



# Fundamentos de IA y Machine Learning

Lección 1: Introducción al Aprendizaje Automático

# Preprocesado y visualización de una base de datos

## CASO PRÁCTICO

En esta actividad vamos a aplicar los conceptos teóricos vistos en clase y disponibles en el manual de la Lección 1. Para ello, se pide realizar los siguientes ejercicios haciendo uso de Excel:

1. Descargue el *dataset* IRIS disponible en <https://gist.github.com/netj/8836201>
2. Describir la base de datos: información, variables o atributos, variable objetivo y número de patrones por cada clase. Puede realizar una búsqueda bibliográfica.
3. Por cada uno de los atributos, calcular:
  - a. Media
  - b. Desviación típica
  - c. Mediana
  - d. Cuartiles
  - e. Mínimo y máximo
4. Visualice la relación entre atributos. ¿Hay alguna relación que sea visualmente significativa? (Use *scatter-plot*).
5. Realice un histograma de los atributos de entrada y la variable objetivo.
6. ¿Qué técnica utilizaría para recuperar los valores perdidos? En caso de que existan, aplíquela.
7. ¿Es un *dataset* balanceado?