

# MAGNA

INSTITUCIÓN DE ESPECIALIZACIÓN PROFESIONAL

CURSO DE FORMACIÓN

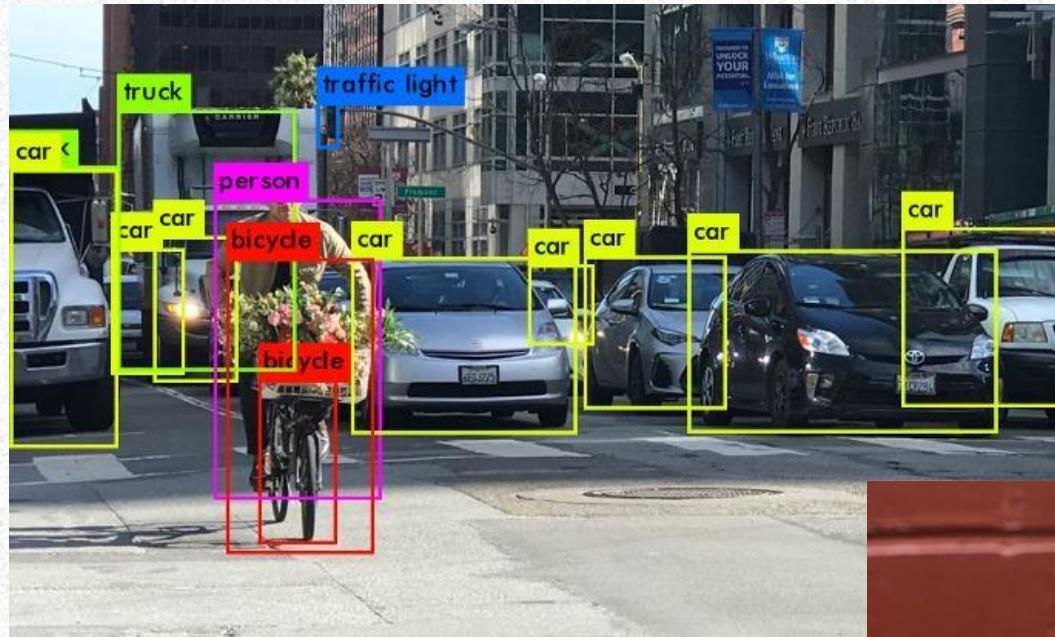


# AUTOMATIZACION INSTRUMENTACION Y CONTROL DE PROCESOS INDUSTRIALES

# Módulo 9

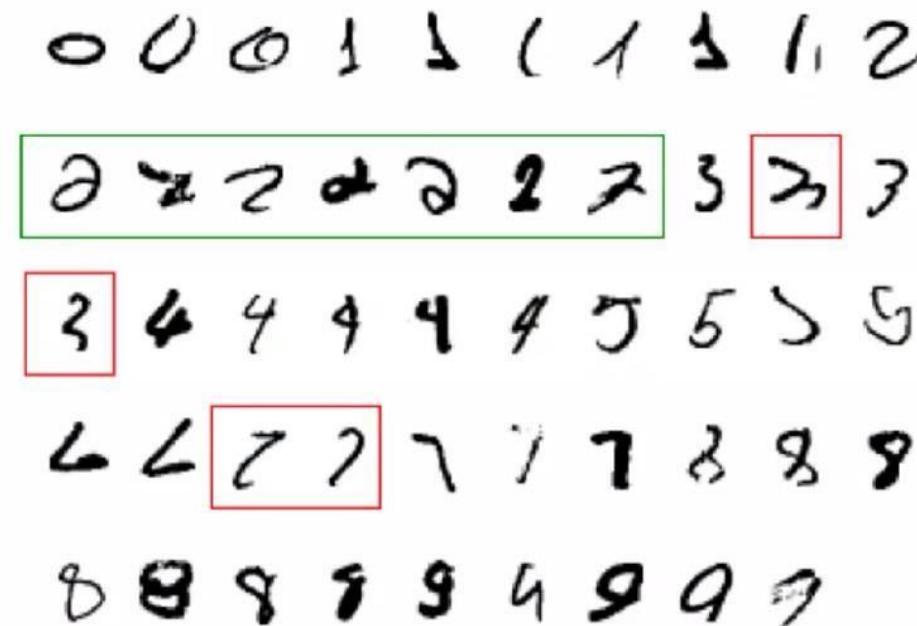
- Machine learning
- Tipos de aprendizaje
- Datos
- Regresión lineal

