

# **Aprendizaje Automático**

## **II. Análisis Exploratorio de Datos**

Juan F. Döppler

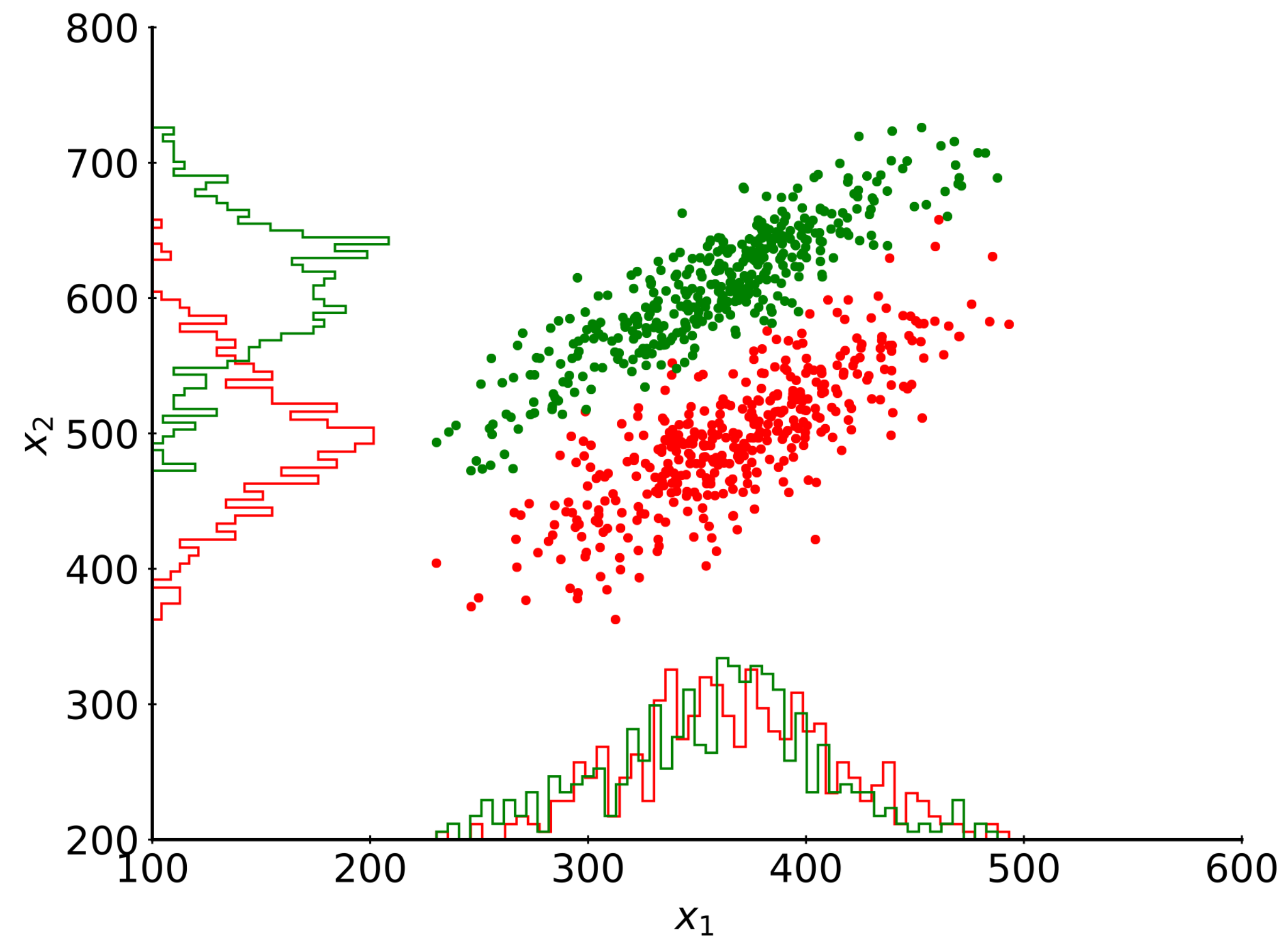
**UNSAM - 21 de septiembre de 2023**

# Contenidos de la materia

- Introducción al Análisis de Datos.
  - Probabilidad y Estadística (y algo de Bayes)
  - Análisis exploratorio de datos
  - Modelos de un solo parámetro
  - Regresión Lineal
- Machine Learning “clásico”.
  - Modelos lineales para regresión y clasificación.
  - Support Vector Machine
  - Árboles de Decisión
  - Métodos de Ensemble (RandomForests, ...)
  - Otros algoritmos
- Redes Neuronales y Aprendizaje Profundo
  - Multi-Layer Perceptron,
  - Convolutional Neural Networks.
  - Otra arquitecturas (RNN, AE, GAN, ...).

