



# Máster Smart Energy

Postgrado de DIGITAL ENERGY

Curso 2022/2023

## Sesión 7 – Modelos de Aprendizaje NO Supervisado

Sara Barja Martínez: [sara.barja@upc.edu](mailto:sara.barja@upc.edu)



# Calendario

	Lunes	Martes	Miércoles	Jueves
ABRIL	10	11	12 S1 – Introducción a Machine Learning	13
	17 S2 – Introducción a Python	18	19 S3 – Estadística descriptiva	20
	24 S4 – Modelos aprendizaje supervisado (I): Clasificación	25	26 S5 – Modelos de aprendizaje supervisado (II): Regresión	27
MAYO	1	2	3 S6 – Aplicación de AI en el sector eléctrico: Odit-e	4
	8 S7 – Modelos de aprendizaje no supervisado	9	10 S8 – Examen final	



## Objetivos de la sesión

- Diferenciar el aprendizaje **Supervisado** y **No Supervisado**.
- Conocer los algoritmos utilizados en **Aprendizaje No Supervisado**.
- **Aplicaciones** del clustering.
- **Métricas** para encontrar el número óptimo de clusters.
- Introducción a **Reinforcement Learning** y sus aplicaciones.
- Repaso general.



## **Diferenciar el aprendizaje Supervisado y No Supervisado.**

Algoritmos utilizados en Aprendizaje No Supervisado.

**Métricas** para encontrar el número óptimo de clusters.

Aplicaciones del clustering.

Introducción a **Reinforcement Learning** y sus aplicaciones.



Recapitulemos...

## MACHINE LEARNING

Aprendizaje  
Supervisado

Regresión

Clasificación

Aprendizaje NO  
Supervisado

Clustering

Reducir dimensiones

Reinforcement  
Learning



## Principales diferencias entre aprendizaje supervisado y NO supervisado

### Aprendizaje Supervisado

### Aprendizaje NO Supervisado

Datos de entrada

Hay datos etiquetados (target)

**No hay** datos etiquetados  
(**no** target)

Objetivo

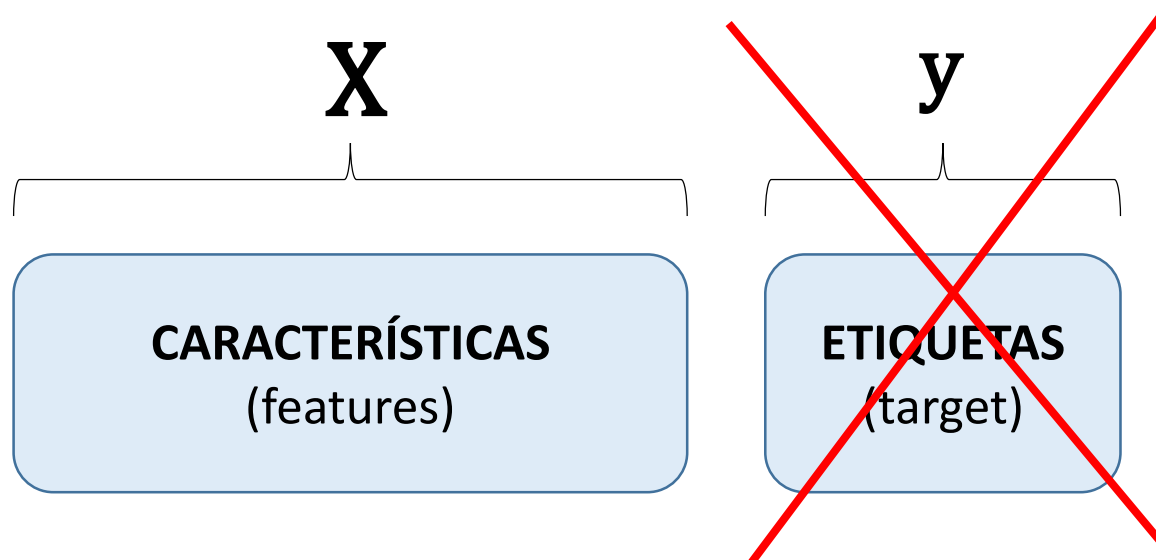
Realizar predicciones de  
clasificación/regresión.

