

# Machine Learning

## Introducción

Christian Oliva Moya  
Pedro Ramón Ventura Gómez



Grado en **Ingeniería Informática** UAM

Máster en **Inteligencia Computacional y Sistemas Interactivos** UAM

Doctorando en **Ingeniería Informática y Telecomunicaciones** UAM

**XAI de redes neuronales profundas**

Profesor Ayudante Dpto. Ingeniería Informática

[christian.oliva@uam.es](mailto:christian.oliva@uam.es)

Director TI en **March Asset Management**

#1 promoción **MIAX** - 7

**Ingeniero Técnico de Telecomunicaciones** UPM

[pventura@march-am.com](mailto:pventura@march-am.com)





**Aprendizaje Automático (Machine Learning)**

**Redes Neuronales (Neural Networks)**

**Aprendizaje Profundo (Deep Learning)**

**Interpretabilidad-Explicabilidad (XAI)**



Investigación



Formación



**Programación (.NET / Python)**

**Gestión de Proyectos**

**Aplicación IA a Finanzas**



Investigación



Formación

# Objetivos del curso

---

## Parte 1: Machine Learning

- Entender los conceptos básicos del Machine Learning (ML)
- Comprender los fundamentos de los algoritmos de ML
- Implementar modelos de ML utilizando la librería **Scikit-Learn**

## Parte 2: Redes Neuronales

- Conocer los fundamentos de las redes neuronales artificiales y el Deep Learning
- Diseñar redes neuronales optimizando los hiperparámetros
- Implementar redes neuronales utilizando las librerías **Tensorflow**, **Keras** y **Pytorch**

# Contenido del curso

---

## Parte 1: Machine Learning

- Introducción al Machine Learning
- Introducción a los modelos de ML: KNN y árboles de decisión
- Evaluación de modelos: métricas y validación
- Preprocesamiento, etiquetado y extracción de características
- Selección de atributos. Reducción de dimensionalidad
- Regresión Lineal
- SVM
- Conjuntos de clasificadores
- Clustering: jerárquico, basado en centroides, mezcla de Gaussianas, densidades

# Contenido del curso

---

## Parte 2: Redes neuronales

- Introducción al Deep Learning
- Definición de una neurona: Regresión lineal y logística
- Arquitectura de una red neuronal básica: Multi-Layer Perceptron
- Aprendizaje: descenso por gradiente y Backpropagation
- Aspectos prácticos del Deep Learning: funciones de activación, funciones de coste, inicialización de pesos, regularización, dropout, técnicas de optimización, etc.
- Tuning de hiperparámetros

# Conocimientos previos

---

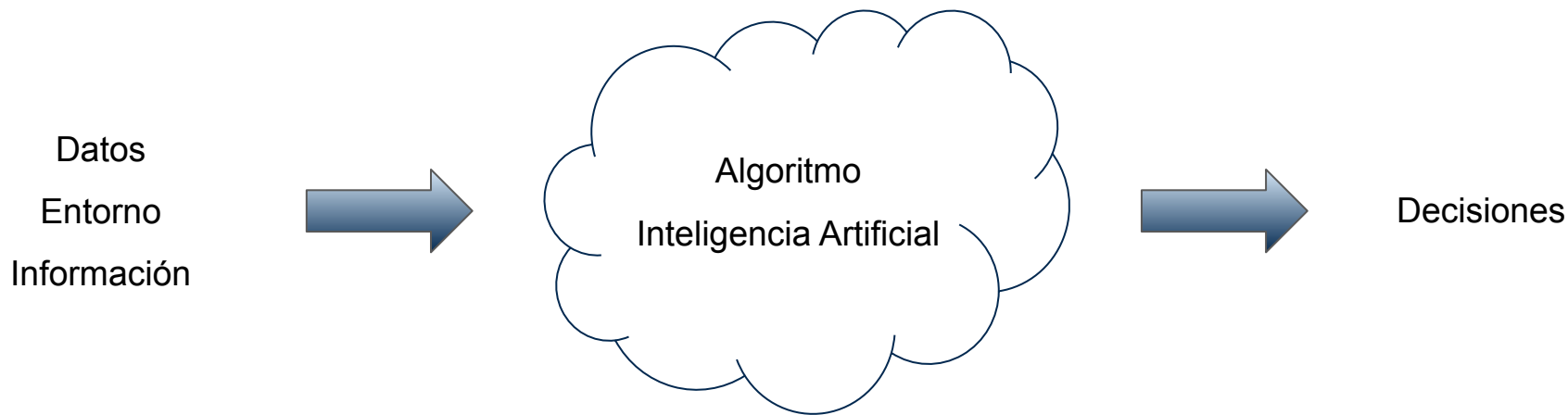
Antes de iniciar el curso, se da por hecho que todos tenéis conocimiento en:

- Cálculo
- Álgebra lineal
- Probabilidad y estadística
- Programación en python
- Manejo de las librerías Pandas, Numpy y Matplotlib



# ¿Qué es la inteligencia artificial?

- La inteligencia artificial (IA) es un campo de la **informática** que se centra en la creación de sistemas que pueden realizar tareas que normalmente requieren **inteligencia humana**.
- La IA se basa en la capacidad de **tomar decisiones automatizadas y lógicas** para abordar problemas y tareas específicas.



# Ejemplo: Inteligencia Artificial (IA)

- Queremos desarrollar un sistema con Inteligencia Artificial para diferenciar salmonetes de sardinas.

¿Cómo lo haríais?



# Ejemplo: Inteligencia Artificial (IA)

---

- Desarrollamos un conjunto de reglas para construir un algoritmo.
  1. Definir el objetivo: *Diferenciar sardinas de salmonetes*
  2. Identificar patrones:



# Ejemplo: Inteligencia Artificial (IA)

---

- Desarrollamos un conjunto de reglas para construir un algoritmo.
  1. Definir el objetivo: *Diferenciar sardinas de salmonetes*
  2. Identificar patrones: *Los salmonetes son rojos y las sardinas no*
  3. Diseñar un conjunto de reglas:



# Ejemplo: Inteligencia Artificial (IA)

- Desarrollamos un conjunto de reglas para construir un algoritmo.
  1. Definir el objetivo: *Diferenciar sardinas de salmonetes*
  2. Identificar patrones: *Los salmonetes son rojos y las sardinas no*
  3. Diseñar un conjunto de reglas:

SI colorPescado ES rojo ENTONCES

DIGO salmonete

SI NO

DIGO sardina



# Ejemplo: Inteligencia Artificial (IA)

- Desarrollamos un conjunto de reglas para construir un algoritmo.

1. Definir el objetivo: *Diferenciar sardinas de salmonetes*

2. Identificar patrones: *Los salmonetes son rojos y las sardinas no*

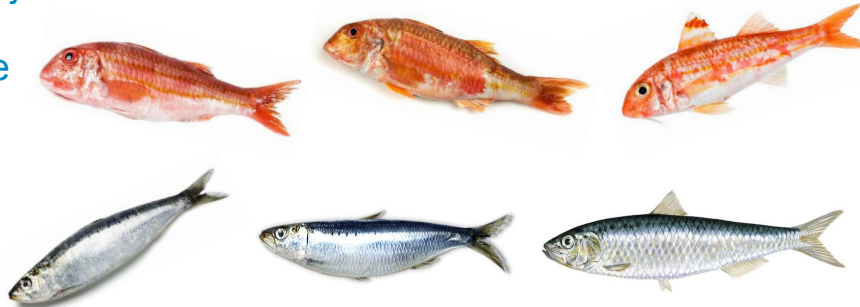
3. Diseñar un conjunto de reglas:

SI colorPescado ES rojo ENTONCES

DIGO salmonete

SI NO

DIGO sardina



# Inteligencia Artificial (IA) vs Aprendizaje Automático (ML)

---

## Inteligencia Artificial (IA)

- Conocimiento experto
- Desarrollo de reglas

## Aprendizaje Automático (ML)

- No es necesario conocimiento experto
- Aprendizaje de las reglas

# Inteligencia Artificial (IA) vs Aprendizaje Automático (ML)

---

## Inteligencia Artificial (IA)

- Conocimiento experto
- Desarrollo de reglas

## Ap ¿Estáis de acuerdo? l

- No es necesario conocimiento experto
- Aprendizaje de las reglas



# Inteligencia Artificial (IA) vs Aprendizaje Automático (ML)

## **Inteligencia Artificial (IA)**

- Conocimiento experto para el desarrollo de las reglas

## **Aprendizaje Automático (ML)**

- No es necesario conocimiento experto para definir las reglas, ya que se aprenden. Sin embargo, ¿cómo son los datos?

