

# TÉCNICAS DE MINERÍA

## PREDICCIÓN

Predice tendencias y patrones de comportamientos en eventos presentes, pasados, futuros

- Segmentación de clientes
- Detector de fraude
- Publicidad predictiva

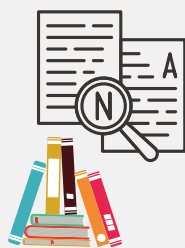


Se encarga de encontrar sub-secuencias interesantes dentro de un conjunto de secuencias.

## PATRONES SECUENCIALES

- TIPOS**
- Itemsets frecuentes
  - Asociaciones
  - Subgrafos
  - Reglas secuenciales
  - Patrones periodicos

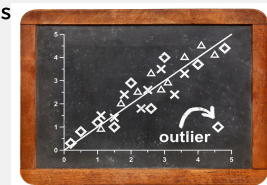
**EJEMPLO:**  
Análisis de textos, encontrar las palabras más utilizadas



## MÉTODOS DE DETECCIÓN

- Desviación Estándar
- Boxplots
- DBSCAN

Observación que se desvía mucho de las otras y despierta sospechas de ser generada por un mecanismo diferente.



## TIPOS

- 1.Outliers que surgen de un error de procedimiento
- 2.Acontecimiento extraordinario
- 3.Variables únicas
- 4.Datos extraordinarios sin explicación

## OUTLIERS

## CLUSTERING

Identifica clusters de acuerdo a una medida de similitud entre ellos.

Se usa en: **marketing, biología, la web**

## TIPOS

### De Partición Jerárquico

1. **Aglomerativo:** comienza cada elemento como cluster
2. **Divisivo:** comienza un solo cluster engloba todos los elementos

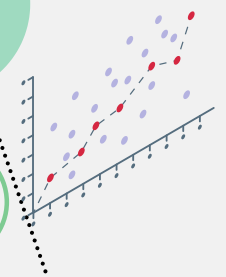


## REGRESIÓN LINEAL

## REGRESIÓN LINEAL

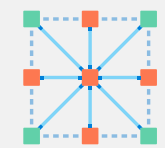
Se tiene una var. dependiente y una independiente

La variable predictora predice el valor de la variable de salida



## REGLAS DE ASOCIACIÓN

Encuentra relaciones dentro de un conjunto de transacciones



**Item** - cada elemento  
**Itemset** - conjunto de items

**SOPORTE:** frecuencia relativa

**CONFIANZA:** mide la fortaleza de la regla

**LIFT:** refleja probabilidad.

## APRIORI

- Algoritmo consiste en
- 1.buscar itemsets frecuentes,
  - 2.convertirlos en RDA