# Machine Learning



# Machine Learning

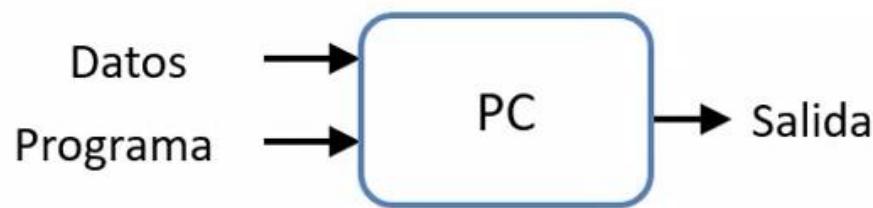
- Ejemplo de cuándo usar machine learning:
  - ¿“Qué” define a un “2” o a un “3”?



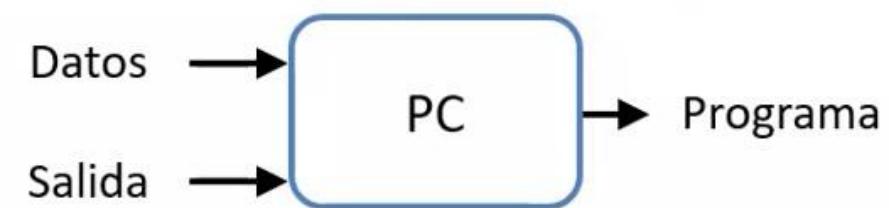
- Clasificación de dígitos: tarea “clásica” de machine learning