# Aprendizaje Automático (ML)

---

- En lugar de programar reglas específicas, el algoritmo aprende por sí mismo los patrones.

## Aprendizaje

*Proceso mediante el cual se adquiere nuevo conocimiento o se modifica el comportamiento a través del ejercicio o la experiencia*

**El programa aprende a partir de los datos**

**El experto prepara correctamente los datos para su procesamiento**

# Aprendizaje Automático (ML)

---

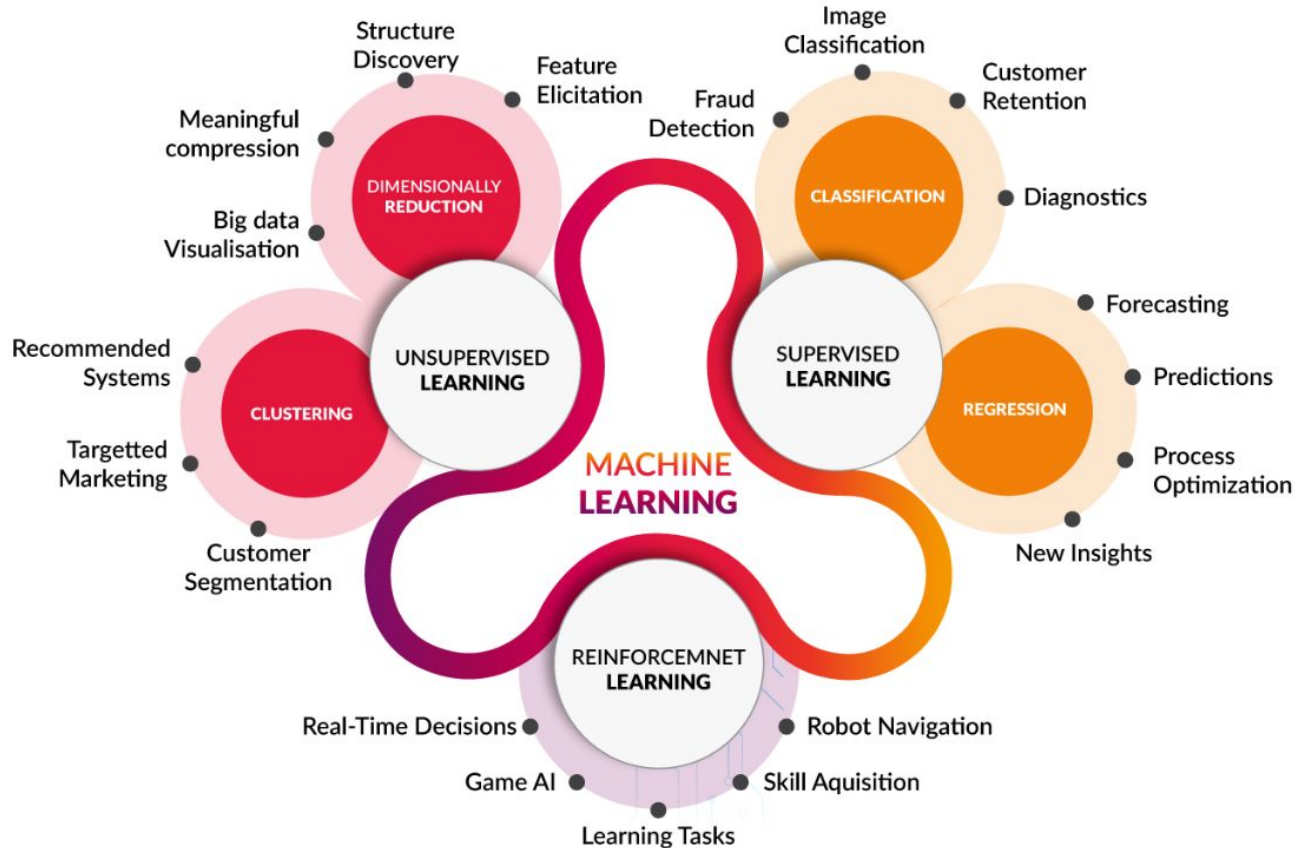
**El programa aprende a partir de los datos**