## Features

$x_1$	$x_2$	t(target)
455	260	0
449.8	425	1
440	320	1
438	280	0
391	387	1
315	470	0
275	440	0
222.3	500	1
....	....	....



# ¿Qué es Aprendizaje automático?

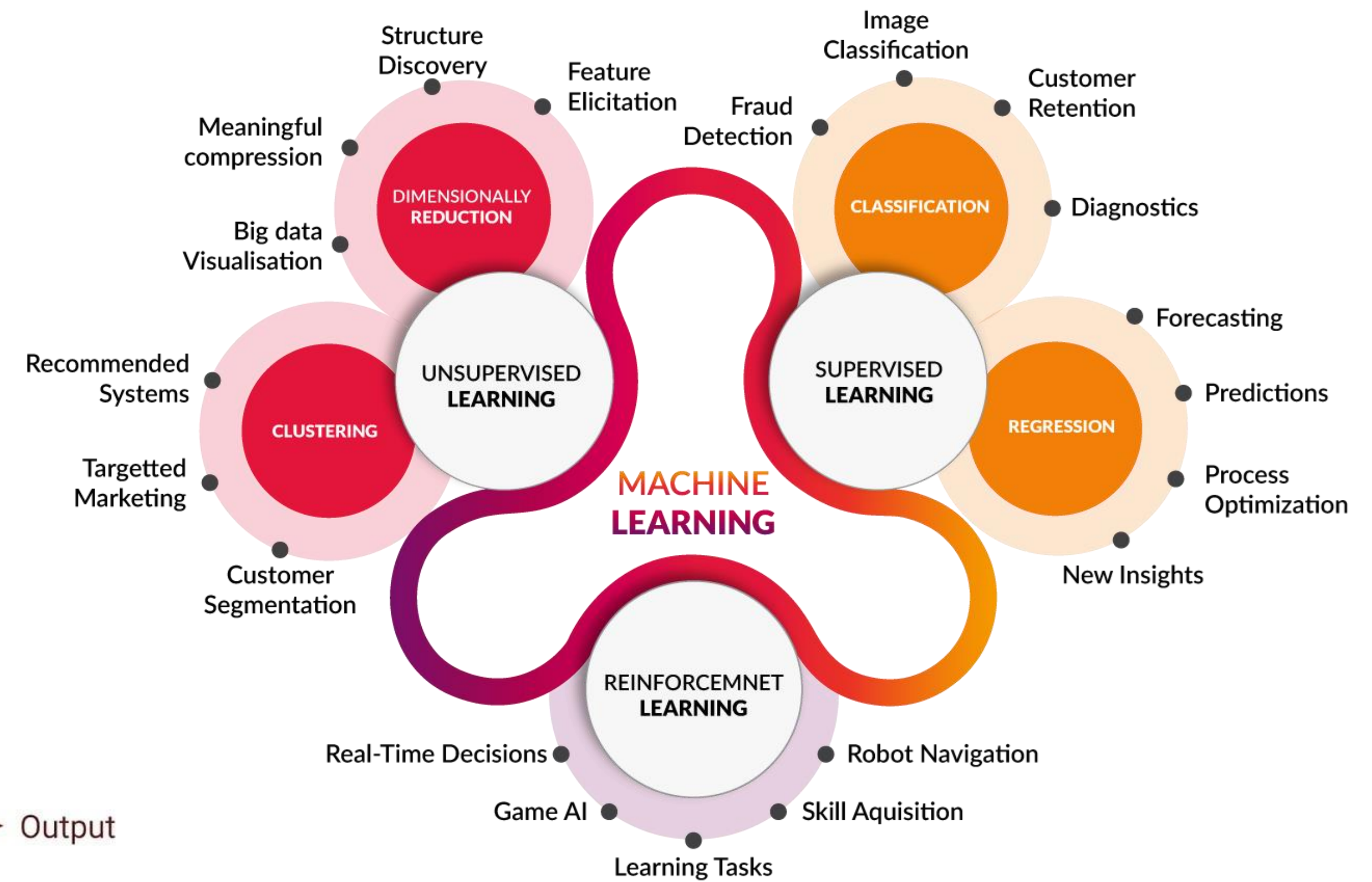
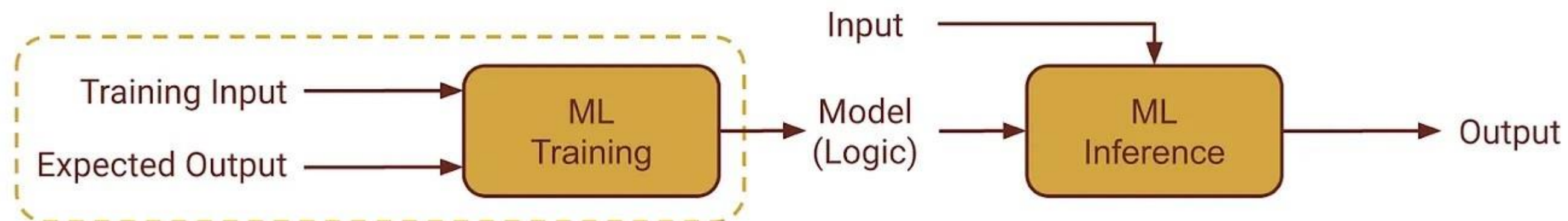


