

---

### **ACTIVIDAD PREVIA**

---

- Revisar el control de lectura, solicitar ejemplos de modelos que sean de Inteligencia Artificial y modelos que no lo sean.

---

### **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 clase:**

- Definir qué es aprendizaje supervisado y no supervisado, identificando los tipos de tareas asociadas (clasificación, regresión y agrupamiento).
- Enumerar los principales modelos utilizados en cada tipo de tarea.
- 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?

---

### **DESARROLLO**

---

**Actividad 1: Definición de aprendizaje supervisado y no supervisado**

En esta actividad se explicarán las definiciones básicas del aprendizaje supervisado y no supervisado mediante clase magistral con presentación visual y ejemplos prácticos para identificar cada tipo de aprendizaje.

**¿Cómo lo haremos?**

- **Conceptos fundamentales:** Se presentarán las definiciones: Aprendizaje supervisado (el modelo aprende a partir de datos etiquetados, buscando una relación entre entradas y salidas) y aprendizaje no supervisado (el modelo identifica patrones en datos no etiquetados).

- **Presentación de ejemplos:** Se mostrarán casos prácticos para que los estudiantes identifiquen el tipo de aprendizaje correspondiente.
- **Materiales de apoyo:** Se utilizará la sección uno del archivo [Resumen02.pdf](#)

### Actividad 2: Tipos de tareas y modelos principales

En esta actividad se explicarán los tipos de tareas del aprendizaje automático y los modelos más utilizados en cada una, mediante clase magistral con cuadro resumen y ejemplos de aplicación.

#### ¿Cómo lo haremos?

- **Clasificación de tareas:** Se presentarán las tres tareas principales:
  - Clasificación
  - Regresión
  - Agrupamiento
- **Análisis de ejemplos:** Se presentarán casos prácticos para que los estudiantes identifiquen el tipo de tarea correspondiente.
- **Materiales de apoyo:** Se utilizará la sección dos del archivo [Resumen02.pdf](#)
- **Planteamiento de ejemplos:** Los estudiantes interactuarán con un chatbot para plantear ejemplos adicionales de cada tipo de tarea y modelo. Utilizar un prompt como:

Plantéame cinco ejemplos prácticos simples y cinco complejos de problemas que se resuelven con clasificación, regresión y agrupamiento, sin indicar directa o indirectamente el tipo de tarea. En total 10 ejemplos u estén mezclados los tipos de tareas. No menciones los nombres de las tareas en las respuestas ni alusiones a etiquetas, clasificación, predicción, etc.

### Actividad 3: Explicación de métricas de distancia

En esta actividad se detallarán las métricas de distancia más comunes utilizadas en aprendizaje automático, mediante exploración de un cuaderno de Jupyter con ejemplos prácticos de cálculo.

#### ¿Cómo lo haremos?

- **Métricas de distancia:** Se presentarán las métricas más comunes:
  - Euclidiana:  $\sqrt{\sum (x_i - y_i)^2}$
  - Manhattan:  $\sum |x_i - y_i|$
  - Mahalanobis: Basada en la covarianza.
  - Minkowski: Generalización de las métricas Euclidiana y Manhattan.
- **Materiales de apoyo:** Se utilizará la sección tres del archivo [Resumen02.pdf](#)

---

**CIERRE**

---

**Verificación de aprendizaje:**

1. ¿Cuál es la diferencia fundamental entre el aprendizaje supervisado y el no supervisado?
2. ¿Qué tipo de problemas se resuelven con aprendizaje supervisado y cuáles con no supervisado?

**Preguntas tipo entrevista:**

1. Explica con un ejemplo práctico cómo decidirías entre usar aprendizaje supervisado o no supervisado para un problema de negocio real.
2. Si tienes datos etiquetados, ¿siempre debes usar aprendizaje supervisado?

**Tarea:** Identificar un conjunto de datos con las 5 características sobre los cuales se pueda aplicar modelos de aprendizaje supervisado y no supervisado, el cual será usado en los talleres durante todo el curso. No se debe usar conjuntos de datos muy conocidos como Iris, MNIST, Titanic, CaBoston Housing Prices, Wine Quality, etc. Se puede usar sitios como Kaggle, UCI Machine Learning Repository, Datos Abiertos del Ecuador, etc.

**Pregunta de investigación:**

1. ¿Qué es la distancia de Mahalanobis y en qué casos se usa? (ver [Distancia euclídea y distancia de Mahalanobis](#))
2. ¿Cuándo se emplea la distancia de Minkowski con  $p$  diferente de 1 y 2?

**Para la próxima clase:** Llevar la descripción del conjunto de datos seleccionado para la tarea, indicando la fuente, número de registros, número de características, tipo de características (numéricas, categóricas, etc.) y una breve descripción del problema que se pretende resolver con los datos.