se
dan en instantes distintos
(pero cercanos) en el tiempo.

Métodos Representativos



GSP



SPADE



AprioriAll



FreeSpan



SPAM



PrefixSpan



ISM



IncSp



ISE



IncSpan



Referencias

Slimani, T. y Lazzez, A. (2013). *Sequential mining: patterns and algorithms analysis*. ArXiv, abs / 1311.0350
<https://arxiv.org/ftp/arxiv/papers/1311/1311.0350.pdf>

Berzal, Fernando. (sin fecha). *Patrones secuenciales*. Data Mining: Minería de Datos.
<https://elvex.ugr.es/idbis/dm/slides/22%20Pattern%20Mining%20-%20Sequences.pdf>

Guevara, Diego. Jaramillo ,Marílyn. Landacay, Katty. (26 de octubre de 2010). *Inteligencia de Negocios: Minería de Datos Secuenciales*. SlideShare. <https://www.slideshare.net/marilynsilvana/mineria-de-datos-secuenciales-5571523>