Encontrar patrones ocultos en  
los datos.  
Reducir dimensiones



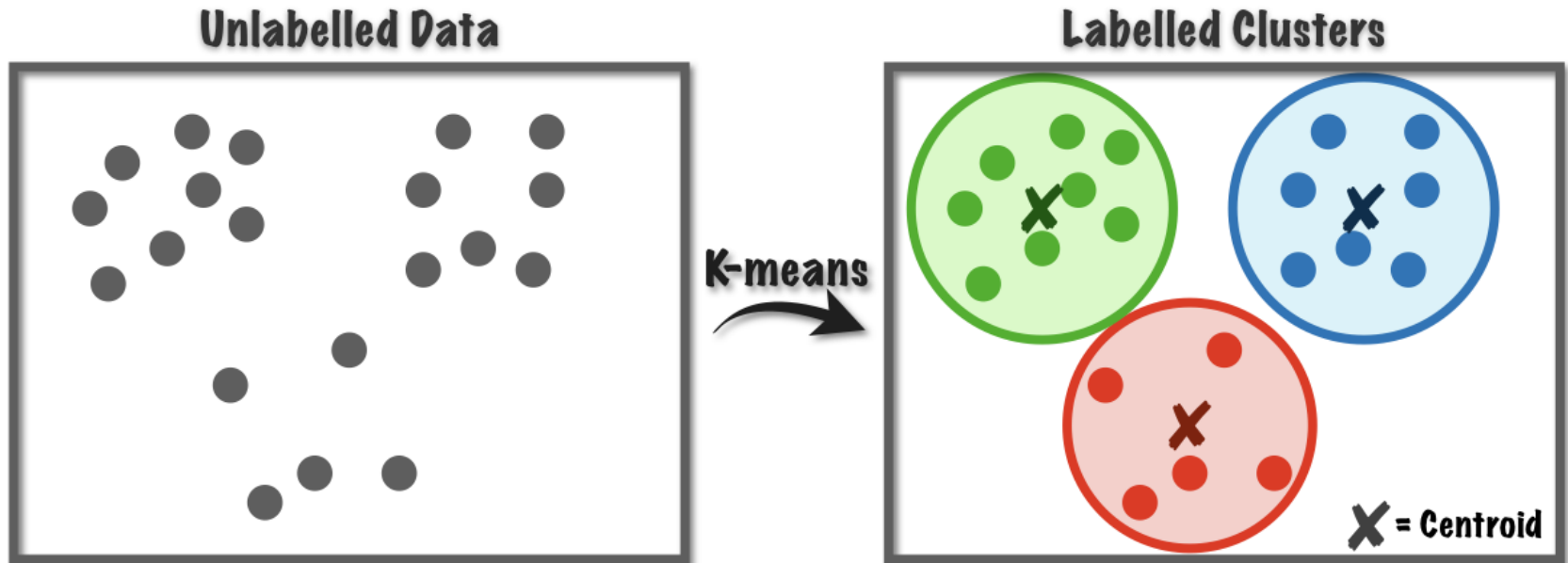
## ¿Qué es el aprendizaje NO supervisado?

El objetivo es entrenar un modelo para **encontrar patrones** en un conjunto de datos sin etiqueta.





¿Qué es el aprendizaje NO supervisado?

