

Aprendizaje No Supervisado

Minería de Datos (CC3074) - 2026

Aprendizaje No Supervisado

Semestre 01, 2026

Aprendizaje Automático

El aprendizaje automático permite que los sistemas aprendan patrones a partir de datos.

Dependiendo de la información disponible, existen distintos tipos de aprendizaje.

Tipos de Aprendizaje

Según la presencia de una variable objetivo:

- Aprendizaje supervisado
- **Aprendizaje no supervisado**
- Aprendizaje semi-supervisado

Aprendizaje No Supervisado

En el aprendizaje no supervisado no existen etiquetas.

Los datos contienen únicamente:

- Variables de entrada

El objetivo es descubrir estructura oculta en los datos.

Etiquetas

No existe una variable respuesta.

No hay una columna que indique la respuesta correcta.

Variables de entrada

Los datos contienen únicamente:

- Edad
- Ingresos
- Gasto mensual

Estas variables describen a cada observación.

Ejemplo

Ejemplo de datos de clientes:

Edad | Ingresos | Gasto ---- | ---- | ---- 23 | 4500 | 1200 45 | 12000 | 3500 31 | 8000
| 2100

No existe una columna llamada “tipo de cliente”.

Objetivo

Responder preguntas como:

- ¿Existen grupos naturales en los datos?
- ¿Qué observaciones son similares?
- ¿Hay patrones que no son evidentes a simple vista?

Características Clave

- No predice una variable objetivo
- Es exploratorio
- Ayuda a entender la estructura del conjunto de datos
- Se usa frecuentemente antes del modelado supervisado

Relación con el EDA

El aprendizaje no supervisado está fuertemente relacionado con el EDA.

Ambos buscan:

- Entender los datos
- Detectar patrones
- Identificar comportamientos similares

Muchas veces se usan de forma complementaria.

Clustering

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.