# Que es machine learning

- Programación tradicional

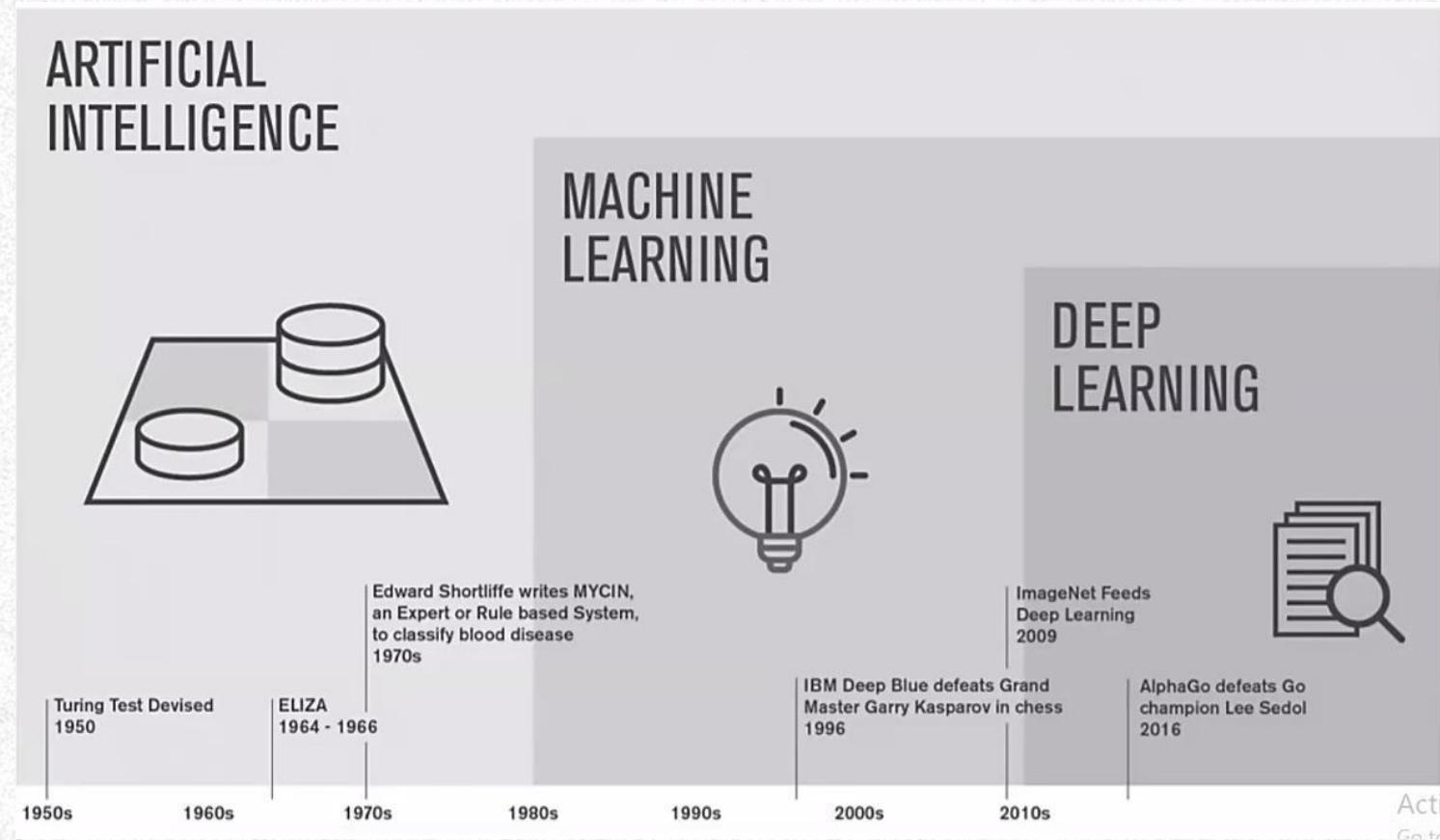


- Machine Learning

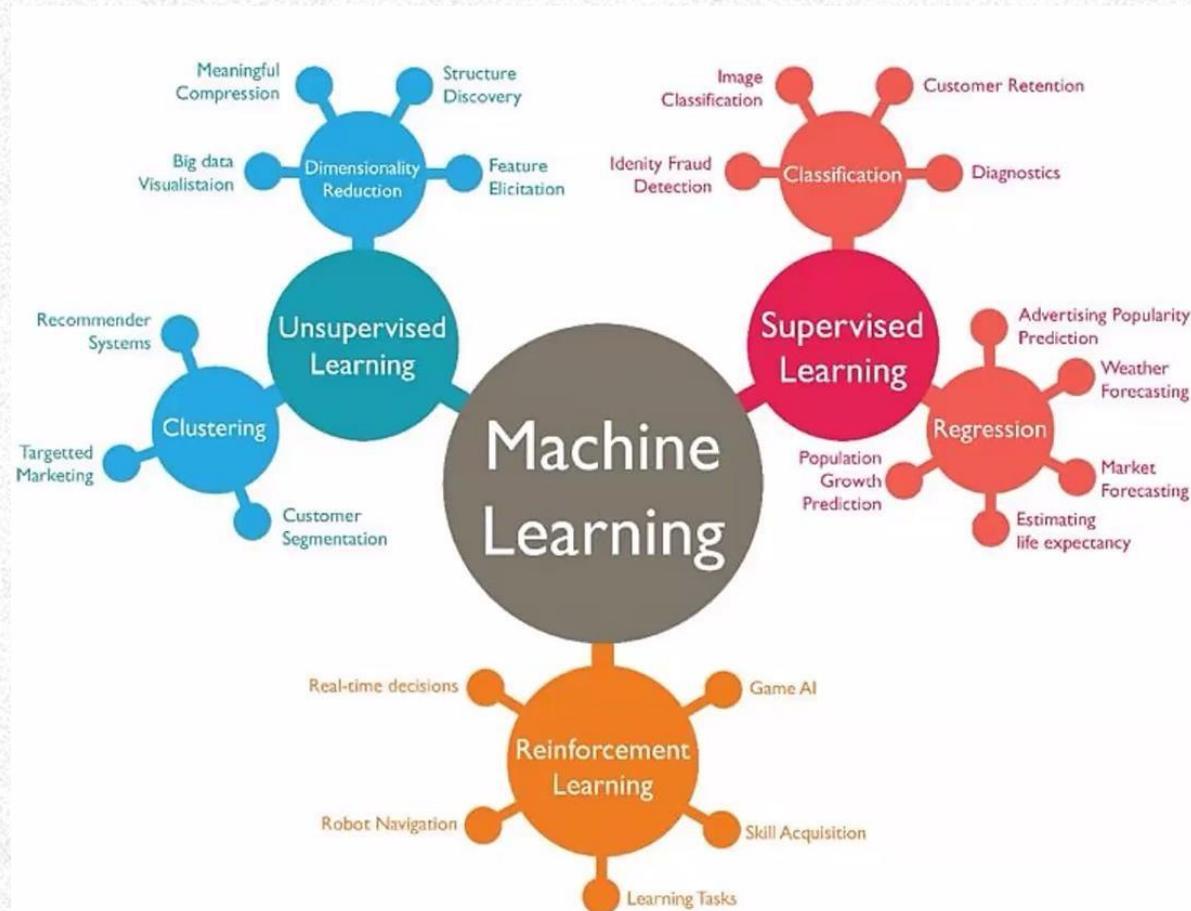


# Que es machine learning

- Las placas de video trabajan con gran cantidad de datos a un poco menos de velocidad de la cpu.



# Machine Learning: Tipos de aprendizaje

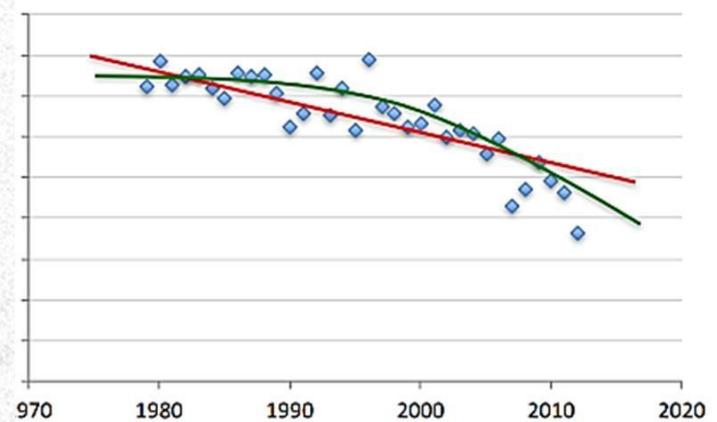
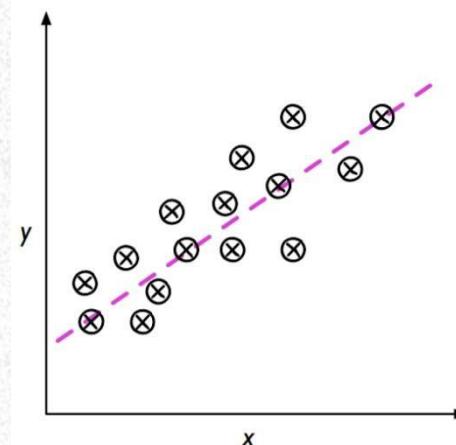


# Machine Learning: Tipos de aprendizaje

- **Tipo de aprendizaje**
  - **Aprendizaje supervisado (inductivo)**
    - Los datos de entrenamiento incluyen las salidas (“respuestas”) correctas
    - Objetivo: encontrar la predicción dada la entrada
    - Tipos: regresión, clasificación
  - **Aprendizaje no supervisado**
    - Los datos de entrenamiento no incluyen salidas (“respuestas”) correctas
    - Objetivo: descubrir patrones similares, estructuras, sub-espacios
    - Tipos: “clustering”, reducción de dimensionalidad
  - **Reinforcement learning (Aprendizaje por refuerzo)**
    - Recompensa dadas una serie de acciones
    - Objetivo: tratar de aprender usando feedback retrasado (recompensa)

# Aprendizaje Supervisado

- Regresión (predicción)
  - Dados  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
  - Objetivo: “Aprender” una función  $h(x)$  que prediga  $y$  dado  $x$  (*función  $h$  que se llama función de hipótesis*)
  - Salida:  $y$  es un número real
  - Ejemplo:
    - Precio de un carro usado



<https://www.youtube.com/watch?v=rLVuedvxOls>

# Aprendizaje Supervisado

Ejemplos clásicos de regresión: Se modela la relación que tiene las entradas con las salidas.