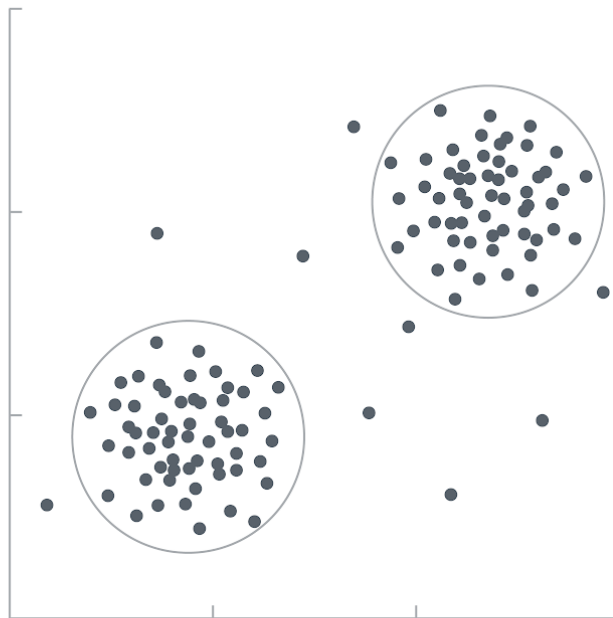




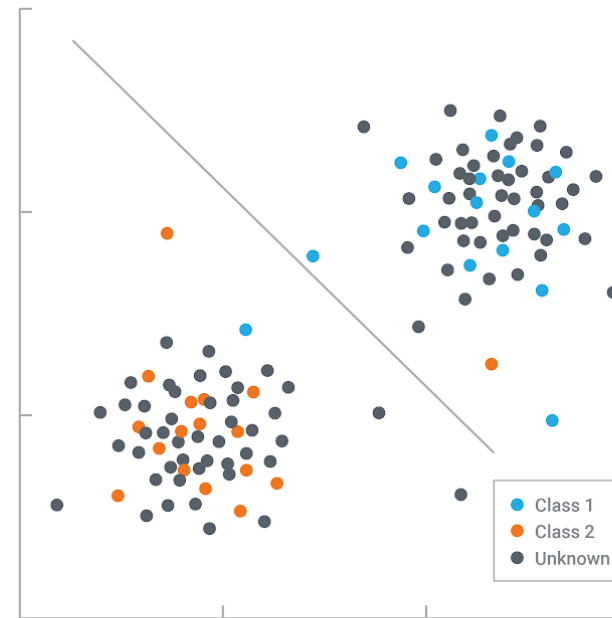


## ¿Qué es el aprendizaje NO supervisado?

**UNSUPERVISED**



**SUPERVISED**





Diferenciar el aprendizaje Supervisado y No Supervisado.

**Algoritmos utilizados en Aprendizaje No Supervisado.**

Métricas para encontrar el número óptimo de clusters.

**Aplicaciones** del clustering.

Introducción a **Reinforcement Learning** y sus aplicaciones.



## Introducción al aprendizaje no supervisado

¿Qué abarca el aprendizaje no supervisado?

### Clustering

- Algoritmos de partición (K-means).
- Algoritmos basados en la jerarquía (Hierarchical).
- Algoritmos basados en la densidad (density).

**Reducción de dimensiones (Análisis de componentes principales (PCA))**



## Introducción al aprendizaje no supervisado

¿Qué abarca el aprendizaje no supervisado?

### Clustering

- Algoritmos de partición (K-means).
- Algoritmos basados en la jerarquía (Hierarchical).
- Algoritmos basados en la densidad (density).

Reducción de dimensiones (Análisis de componentes principales (PCA))



## Clustering

**Objetivo:** agrupar personas o elementos basándonos en sus características.

**Definición:** Es la tarea de dividir el conjunto de datos de entrada (dataset) en varios grupos, de manera que los datos que pertenecen a un mismo grupo sean parecidos entre ellos y diferentes de los datos de otros grupos.

- Sirve para comprender la estructura de los datos, o como preprocesamiento para el aprendizaje supervisado (por ejemplo entrenar un clasificador para cada grupo).
- El clústering no necesita etiquetas, que suelen ser costosas de obtener.



sample



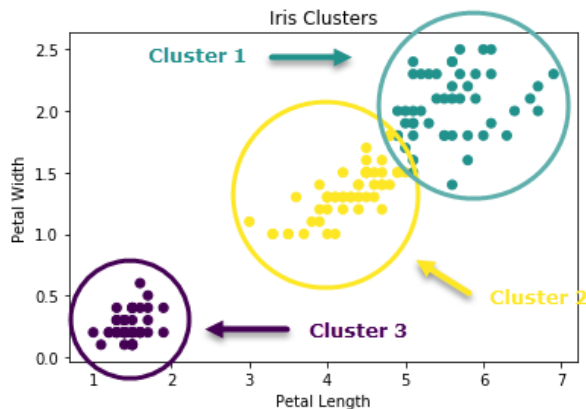
Cluster/group



## Algoritmos de Clustering

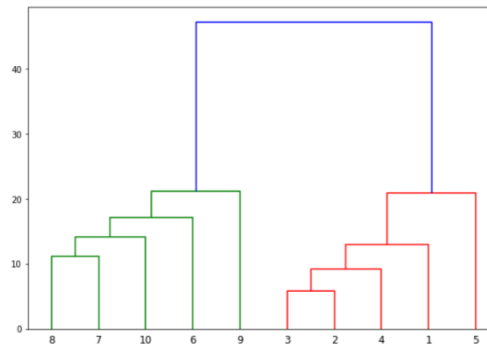
### Algoritmos de partición

- **K-means**
- Este algoritmo divide las observaciones en ***k clusters*** (grupos) y cada partición forma un cluster.



