

---

## RESULTADO DE APRENDIZAJE

---

### RdA de la asignatura:

- **RdA 2:** Aplicar modelos de aprendizaje automático supervisado y no supervisado, así como su validación y optimización, en la resolución de problemas tanto reales como simulados.

### RdA de la actividad:

- Describir y diferenciar los algoritmos jerárquicos aglomerativos y divisivos, junto con sus ventajas y desventajas.
- Analizar y construir dendogramas para entender los resultados de un agrupamiento jerárquico.
- Aplicar diferentes criterios de enlace en algoritmos de agrupamiento jerárquico.
- Implementar algoritmos de agrupamiento jerárquico en un conjunto de datos utilizando Python.

---

## INTRODUCCIÓN

---

**Pregunta inicial:** ¿Cómo podemos agrupar datos de manera jerárquica y visualizar las relaciones entre los grupos?

---

## DESARROLLO

---

### Actividad 1: Presentación teórica sobre algoritmos de agrupamiento jerárquico

se introducirá a los estudiantes los fundamentos del agrupamiento jerárquico, incluyendo los métodos aglomerativos y divisivos, y los criterios de enlace para calcular las distancias entre grupos. También se explicará el uso de los dendogramas como herramienta visual para entender las jerarquías de los grupos.

### ¿Cómo lo haremos?

- **Introducción a los algoritmos jerárquicos:** Explicaremos los principios de los algoritmos aglomerativos y divisivos. Se destacará cómo los métodos aglomerativos comienzan con cada punto como un clúster individual y los unen iterativamente, mientras que los divisivos parten de un clúster único y lo dividen sucesivamente.

- **Dendogramas:** Introduciremos los dendogramas como una herramienta para visualizar la jerarquía de los agrupamientos. Analizaremos cómo interpretar las uniones y las distancias representadas en el gráfico.
- **Criterios de enlace:** Explicaremos los tres criterios principales para determinar las distancias entre clústeres:
  - **Enlace sencillo (Single Linkage):** Distancia mínima entre dos clústeres.
  - **Enlace completo (Complete Linkage):** Distancia máxima entre dos clústeres.
  - **Enlace promedio (Average Linkage):** Promedio de todas las distancias entre puntos de los dos clústeres.

**Verificación de aprendizaje:** Los estudiantes deberán:

- Identificar correctamente las diferencias entre los métodos aglomerativos y divisivos.
- Explicar qué representa un dendograma y cómo se interpreta.
- Comparar los efectos de los diferentes criterios de enlace en un ejemplo dado.

### Actividad 2 : Agrupamiento jerárquico con una lista de números

Se realizará un ejercicio práctico para agrupar una lista de números utilizando un enfoque jerárquico. Se construirán grupos utilizando diferentes criterios de enlace y se visualizará el proceso en un dendograma manual.

#### ¿Cómo lo haremos?

- **Datos iniciales:** Se trabajará con la siguiente lista de números:  
[10, 4, 20, 30, 38, 87, 82, 56, 66, 70]
- **Cálculo de distancias:** Se calcularán las distancias absolutas entre todos los pares de números para construir una matriz de distancias inicial. La distancia entre dos números será la diferencia absoluta entre ellos.
- **Construcción del dendograma:** Se fusionarán los números o grupos más cercanos según el criterio de enlace seleccionado:
  - **Enlace sencillo:** Se tomará la distancia mínima entre elementos de los dos grupos.
  - **Enlace completo:** Se tomará la distancia máxima entre elementos de los dos grupos.
  - **Enlace promedio:** Se calculará el promedio de todas las distancias entre elementos de los dos grupos.
- **Representación manual:** A medida que se fusionan los grupos, se registrarán los pasos y se dibujará un dendograma manualmente en el pizarrón o en hojas de trabajo distribuidas a los estudiantes.

**Verificación de aprendizaje:** Replicar el proceso con otro conjunto de números.

### Actividad 3 : Visualización de video sobre algoritmos de agrupamiento jerárquico

Se utilizará un video educativo para que los estudiantes comprendan los conceptos fundamentales del agrupamiento jerárquico, incluyendo los métodos aglomerativos y divisivos, los criterios de enlace, y la interpretación de dendogramas.

#### ¿Cómo lo haremos?

- **Proyección de video:** Se presentará un video educativo que explique los fundamentos del agrupamiento jerárquico.

Enlace al video: [Hierarchical Cluster Analysis](#).

- **Interacción con ChatGPT:** Después del video, cada estudiante formulará preguntas a ChatGPT relacionadas con términos o conceptos que no comprendieron durante la visualización.

**Verificación de aprendizaje:** Se realizará una discusión grupal donde los estudiantes compartirán:

- Las preguntas que realizaron a ChatGPT.
- Las respuestas que obtuvieron y cómo estas les ayudaron a comprender mejor el tema.

### Actividad 4: Implementación directa del agrupamiento jerárquico

Los estudiantes ejecutarán el siguiente código para generar un agrupamiento jerárquico sobre el conjunto de datos y visualizar el dendograma.

**Datos iniciales:** Se trabajará con el siguiente conjunto de datos:

[10, 4, 20, 30, 38, 87, 82, 56, 66, 70]

#### Código a ejecutar:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.cluster.hierarchy import dendrogram, linkage

# Datos de entrada
data = np.array([10, 4, 20, 30, 38, 87, 82, 56, 66, 70]).reshape(-1, 1)

# Aplicar el agrupamiento jerárquico con criterio de enlace promedio
linkage_matrix = linkage(data, method='average')

# Visualizar el dendograma
plt.figure(figsize=(8, 5))
dendrogram(linkage_matrix, labels=data.flatten())
```

```
plt.title('Dendograma – Enlace Promedio')
plt.xlabel('Puntos de datos')
plt.ylabel('Distancia')
plt.show()
```

**Resultados esperados:** Los estudiantes deben analizar el dendograma generado e identificar:

- El orden en que se formaron los clusters.
- Las distancias a las que se unieron los diferentes grupos.

**Verificación de aprendizaje:** Responder las siguientes preguntas basadas en los resultados:

- ¿Cuál fue la distancia más grande entre dos clusters antes de unirse?
- ¿Qué estructura jerárquica muestran los datos?
- ¿Cómo podría cambiar el resultado si utilizáramos un criterio de enlace diferente (e.g., enlace sencillo o completo)?

---

### CIERRE

---

**Tarea:** Implementar el algoritmos en los datos seleccionados la clase anterior.

**Pregunta de investigación:**

1. ¿Qué información podemos obtener de un dendograma que no podemos obtener directamente de otros modelos?
2. ¿Cómo afecta el criterio de enlace (sencillo, completo o promedio) a la forma final de los grupos en el agrupamiento jerárquico?
3. ¿En qué casos prácticos sería más útil un método divisivo en lugar de uno aglomerativo?