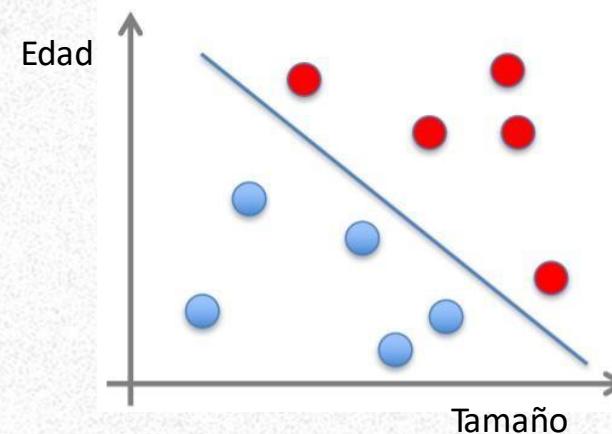
## Clasificación

- Dados  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
- Objetivo: “Aprender” una función  $h(x)$  que prediga  $y$  dado  $x$

-Salida:  $y$  es categórico (ejemplo: 0 o 1)

Ejemplo:

- Determinar si un tumor es benigno (0) o maligno (1) con base en su tamaño y en la edad del paciente
- La regresión es con datos de valores de tipo real



# Aprendizaje Supervisado

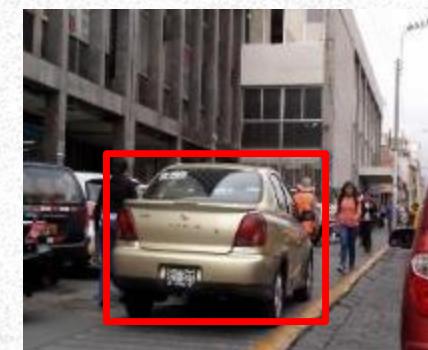
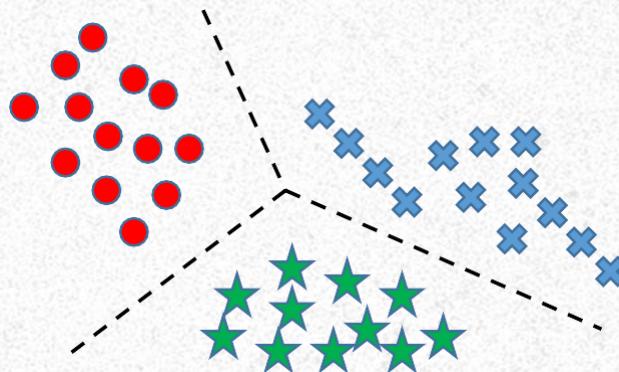
## Clasificación: Métodos usados

- Regresión logística
- Redes neuronales
- Support Vector Machines (SVM)
- Naive Bayes
- Redes Bayesianas
- Árboles de decisión
- K-nearest neighbor

## - Aplicaciones

- Reconocimiento de rostros
- Reconocimiento de objetos
- Reconocimiento de habla
- Diagnóstico médico
- Anuncios por internet

Clasificación de tres clases,  
La clasificación se hace en categorías



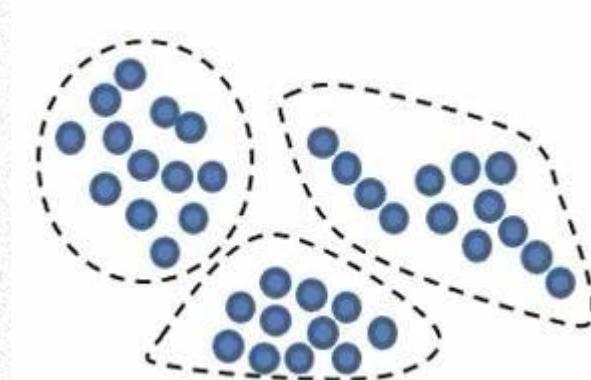
# Aprendizaje No Supervisado

“Clustering” (agrupamiento)