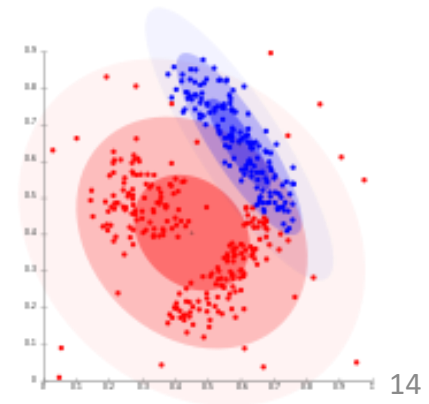
### Algoritmos basados en la jerarquía

- **Hierarchical Clustering**
- Los cúmulos forman una estructura tipo árbol basada en la jerarquía.
- Se divide en dos categorías:
  - **Aglomerado**
  - **Divisivo**



### Algoritmos basados en la densidad

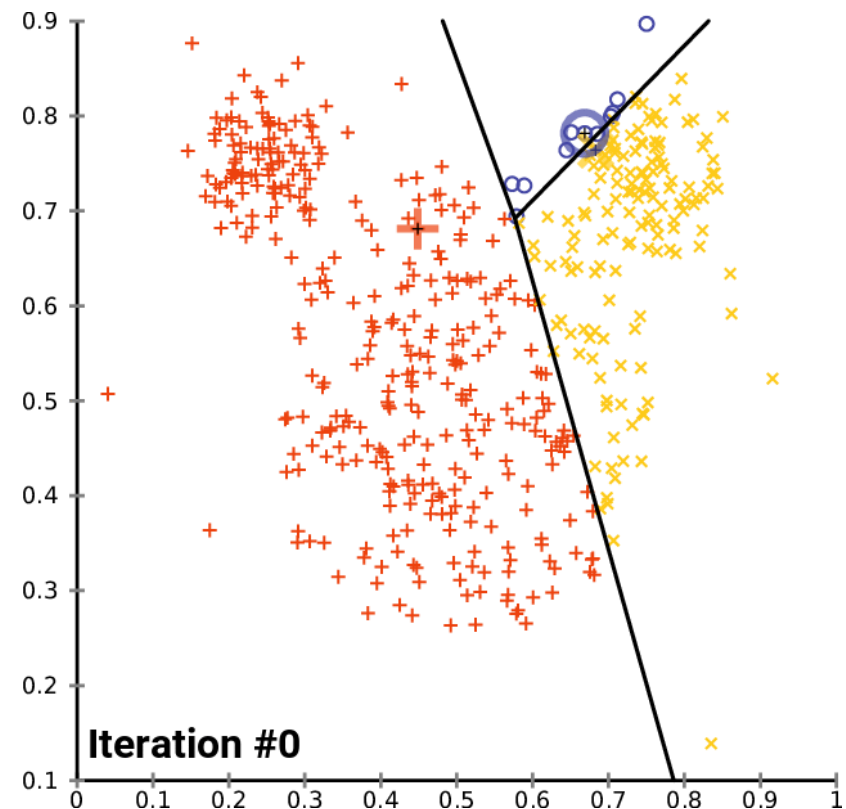
- **DBSCAN, OPTICS**
- Estos algoritmos consideran que los clusters son una región densa que tiene cierta similitud y son diferentes de la región más baja y densa del espacio.





## Método de partición: K-means

- Agrupa las observaciones en  $k$  clusters (el valor de  $K$  debe ser especificado previamente).
- El algoritmo k-means hace lo siguiente:
  - Determina de forma iterativa los mejores puntos centrales  $K$  (conocidos como **centroides**).
  - Asigna cada observación al centroide más cercano. Los ejemplos más cercanos al mismo centroide pertenecen al mismo grupo.
- El algoritmo k-means selecciona las ubicaciones del centroide para minimizar el cuadrado acumulativo de las distancias desde cada ejemplo hasta su centroide más cercano.

