

Clustering

Minería de Datos (CC3074) - 2026

Clustering

Semestre 01, 2026

El clustering es una técnica de aprendizaje no supervisado.

Consiste en:

- Agrupar observaciones similares
- Formar grupos llamados clusters

Idea General del Clustering

Un buen clustering busca:

- Alta similitud dentro del grupo
- Baja similitud entre grupos

No existen etiquetas correctas o incorrectas.

Usos del Clustering

Aplicaciones comunes:

- Segmentación de clientes
- Análisis de comportamiento
- Exploración inicial de datos
- Apoyo en la toma de decisiones

Interpretación de los Grupos

En clustering, los grupos deben ser interpretables.

El analista debe:

- Analizar las características de cada grupo
- Dar significado a los clusters
- Relacionarlos con el contexto del problema

Importante

- No siempre produce una única solución correcta
- Depende de los datos y de las variables usadas
- Es sensible a la escala y a los outliers

La interpretación es clave.