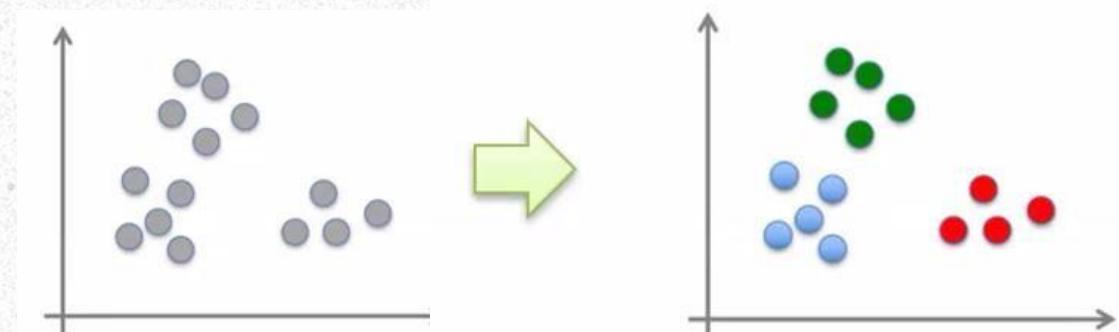
- Dados  $x_1, x_2, \dots, x_n$  (sin etiquetas o salidas deseadas)
- Objetivo: encontrar la estructura escondida (grupos) dentro de todos los x

Ejemplo

- Segmentación de mercado
- Análisis de redes sociales
- Métodos: k-means, clustering aglomerativo, mean-shift, espectral



Recomendación de Netflix

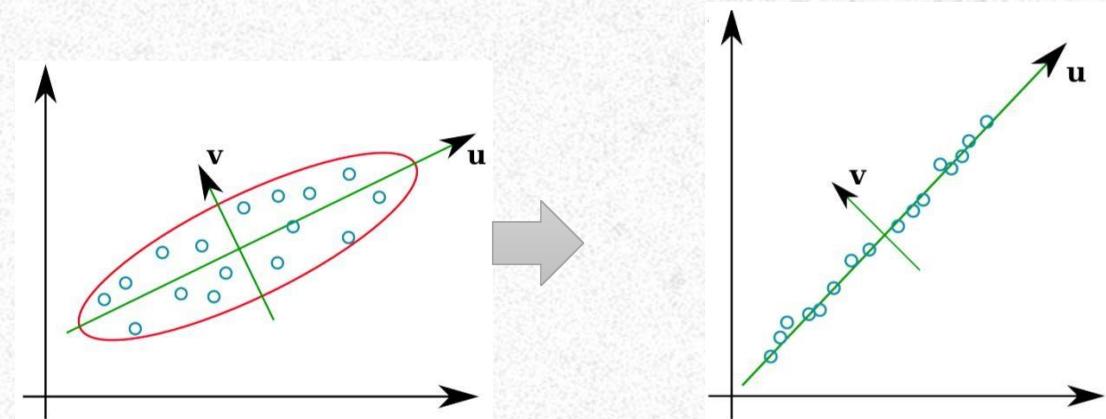


Métodos: k-means, clustering aglomerativo, mean-shift, espectral

# Aprendizaje No Supervisado

## Reducción de Dimensionalidad

- Dados:  $x_1, x_2, \dots, x_n$ , cada uno con  $m$  atributos ( $x_i \in \mathbb{R}^m$ )
- Objetivo:
  - Reducir el tamaño de cada  $x$  a solo  $d$  atributos ( $d < m$ )
  - Se debe preservar los atributos más importantes