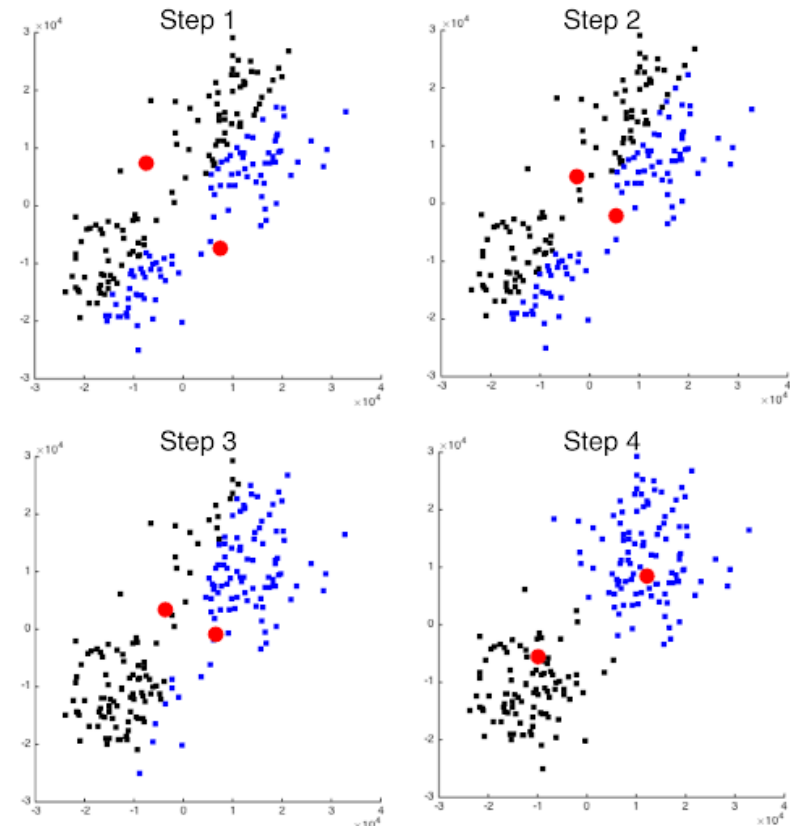




## Método de partición: K-means

### Pasos a seguir

- Para **K** número de clusters, elige un centroide al azar.
- Calcula la distancia a los otros puntos del centroide.
- Ajustar el centroide iterativamente, y asigna cada punto de datos al centroide más cercano.
- Continúa ajustando hasta que el centroide no se mueva.







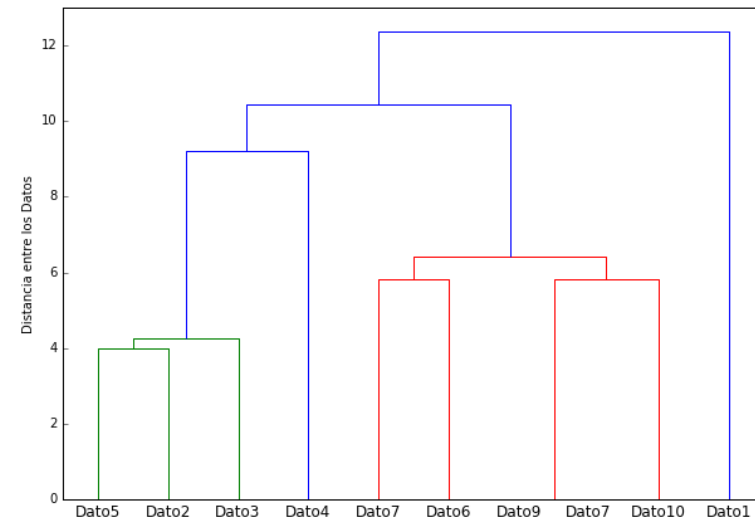
## Método jerárquico: clustering jerárquico

El algoritmo de clúster jerárquico agrupa los datos basándose en la distancia entre cada uno y buscando que los datos que están dentro de un clúster sean los más similares entre sí.