**El experto prepara correctamente los datos para su procesamiento**

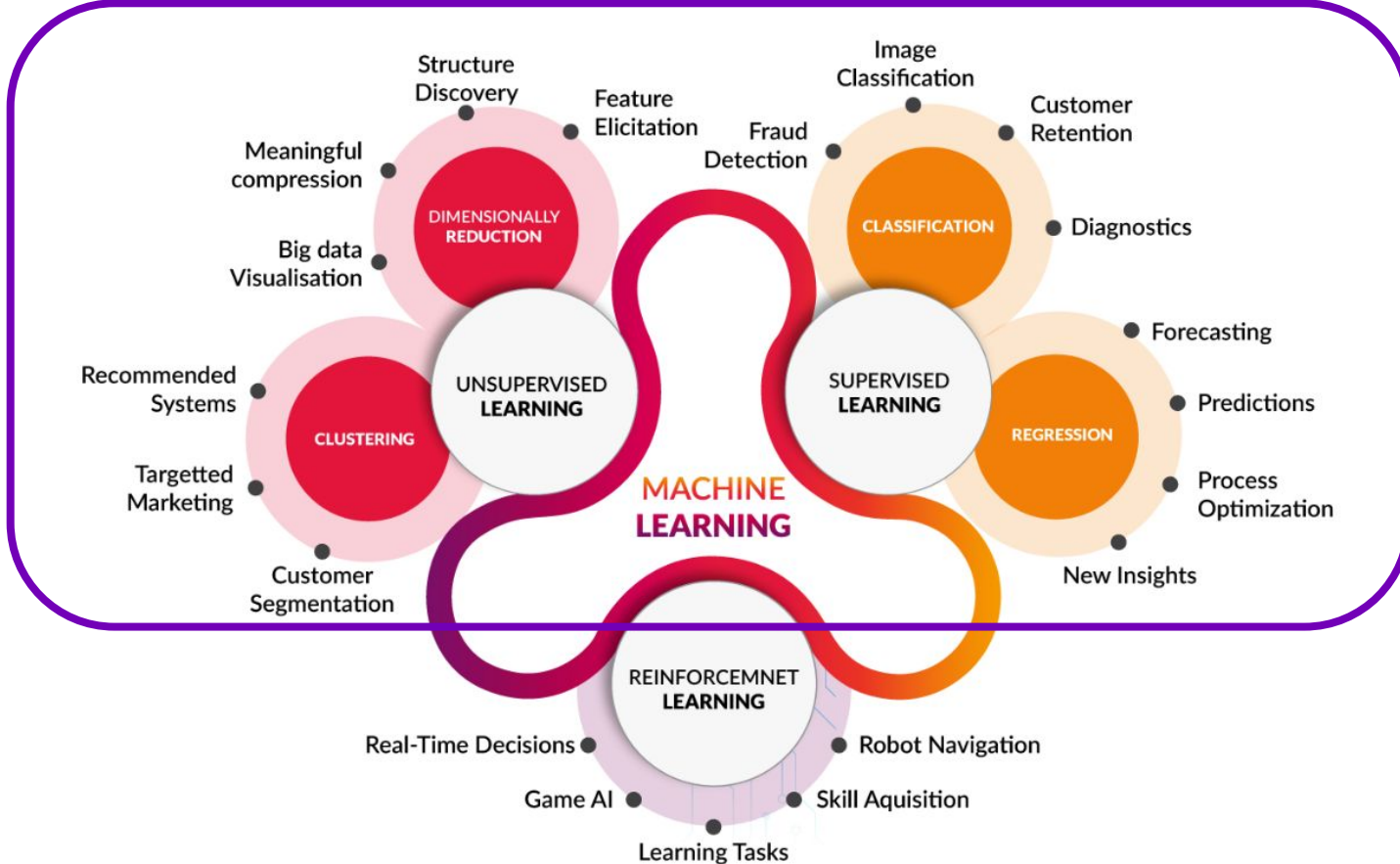
## **Ejemplo**

*Cuando trabajamos con los OHLC de los activos, ¿se utilizan los precios o sus retornos?*

# Aprendizaje Automático (ML)

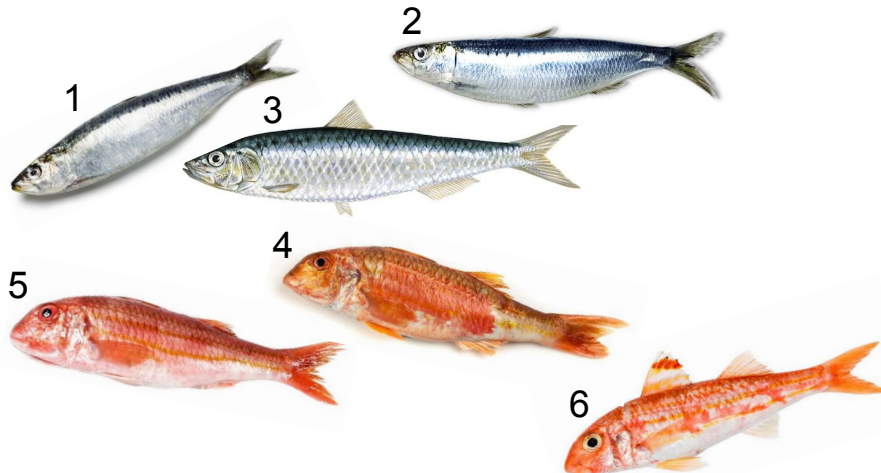


# Aprendizaje Automático (ML)



# Aprendizaje Automático (ML)

- De cada imagen de un pescado, extraemos los dos siguientes atributos:
  - Número de píxeles con un tono azulado
  - Número de píxeles con un tono rojizo



ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332

# Aprendizaje Automático - Supervisado vs No supervisado

- **¿Qué sabemos inicialmente de los datos?**

Solamente conocemos el número de píxeles con tono azul y rojo.

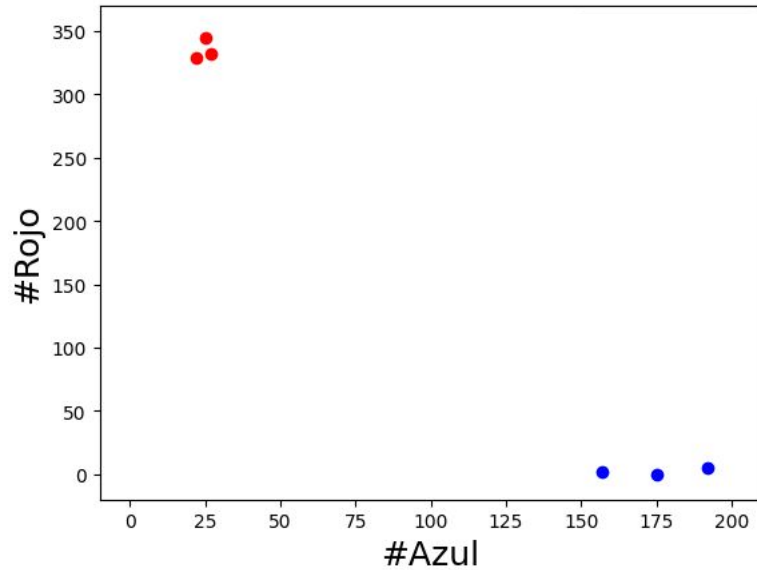
- **¿Qué podemos hacer?**

- ¿Podemos intentar visualizar estos datos?
- ¿Podemos decir si son sardinas o salmonetes?
- ¿Podemos intentar agruparlos?

ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332

# Aprendizaje Automático - Supervisado vs No supervisado

- ¿Podemos intentar visualizar estos datos?

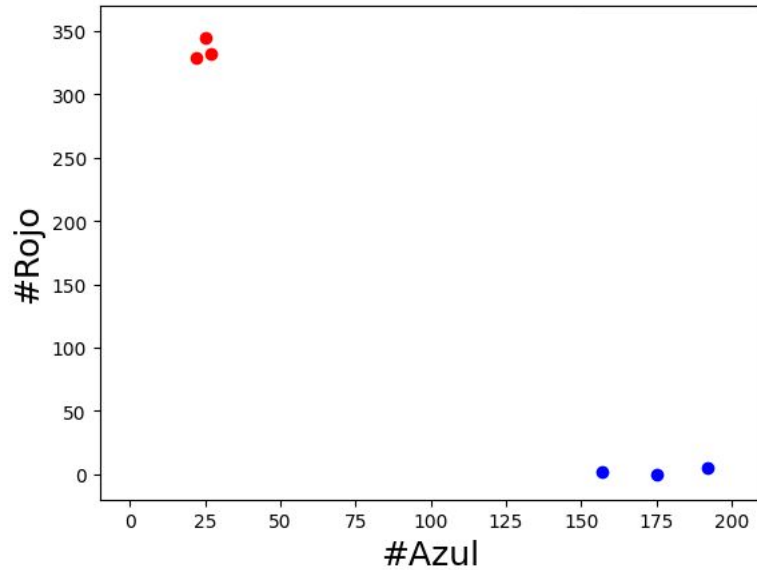


ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332



# Aprendizaje Automático - Supervisado vs No supervisado

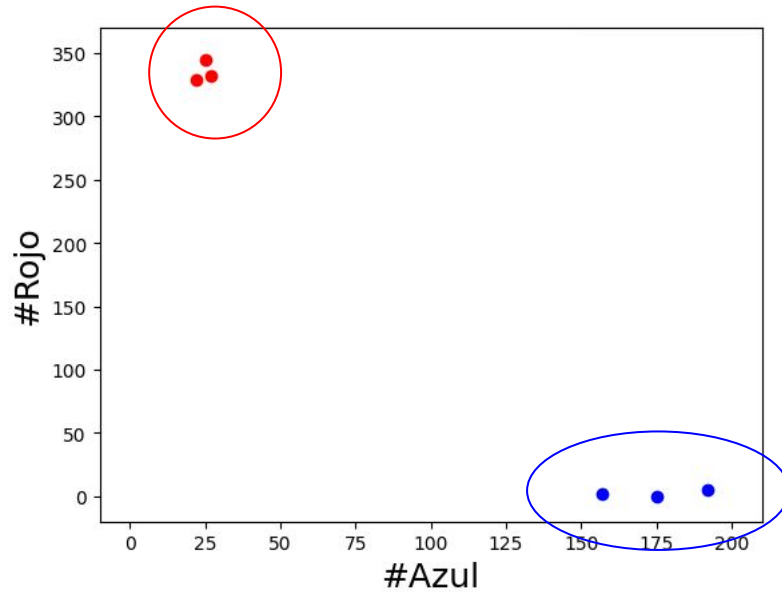
- ¿Podemos decir si son sardinas o salmonetes? No, no sabemos lo que son.



ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332

# Aprendizaje Automático - Supervisado vs No supervisado

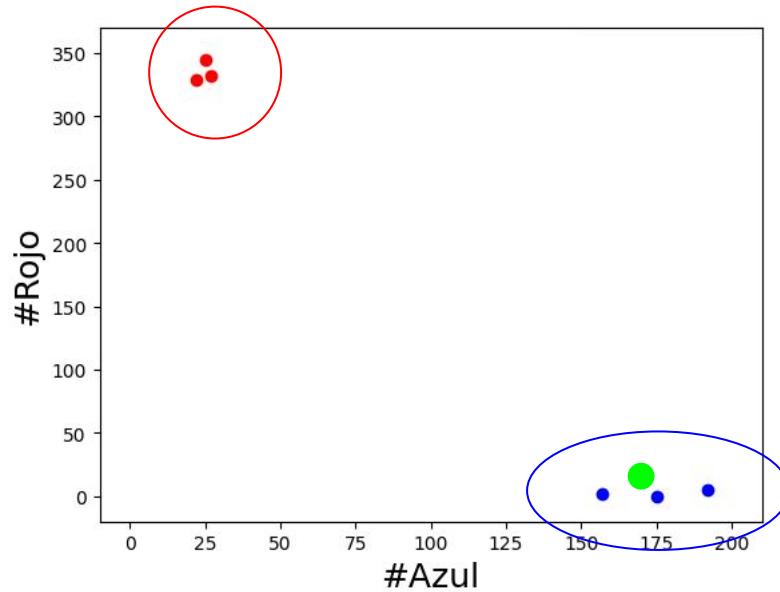
- ¿Podemos intentar agruparlos? Sí.



ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332

# Aprendizaje Automático - Supervisado vs No supervisado

- Si viene un nuevo dato, ¿a qué grupo pertenece?



ID	#Azul	#Rojo
7	162	22

ID	#Azul	#Rojo
1	175	0
2	192	5
3	157	2
4	25	345
5	22	329
6	27	332

# Aprendizaje Automático - Supervisado vs No supervisado

- **Aprendizaje No Supervisado: Agrupación (Clustering)**

Si no tenemos más información que su aspecto, podemos agrupar los datos según patrones similares.

Azules



Rojos



# Aprendizaje Automático - Supervisado vs No supervisado

- ¿Qué sabemos inicialmente de los datos?

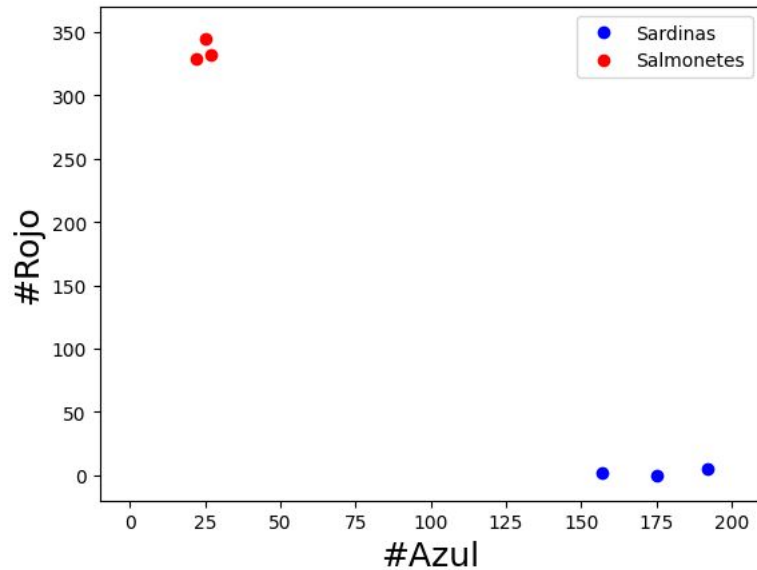
Ahora también conocemos qué es cada pescado.

- ¿Qué podemos hacer?
  - ¿Podemos intentar visualizar estos datos?
  - ¿Podemos decir si son sardinas o salmonetes?
  - ¿Podemos intentar agruparlos?

ID	#Azul	#Rojo	Clase
1	175	0	Sardina
2	192	5	Sardina
3	157	2	Sardina
4	25	345	Salmonete
5	22	329	Salmonete
6	27	332	Salmonete

# Aprendizaje Automático - Supervisado vs No supervisado

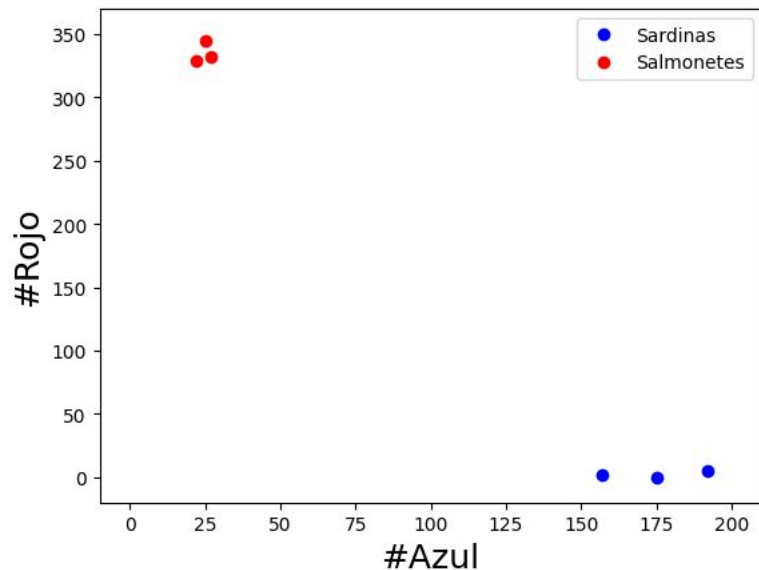
- ¿Podemos intentar visualizar estos datos?



ID	#Azul	#Rojo	Clase
1	175	0	Sardina
2	192	5	Sardina
3	157	2	Sardina
4	25	345	Salmonete
5	22	329	Salmonete
6	27	332	Salmonete

# Aprendizaje Automático - Supervisado vs No supervisado

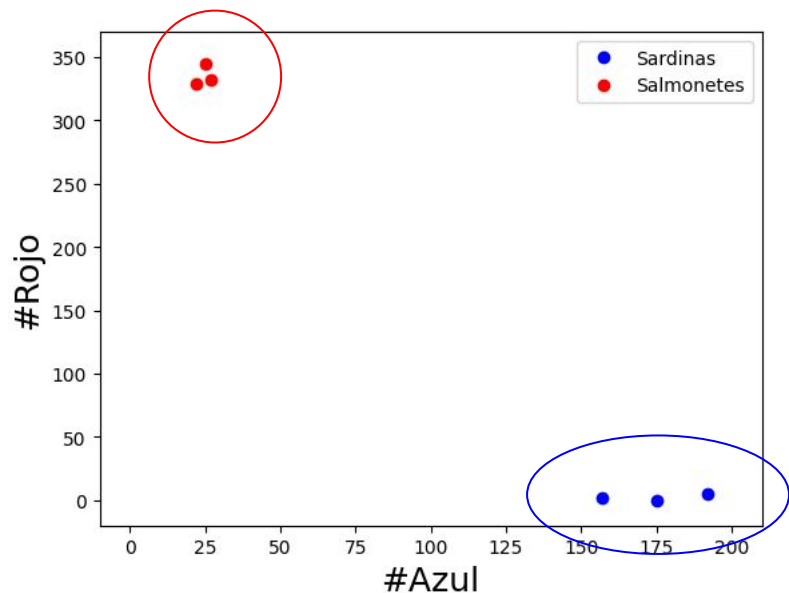
- ¿Podemos decir si son sardinas o salmonetes? Sí.