## Métodos:

- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

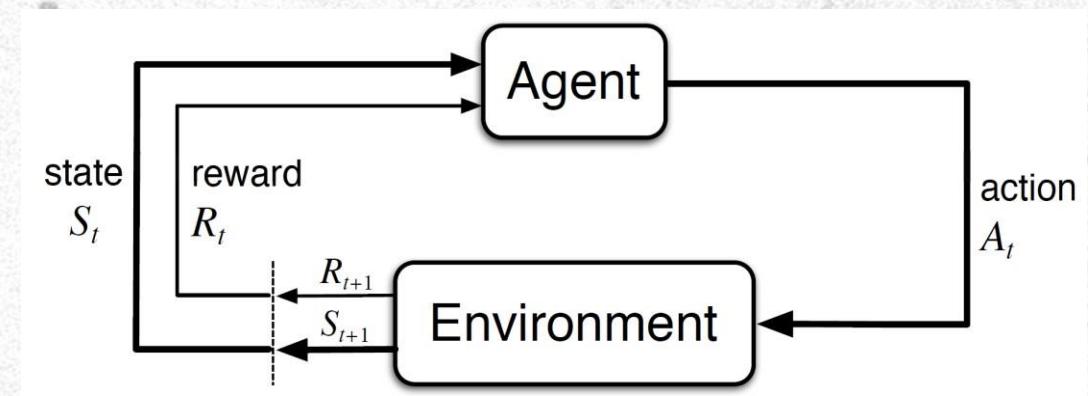
# Reinforcement Learning

Dados:

- Secuencia de estados:  $s_1, s_2, s_3, s_4, \dots$
- Secuencias de acciones:  $a_1, a_2, a_3, a_4, \dots$
- Secuencia de recompensas:  $r_1, r_2, r_3, r_4, \dots$   
(y recompensa final)

Objetivo:

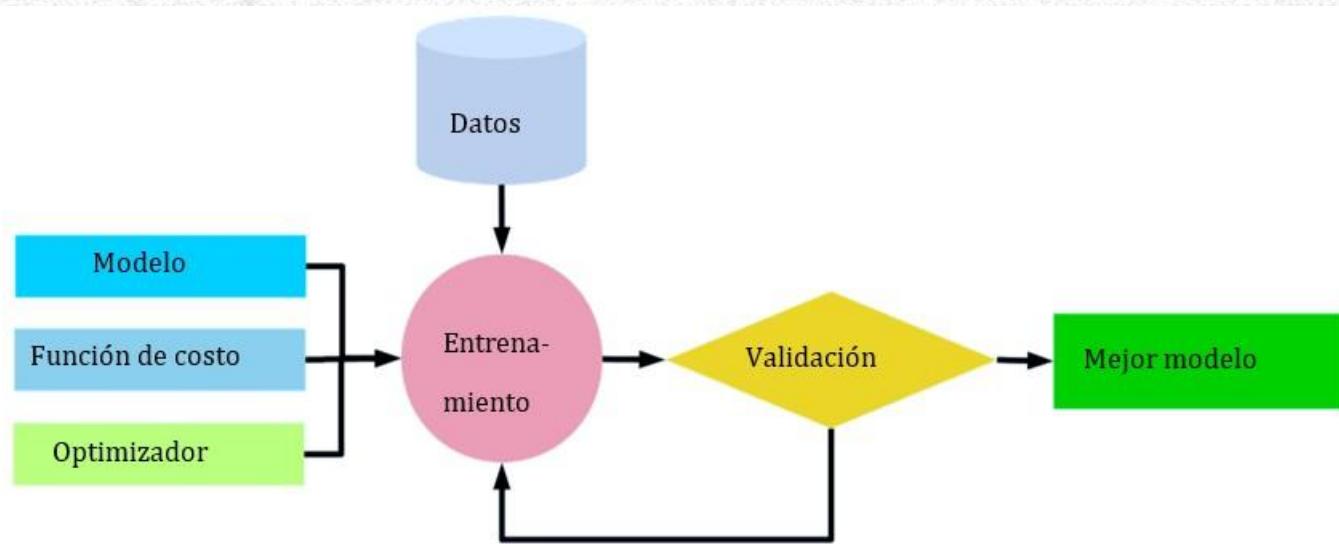
- Encontrar la “política” (policy) óptima que lleva a un estado deseado



- Ejemplos: juegos, robot en un laberinto, hacer volar un helicóptero

# Esquema General de Aplicación

- Entrenamiento (*training*):