**Tipos de clustering jerárquico** dependiendo de la dirección en la que el algoritmo ejecute el agrupamiento:

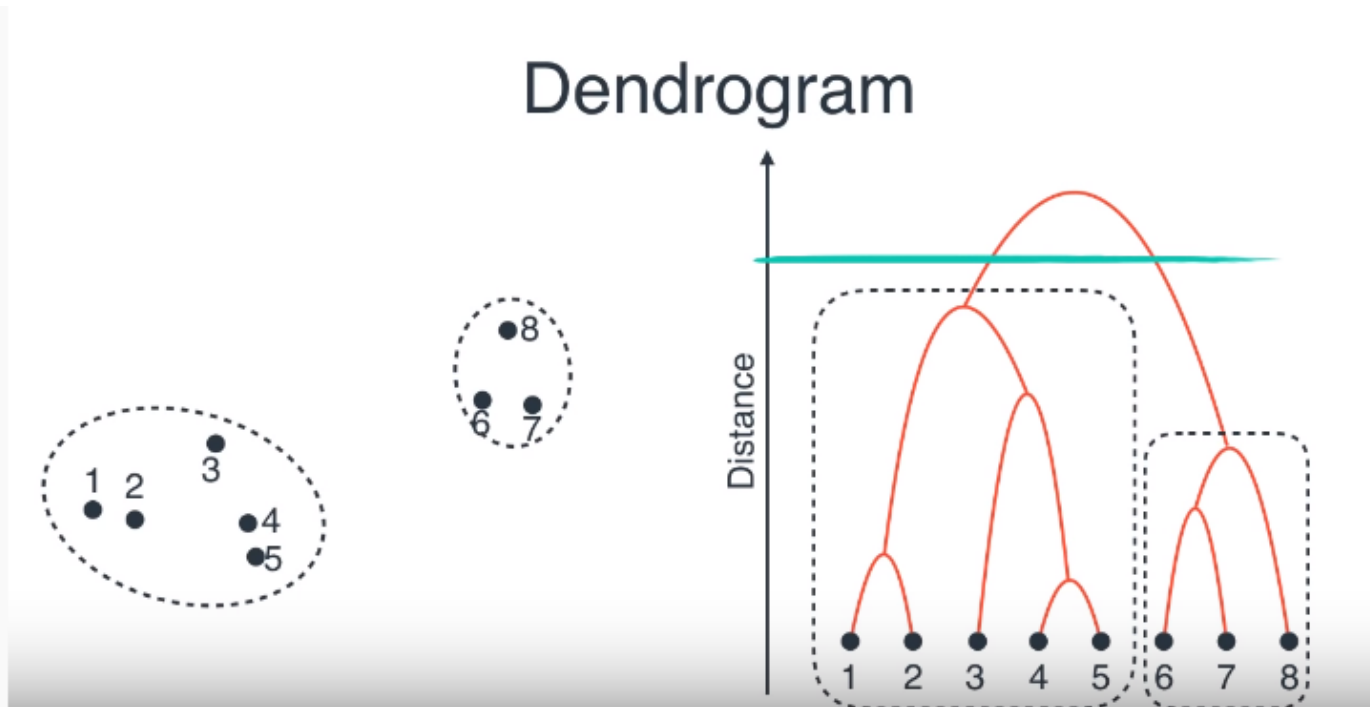
- **Tipo Aglomerativo:** Empezamos a agrupar desde cada elemento individual. Al inicio cada punto o dato está en un clúster separado. A cada paso, los dos clústers más cercanos se fusionan. Al final del proceso solo queda un único clúster que aglutina todos los elementos.
- **Divisible:** Comenzamos a la inversa, partimos de un único clúster que aglomera todos los datos y vamos dividiendo en clústers más pequeños.