ID	#Azul	#Rojo	Clase
1	175	0	Sardina
2	192	5	Sardina
3	157	2	Sardina
4	25	345	Salmonete
5	22	329	Salmonete
6	27	332	Salmonete

# Aprendizaje Automático - Supervisado vs No supervisado

- ¿Podemos intentar agruparlos? Sí, pero, ¿para qué, si ya sabemos lo que son?

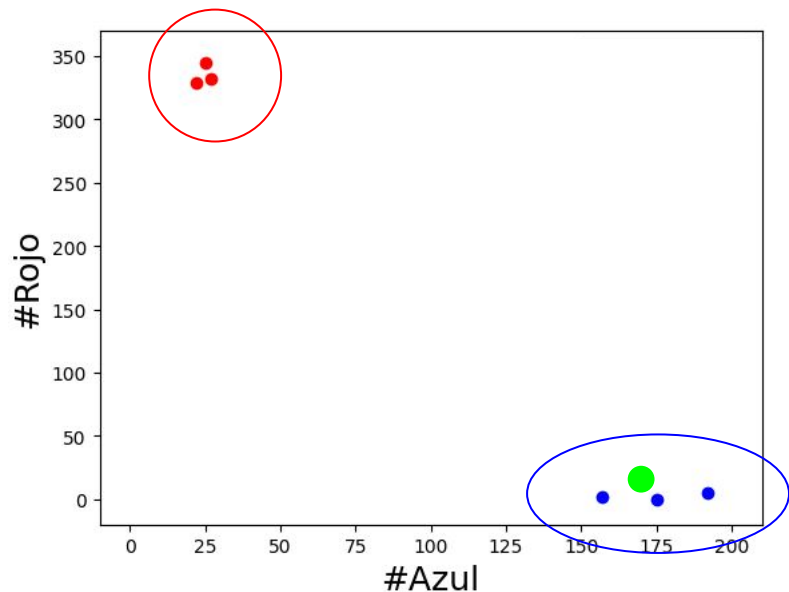


ID	#Azul	#Rojo	Clase
1	175	0	Sardina
2	192	5	Sardina
3	157	2	Sardina
4	25	345	Salmonete
5	22	329	Salmonete
6	27	332	Salmonete



# Aprendizaje Automático - Supervisado vs No supervisado

- Si viene un nuevo dato, ¿qué es?



ID	#Azul	#Rojo
7	162	22

ID	#Azul	#Rojo	Clase
1	175	0	Sardina
2	192	5	Sardina
3	157	2	Sardina
4	25	345	Salmonete
5	22	329	Salmonete
6	27	332	Salmonete

# Aprendizaje Automático - Supervisado vs No supervisado

- **Aprendizaje Supervisado**

Si sabemos qué son los datos, es decir, conocemos su **clase**, podemos identificarlos.

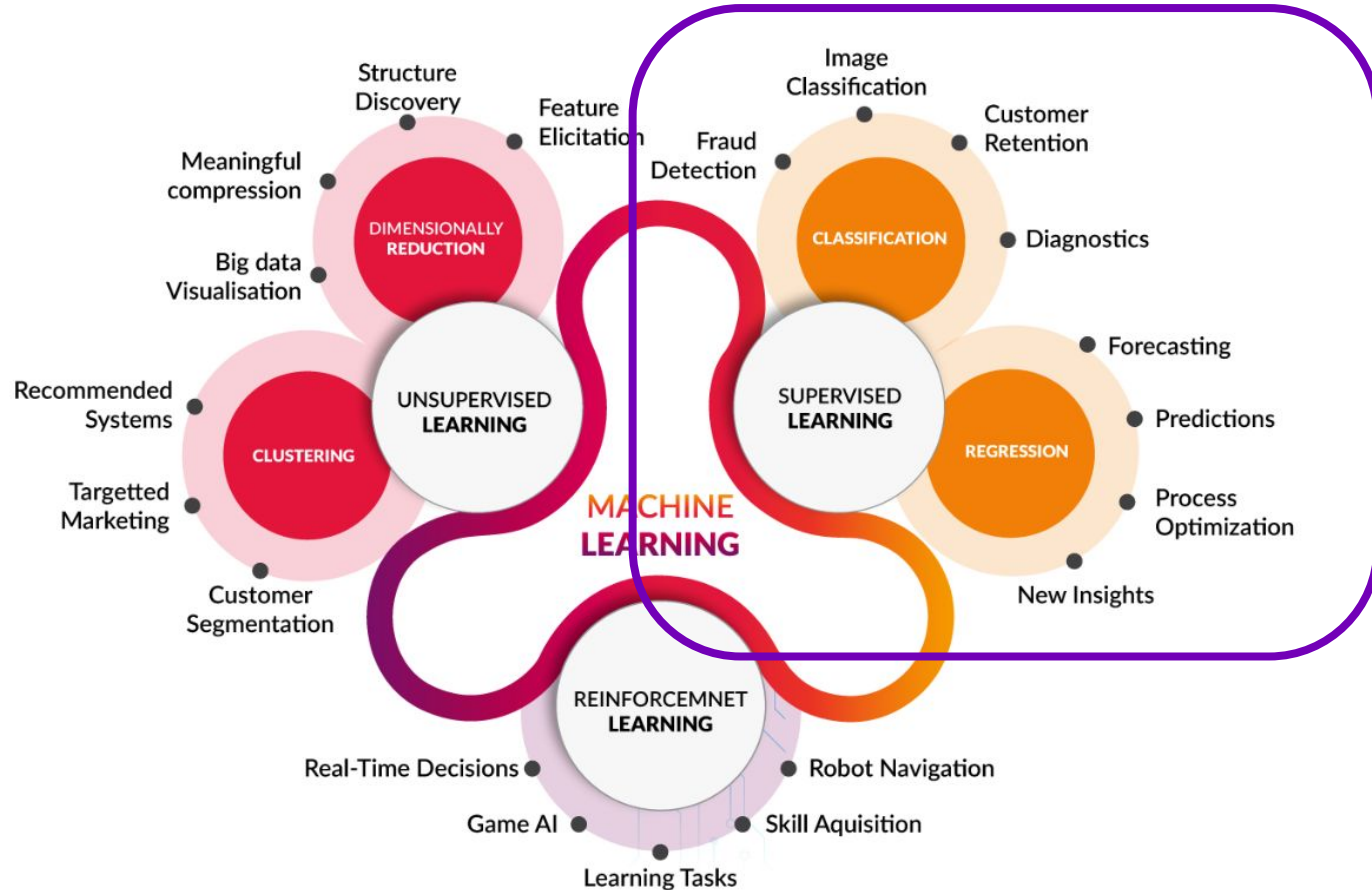
Sardinas



Salmonetes



# Aprendizaje Automático (ML)



# ML Supervisado - Regresión vs Clasificación

---

## Clasificación

- Predecir la **clase** de un dato
- Responde a la pregunta:

¿Qué eres?

- Devuelve una probabilidad  $[0, 1]$

## Regresión

- Predecir el **valor** de un dato
- Responde a la pregunta:

¿Cuánto vales?

- Devuelve un valor Real  $[-\infty, \infty]$

# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Diferenciar sardinas de salmonetes

¿Regresión o Clasificación?