# ¿Qué es Aprendizaje automático?

**Traditional Programs:** Define algo/logic to compute output



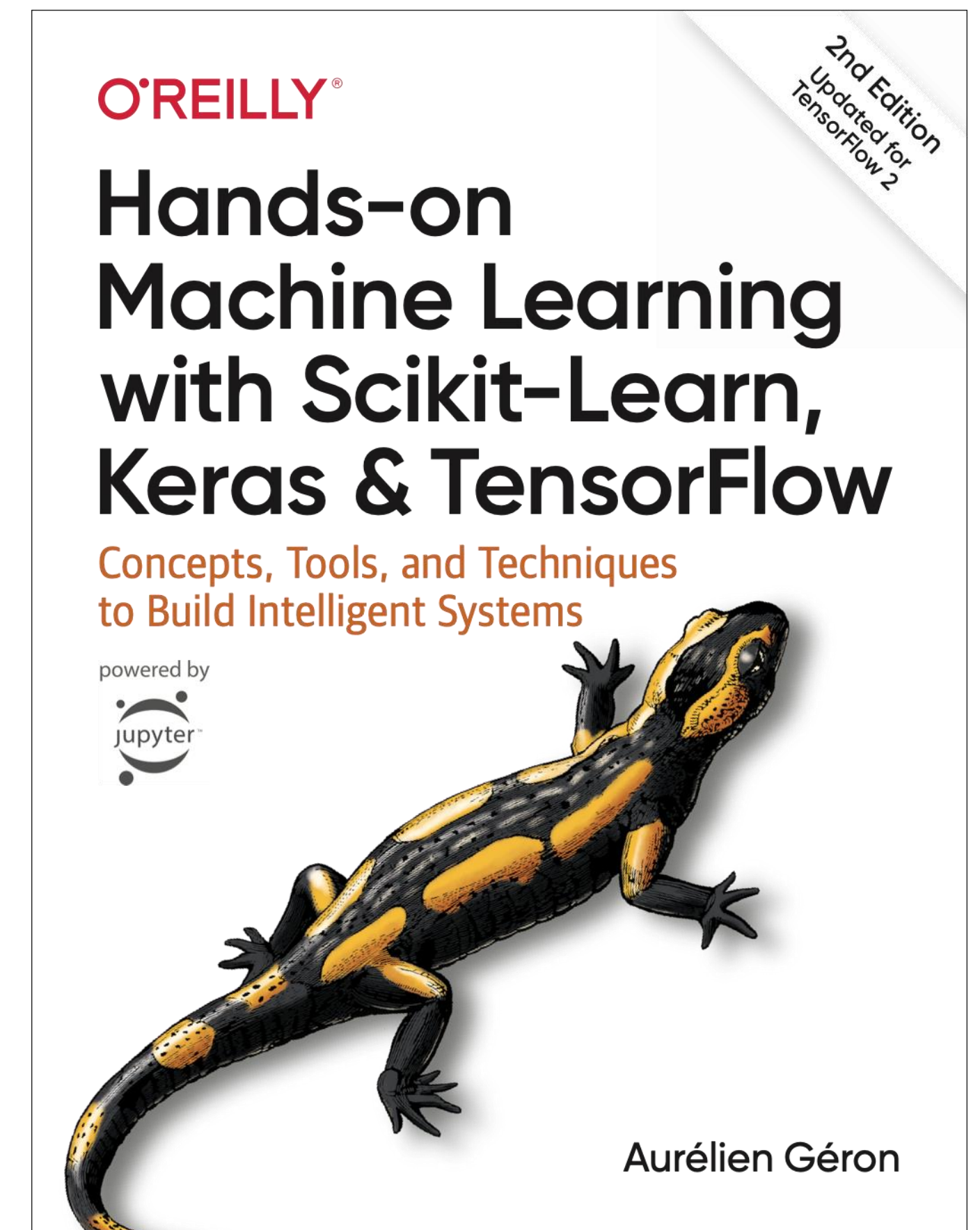
**Machine Learning:** Learn model/logic from data



# Pasos de un proyecto

Credit: A. Géron

1. Look at the big picture.
2. Get the data.
3. Discover and visualize the data to gain insights.
4. Prepare the data for Machine Learning algorithms.
5. Select a model and train it.
6. Fine-tune your model.
7. Present your solution.
8. Launch, monitor, and maintain your system.