**Dendograma:** manera de representar un clústering jerárquico.



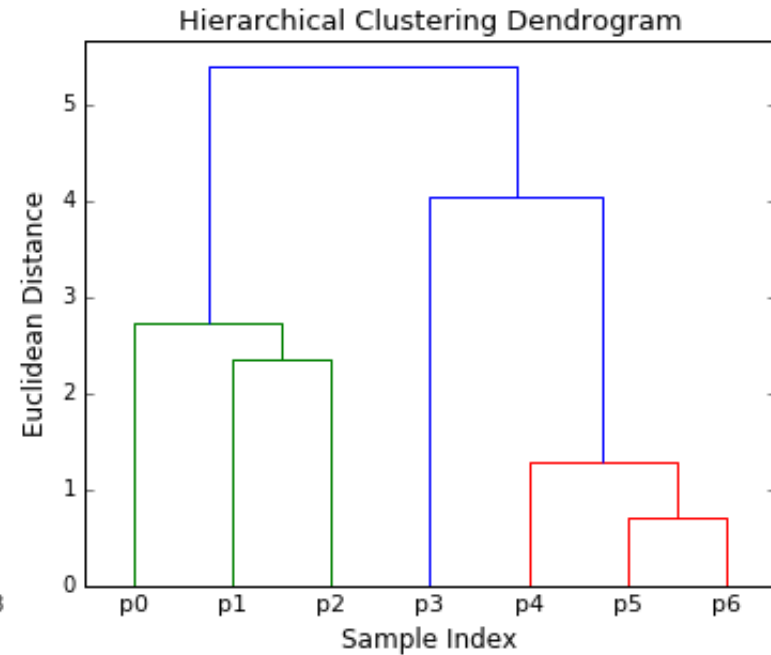
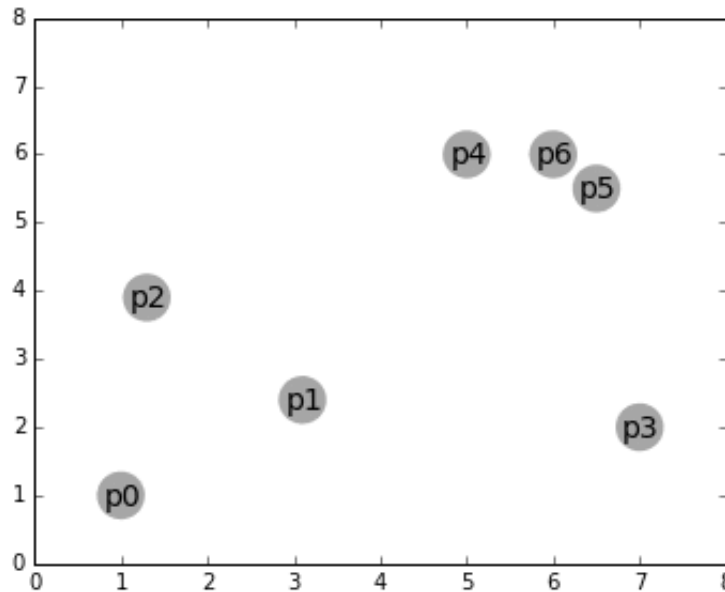


## Método jerárquico: clustering jerárquico aglomerativo



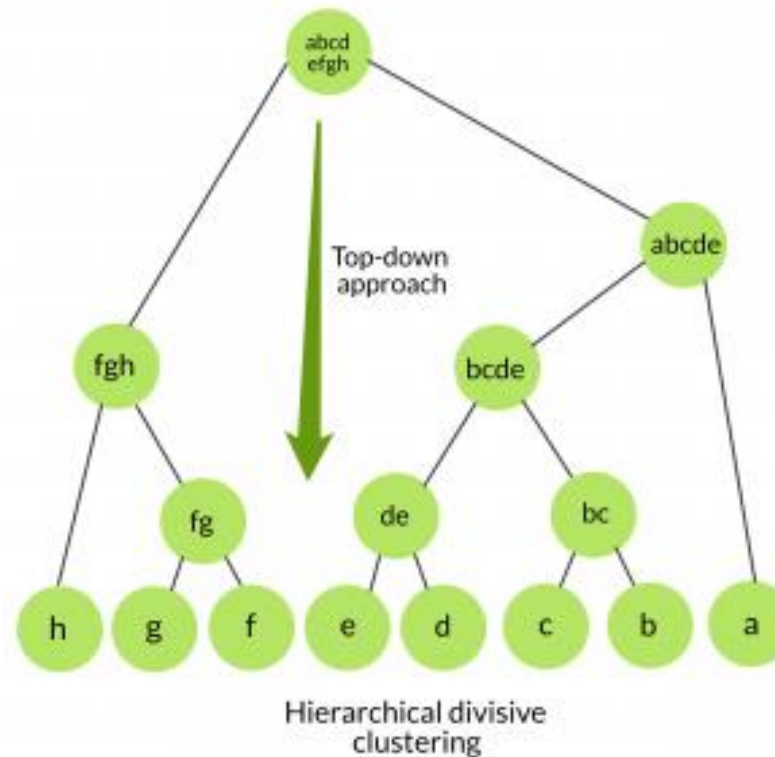


## Método jerárquico: clustering jerárquico aglomerativo





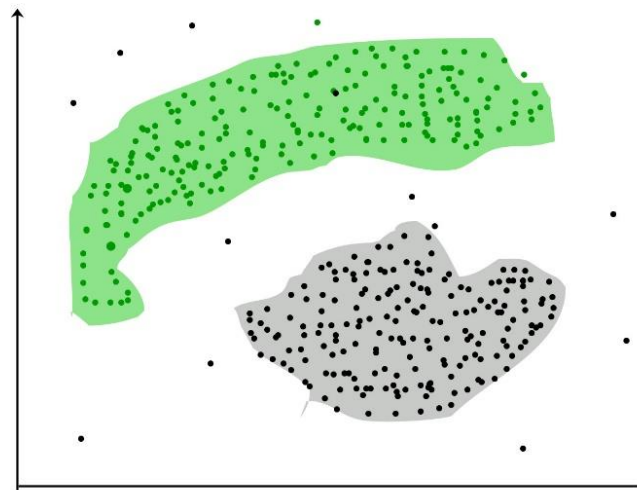
## Método jerárquico: clustering jerárquico divisible