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Diferenciar sardinas de salmonetes

## Clasificación

- Respondo a la pregunta:

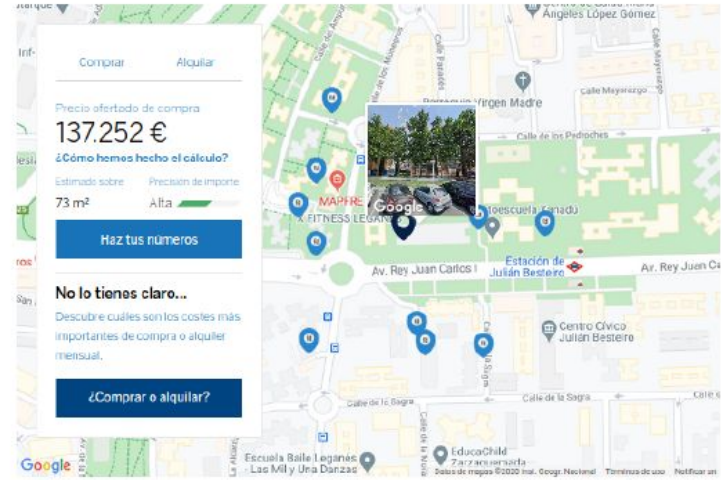
¿Qué eres?



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Estimación del valor de un inmueble

¿Regresión o Clasificación?



Vivienda		Barrio	
Características de la vivienda			
Conoce los datos de esta vivienda según la Dirección General del Catastro.	TIPO DE VIVIENDA	SUPERFICIE VIVIENDA	CONSTRUIDA EN
	Piso	68 m²	1974
	USO	SUPERFICIE TOTAL	
	Residencial	73 m²	
<a href="#">¿Cómo calculamos estos valores?</a>			

fuelle: BBVA Valora

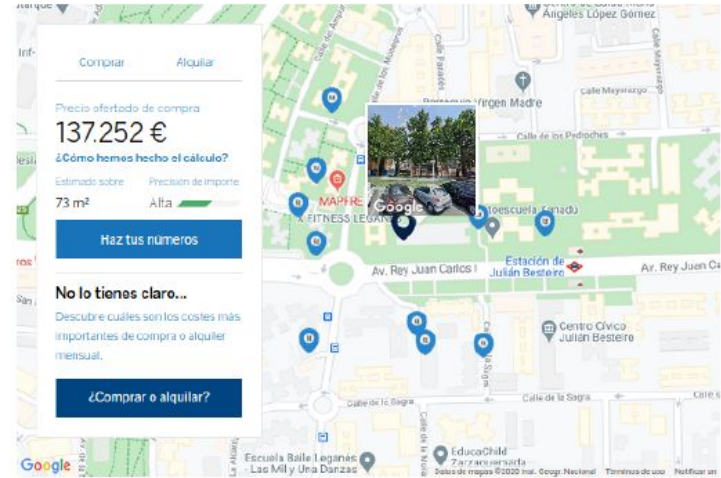
# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Estimación del valor de un inmueble

## Regresión

- Responde a la pregunta:

**¿Cuánto vales?**



Vivienda

Barrio

### Características de la vivienda

Conoce los datos de esta vivienda según la Dirección General del Catastro.

TIPO DE VIVIENDA

Piso

SUPERFICIE VIVIENDA

68 m²

CONSTRUIDA EN

1974

USO

Residencial

SUPERFICIE TOTAL

73 m²

¿Cómo calculamos estos valores?

fuelle: BBVA Valora



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Predecir si una empresa quebrará el próximo año

¿Regresión o Clasificación?



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Predecir si una empresa quebrará el próximo año

## **Clasificación**

- Responde a la pregunta:

**¿Qué eres?**

*Das una probabilidad de quebrar*



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Tomar una decisión de compra-venta a partir de OHCL

¿Regresión o Clasificación?



# ML Supervisado - Regresión vs Clasificación - Ejemplos

- Tomar una decisión de compra-venta a partir de OHCL

## Clasificación

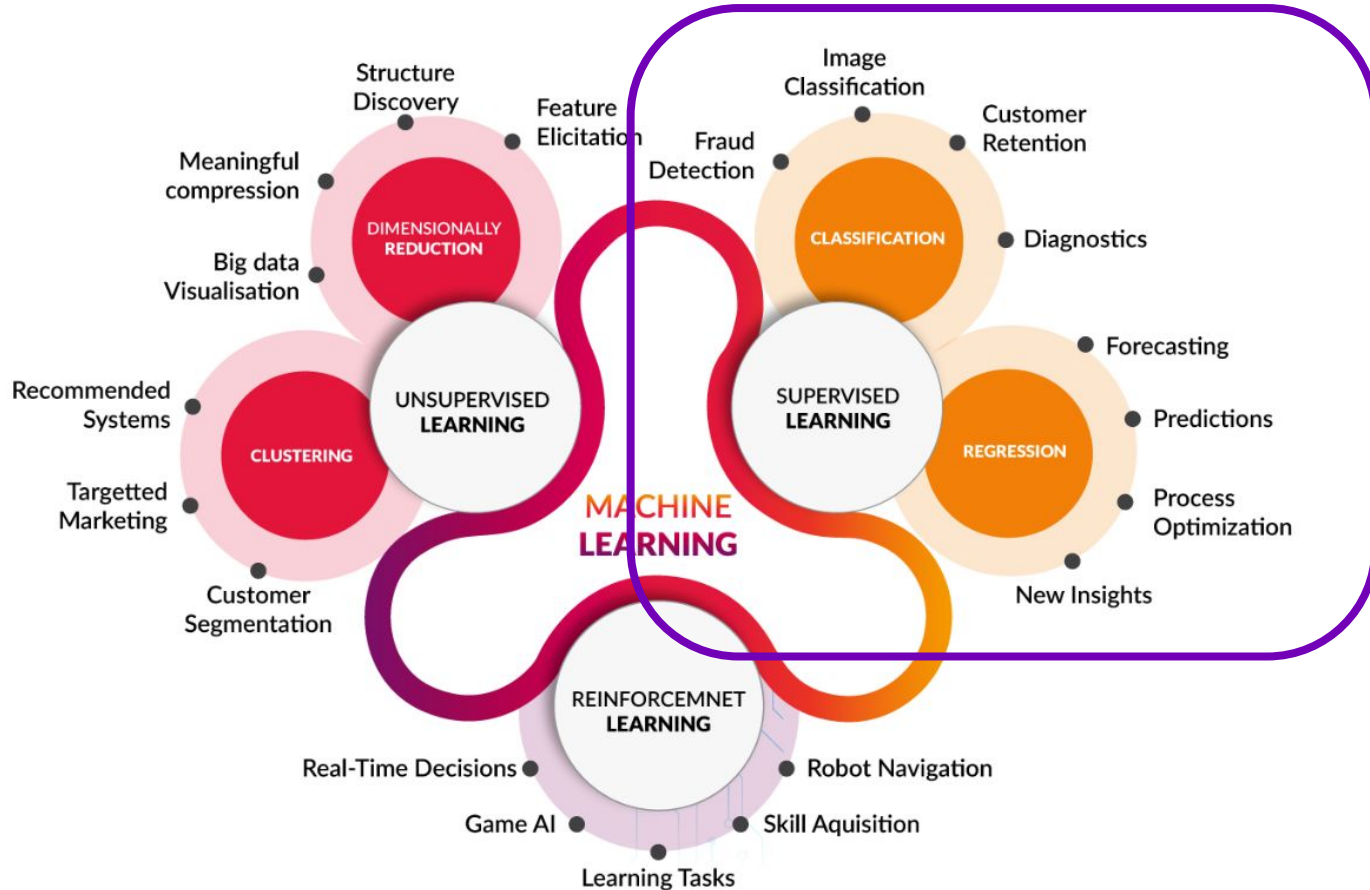
- Responde a la pregunta:

¿Qué eres?