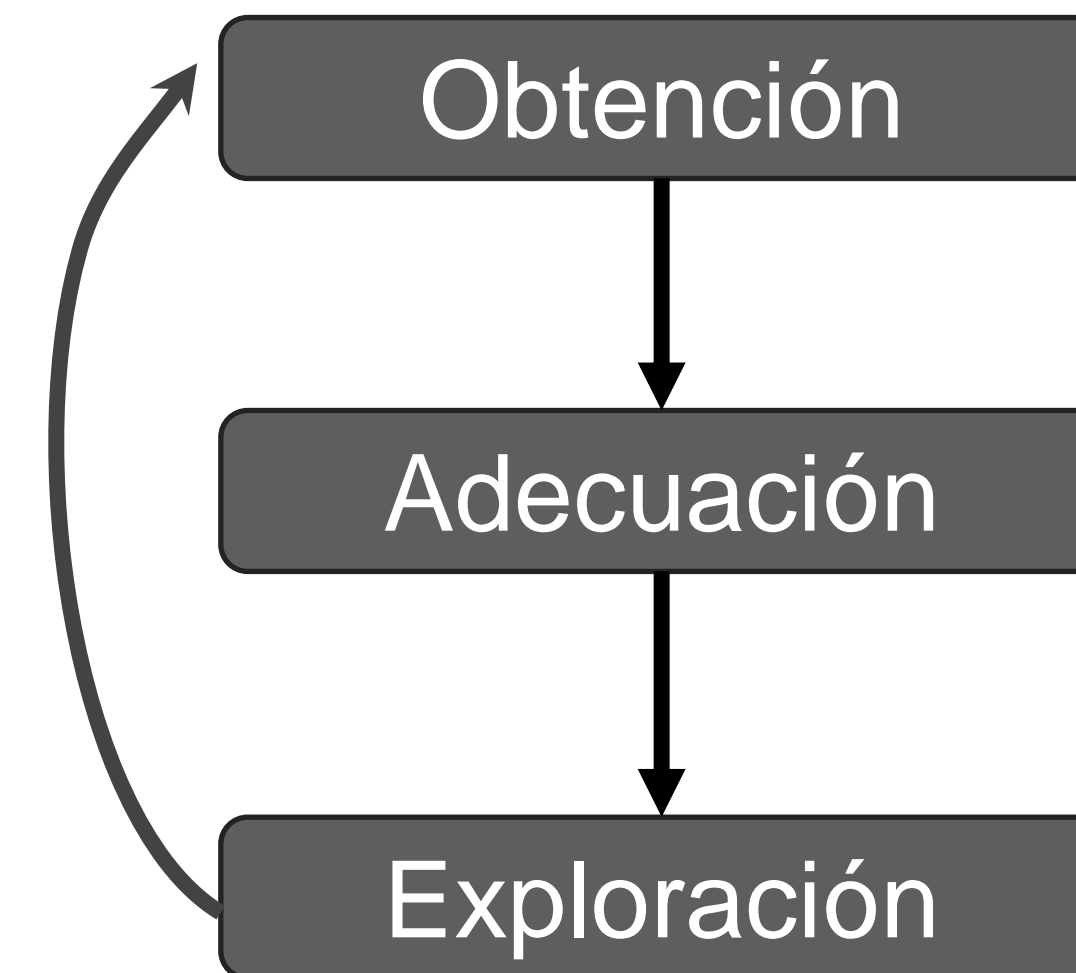




# Pasos de un proyecto

Credit: A. Géron

1. Look at the big picture.
- 2. Get the data.**
- 3. Discover and visualize the data to gain insights.**
- 4. Prepare the data for Machine Learning algorithms.**
5. Select a model and train it.
6. Fine-tune your model.
7. Present your solution.
8. Launch, monitor, and maintain your system.



# ¿Por qué hacer AED?



- ¿Qué tipo de datos son?
- ¿Qué valores toman?
- ¿Cuántos son?
- ¿Están completos?
- Métricas estadísticas
- Gráficos (distribuciones de variables, bivariados, etc)
- ¿Hay valores extremos?
- ¿Correlaciones?
- ...

- ¿Qué tipo de datos son?
  - ¿Qué valores toman?
  - ¿Cuántos son?
  - ¿Están completos?
  - Métricas estadísticas
  - Gráficos (distribuciones de variables, bivariados, etc)
  - ¿Hay valores extremos?
  - ¿Correlaciones?
  - ...
- 
- ¿Cuándo frenar?

# **¡A los notebooks!**

**Por favor, abran 02\_AED.ipynb**