



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN®



UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE CIENCIAS DE FISICO  
MATEMATICAS

## **Resumen**

Omar Alejandro Delgado Lozano

Matricula: 1647656

# **Contenido**

Reglas de asociación

Regresion Lineal

Patrones Secuenciales

Clustering

Prediccion

Outliers

## ***Reglas de asociación***

Las reglas de asociación se emplean para descubrir patrones de objetos o atributos que suelen ocurrir juntos, a partir del estudio de bases de datos transaccionales.

Concretamente, una regla de asociación es una implicación de la forma  $X \Rightarrow Y$ , donde  $X$  e  $Y$  son dos conjuntos disjuntos de *items*. Esto significa que si encontramos todos los *items* en  $X$  en una transacción, esperamos encontrar también los *items* en  $Y$ .

## ***Regresión Lineal***

La regresión lineal es un campo de estudio que enfatiza la relación estadística entre dos variables continuas conocidas como variables de predicción y respuesta

- La variable predictora se denota con mayor frecuencia como  $x$  y también se conoce como variable independiente.
- La variable de respuesta se denota con mayor frecuencia como  $y$  y también se conoce como variable dependiente.

La regresión nos permite además, determinar el grado de dependencia de las series de valores  $X$  e  $Y$ , prediciendo el valor  $y$  estimado que se obtendría para un valor  $x$  que no esté en la distribución

### **Patrones Secuenciales**

Las secuencias son un importante tipo de datos, que representa una clase especial de estructura donde importa el orden que ocupan los elementos. Este orden, puede o no estar relacionado con el factor tiempo.

Es un tema que ocupa el encontrar patrones estadísticamente relevantes entre ejemplos de datos donde los valores se entregan en una secuencia se presume que los valores son discretos y, por lo tanto, la minería de series de tiempo está estrechamente relacionada, pero generalmente se considera una actividad diferente. La minería de patrones secuenciales es un caso especial de minería de datos estructurados.

## ***Clustering***

Conocido como agrupamiento, es una de las técnicas de minería de datos, el proceso consiste en la división de los datos en grupos de objetos similares. Cuando se representan la información obtenida a través de clusters se pierden algunos detalles de los datos, pero a la vez se simplifica dicha información.

Técnica en la que el aprendizaje realizado es no supervisado. Desde un punto de vista práctico. El *clustering* juega un papel muy importante en aplicaciones de minería de datos, tales como exploración de datos científicos, recuperación de la información y minería de texto, aplicaciones sobre bases de datos espaciales.

## ***Prediccion***

En una consulta de predicción por lotes, puede asignar el modelo a un origen externo de datos, agrupa una variedad de técnicas estadísticas de modelización, aprendizaje automático y minería de datos que analiza los datos actuales e históricos reales para hacer predicciones acerca del futuro o acontecimientos no conocidos.

## ***Outliers***

Un valor más extremo (outlier) es un valor en un conjunto de datos que es muy diferente de los otros valores son valores excepcionalmente lejanos del centro. En la mayoría de los casos, los outliers tienen influencia en la media, pero no en la mediana, o la moda. Los outliers son importantes en su efecto en la media. No hay una regla para identificar a los outliers es una observación que es numéricamente distante del resto de los datos. Las estadísticas derivadas de los conjuntos de datos que incluyen valores atípicos serán frecuentemente engañosa Los valores atípicos pueden ser indicativos de datos que pertenecen a una población diferente del resto de las muestras establecidas.