*Eres compra o eres venta*



# Aprendizaje Automático (ML)



# ML Supervisado - Paramétricos vs No Paramétricos

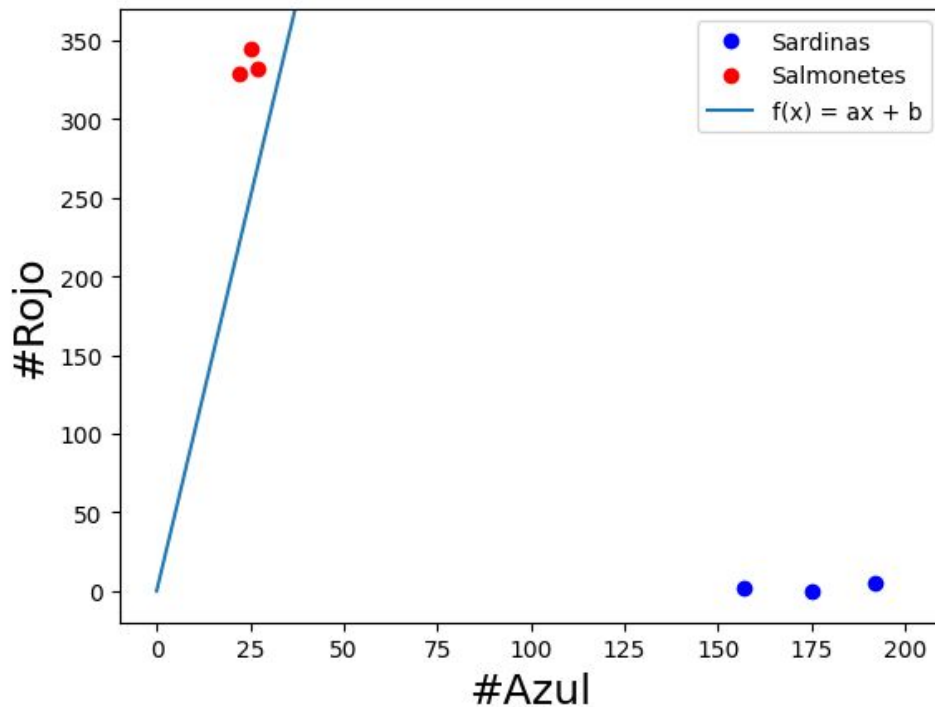
Volviendo al ejemplo con los peces, definimos una **frontera de decisión** con una línea recta:

- ¿Qué forma tiene una línea recta?

$$f(x) = a \cdot x + b$$

Si  $a = 10.0$  y  $b = 0.0$

- Todo lo que hay **por encima de la recta** serán **salmonetes** y todo lo que hay **por debajo de la recta** serán **sardinas**.



# ML Supervisado - Paramétricos vs No Paramétricos

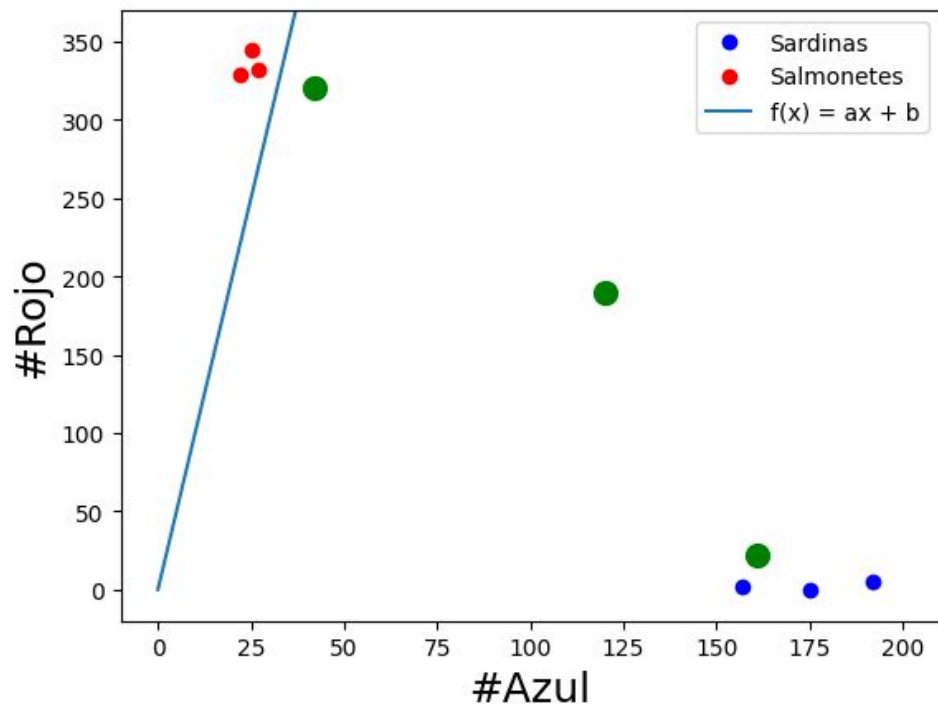
Volviendo al ejemplo con los peces, definimos una frontera de decisión con una línea recta:

- ¿Qué forma tiene una línea recta?

$$f(x) = a \cdot x + b$$

Si  $a = 10.0$  y  $b = 0.0$

- Todo lo que hay **por encima de la recta** serán **salmonetes** y todo lo que hay **por debajo de la recta** serán **sardinas**.
- ¿Qué pasa con los puntos verdes?



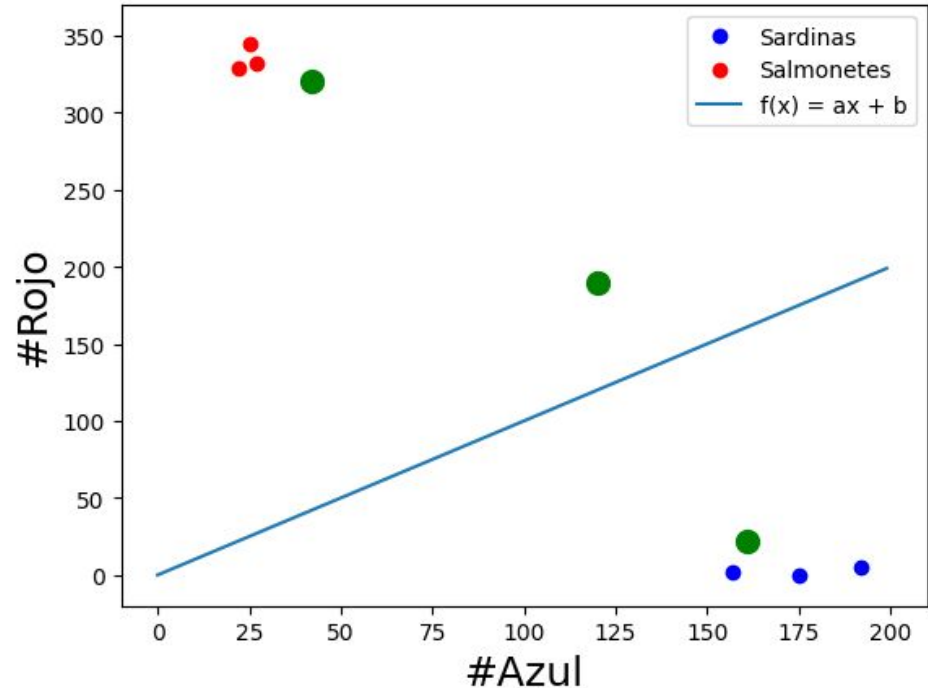
# ML Supervisado - Paramétricos vs No Paramétricos

- Si cambiamos los parámetros  $a$  y  $b$

$$f(x) = a \cdot x + b$$

Si  $a = 1.0$  y  $b = 0.0$

- ¿Qué pasa ahora con los puntos verdes?