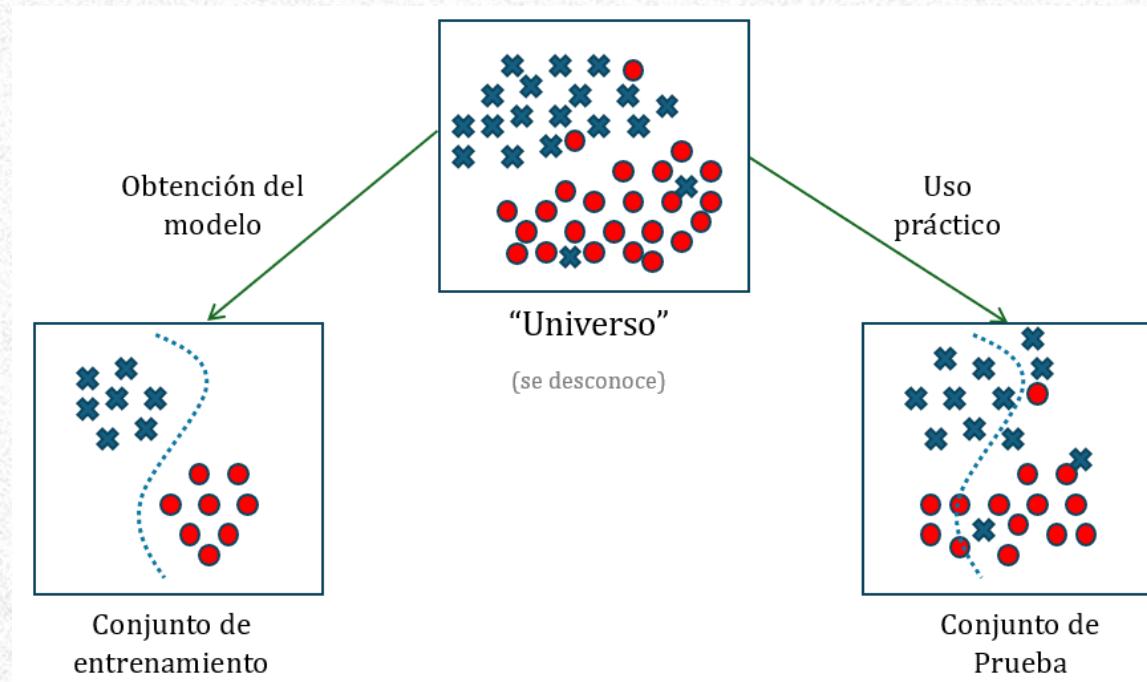
- Prueba (*testing*):
  - El mejor modelo se utiliza para los datos de prueba (*testing*)

# Datos

Los datos tienen diversos formatos

Texto, números, gráficos, tablas, imágenes, videos, etc.

División de datos:



# División de los datos

## Conjunto de entrenamiento (training set)

- Usado para aprender los parámetros del modelo (y determinar el mejor modelo)

## Conjunto de prueba (test set)

- Usado para realizar la prueba de desempeño final
- NO se debe usar este conjunto para “afinar” el modelo entrenado
- Se usa solo al final del proceso

## Conjunto de validación cruzada (cross-validation set)

- Permite “afinar” el modelo (saber qué funciona mejor)
- Es como un conjunto de prueba “falso”
- Datos se dividen en:



# ¿Qué es la regresión lineal?

- La regresión lineal es una técnica de análisis de datos que predice el valor de datos desconocidos mediante el uso de otro valor de datos relacionado y conocido. Modela matemáticamente la variable desconocida o dependiente y la variable conocida o independiente como una ecuación lineal. Por ejemplo, supongamos que tiene datos sobre sus gastos e ingresos del año pasado. Las técnicas de regresión lineal analizan estos datos y determinan que tus gastos son la mitad de tus ingresos. Luego calculan un gasto futuro desconocido al reducir a la mitad un ingreso conocido futuro.

# ¿Por qué es importante la regresión lineal?

- Los modelos de regresión lineal son relativamente simples y proporcionan una fórmula matemática fácil de interpretar para generar predicciones. La regresión lineal es una técnica estadística establecida y se aplica fácilmente al software y a la computación. Las empresas lo utilizan para convertir datos sin procesar de manera confiable y predecible en inteligencia empresarial y conocimiento práctico. Los científicos de muchos campos, incluidas la biología y las ciencias del comportamiento, ambientales y sociales, utilizan la regresión lineal para realizar análisis de datos preliminares y predecir tendencias futuras. Muchos métodos de ciencia de datos, como el machine learning y la inteligencia artificial, utilizan la regresión lineal para resolver problemas complejos.