

---

## RESULTADO DE APRENDIZAJE

---

### RdA de la asignatura:

- **RdA 1:** Plantear los conceptos fundamentales del aprendizaje automático, incluyendo los principios básicos, técnicas de preprocesado de datos, métodos de evaluación y ajuste de modelos, destacando su importancia en el análisis y resolución de problemas de datos.

### RdA de la actividad:

1. Definir qué es aprendizaje supervisado y no supervisado.
2. Identificar los tipos de tareas asociadas (clasificación, regresión y agrupamiento).
3. Enumerar los principales modelos utilizados en cada tipo de tarea.
4. Comprender las métricas comunes utilizadas para medir distancias.

---

## INTRODUCCIÓN

---

**Pregunta inicial:** Si tuvieras un conjunto de datos de imágenes de flores, ¿qué datos se necesitaría para construir un modelo que identifique su tipo automáticamente?

- Breve presentación sobre el contexto histórico del aprendizaje automático y su relevancia actual.

---

## DESARROLLO

---

### Actividad 1: Definición de aprendizaje supervisado y no supervisado (10 minutos)

Explicaremos las definiciones básicas:

- **Supervisado:** El modelo aprende a partir de datos etiquetados, buscando una relación entre entradas y salidas.
- **No supervisado:** El modelo identifica patrones en datos no etiquetados.

**Recursos:** Presentación visual con ejemplos.

**Verificación de aprendizaje:** Presentación de ejemplos para determinar tipo de aprendizaje.

**Actividad 2: Tipos de tareas y modelos principales (20 minutos)**

Explicaremos los tipos de tareas y los modelos más utilizados:

- **Clasificación:** Regresión logística, SVM, Árboles de decisión, Redes neuronales.
- **Regresión:** Regresión lineal, Regresión polinómica, KNN, Redes neuronales.
- **Agrupamiento:** Modelos: K-Means, DBSCAN, Hierarchical Clustering.

**Recursos:** Cuadro resumen.

**Verificación de aprendizaje:** Presentación de ejemplos para determinar tipo de tarea.

**Actividad 3: Explicación de métricas de distancia (15 minutos)**

Detallaremos las métricas más comunes:

- Euclidiana:  $\sqrt{\sum (x_i - y_i)^2}$
- Mahalanobis: Basada en la covarianza.
- Hamming: Comparación binaria.
- Minkowski: Generalización de las métricas Euclidiana y Manhattan.

**Ejercicio práctico:** Comparar distancias entre puntos con diferentes métricas en Python.

```
import numpy as np
from scipy.spatial.distance import euclidean, mahalanobis, hamming, minkowski

# Puntos de ejemplo
x = np.array([1, 2])
y = np.array([4, 6])

# Distancias
print("Distancia Euclidiana:", euclidean(x, y))
print("Distancia Minkowski (p=3):", minkowski(x, y, p=3))

# Ejemplo de distancia de Hamming
x_bin = np.array([1, 0, 1])
y_bin = np.array([0, 1, 1])
print("Distancia Hamming:", hamming(x_bin, y_bin))
```

---

**CIERRE**

---

**Tarea:** No se deja tarea.

**Pregunta de investigación:**

- ¿Qué es la distancia de Mahalanobis y en qué casos se usa?
- ¿Cómo afecta la escala de las variables a las métricas de distancia? ¿Cómo podemos solucionarlo?
- ¿Qué es el método del codo (Elbow Method) en KMeans?