# ML Supervisado - Paramétricos vs No Paramétricos

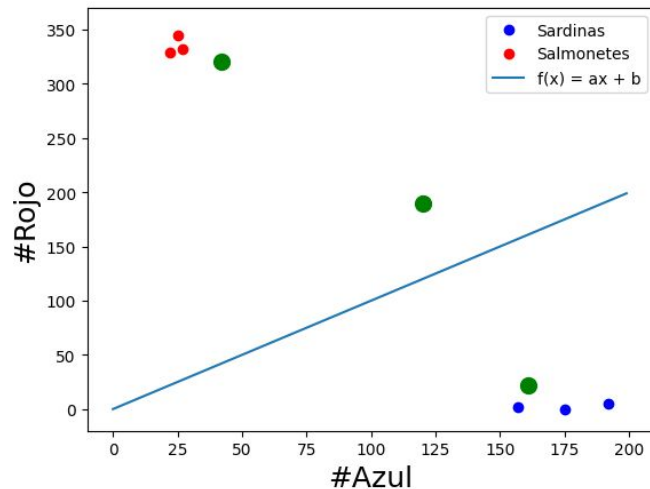
- ¿Qué forma tiene una línea recta?

$$f(x) = a \cdot x + b$$

Si  $a = 1.0$  y  $b = 0.0$

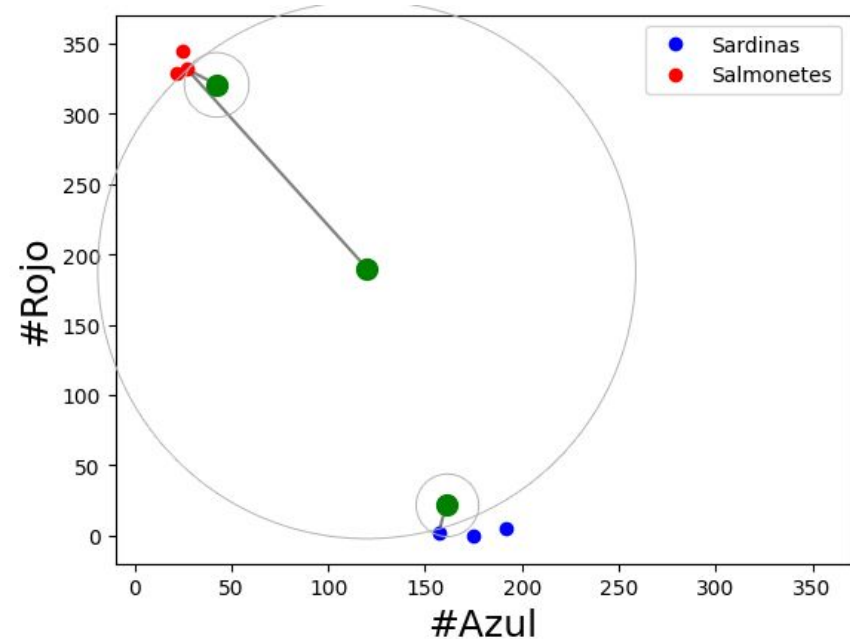
La frontera de decisión de este modelo **depende** de los valores de  $a$  y  $b$  para ser capaz de separar los puntos rojos de los azules con una **línea recta**. Por tanto:

- Es un **clasificador lineal** (separa con una línea recta)
- Es **paramétrico** (depende de  $a$  y  $b$ )



# ML Supervisado - Paramétricos vs No Paramétricos

- Asumamos lo siguiente: Un punto nuevo será lo mismo que a lo que más se parezca (**similitud**)
- Fijémonos en el **punto más cercano** en el plano para clasificar:
- ¿Cómo es la frontera de decisión?

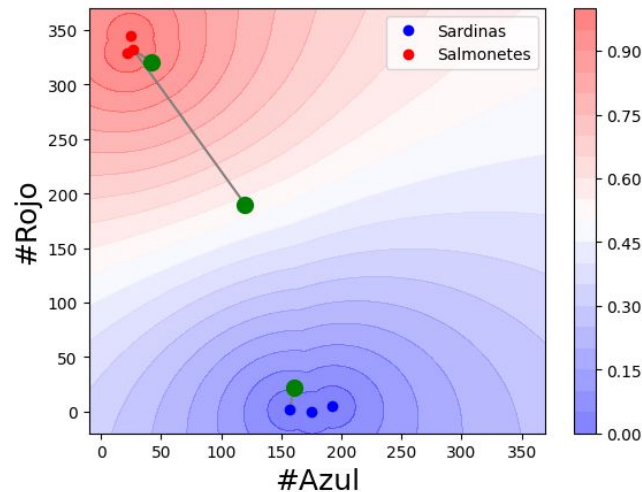


# ML Supervisado - Paramétricos vs No Paramétricos

- Asumamos lo siguiente: Un punto nuevo será lo mismo que a lo que más se parezca (similitud)
- Fijémonos en el punto más cercano en el plano para clasificar:

Este modelo **no tiene** ningún parámetro ajustable. Es capaz de partir el espacio de muestra para separar los puntos rojos de los azules considerando el punto más cercano. Por tanto:

- Es un **clasificador no lineal**
- Es **no paramétrico** (se basa en la similitud)



# ML Supervisado - Paramétricos vs No Paramétricos

---

## Paramétricos

- Número fijo de parámetros  
Se ajustan durante el entrenamiento
- Requiere una fase de entrenamiento  
Suele ser lenta
- Predicción rápida  
Depende de los parámetros del modelo

## No Paramétricos

- No tienen parámetros
- Utilizan diversas técnicas  
Búsqueda, similitud, partición, etc.
- No hay entrenamiento
- Predicción lenta  
Depende de los datos

# ML Supervisado - Paramétricos vs No Paramétricos

---

## Paramétricos

- Regresión lineal o logística
- Redes neuronales
- SVM

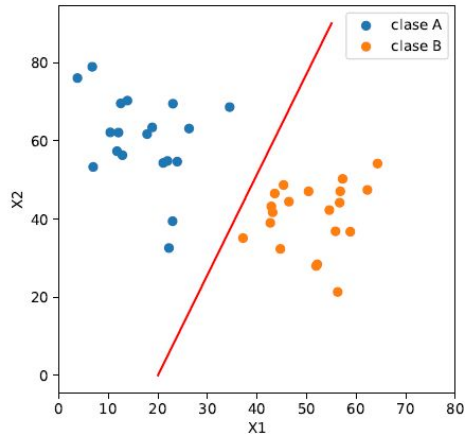
## No Paramétricos

- Árboles de decisión
- K-vecinos cercanos (KNN)

# ML Supervisado - Clasificador Lineal vs No Lineal

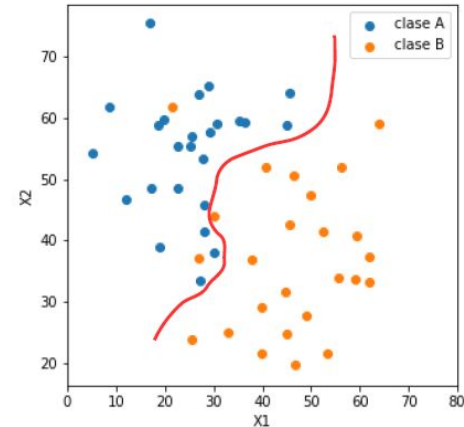
## Lineal

- Separa el espacio con una función lineal (línea recta)



## No Lineal

- Separa el espacio con algo que no es una línea recta



# Alcance Inter-disciplinar

---