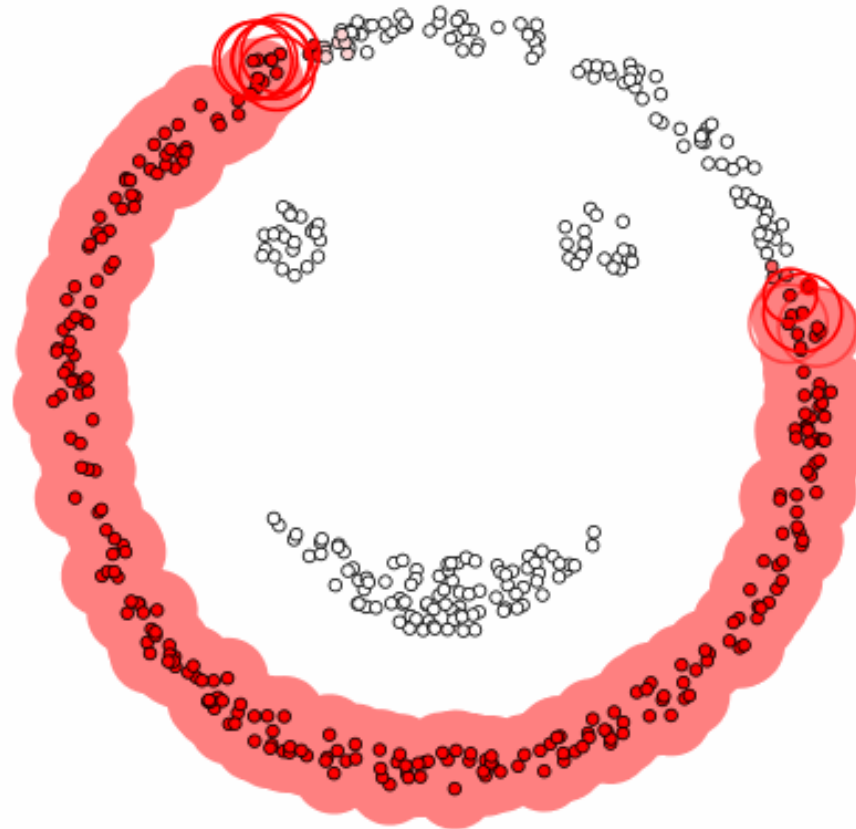
## Método basado en la densidad

- Un clúster en una región densa de puntos, separada por regiones poco densas de otras regiones densas.
- Útiles cuando los clusters tienen formas irregulares, están entrelazados o hay ruido/outliers en los datos. Se elimina el ruido del conjunto de datos.
- Algoritmo basado en densidad: **DBSCAN** (Density Based Spatial Clustering of Applications with Noise)





## Método basado en la densidad





## Método basado en la densidad: DBSCAN

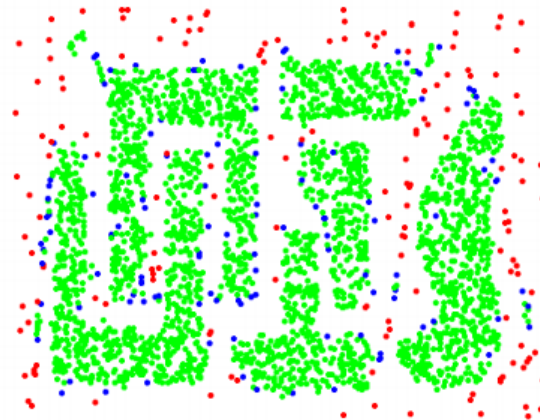
**DBSCAN:** Detecta regiones densas de puntos separadas de otras regiones densas por regiones poco densas

**Densidad:** Número de puntos en un radio específico

**Dataset**



**DBSCAN Clustering**



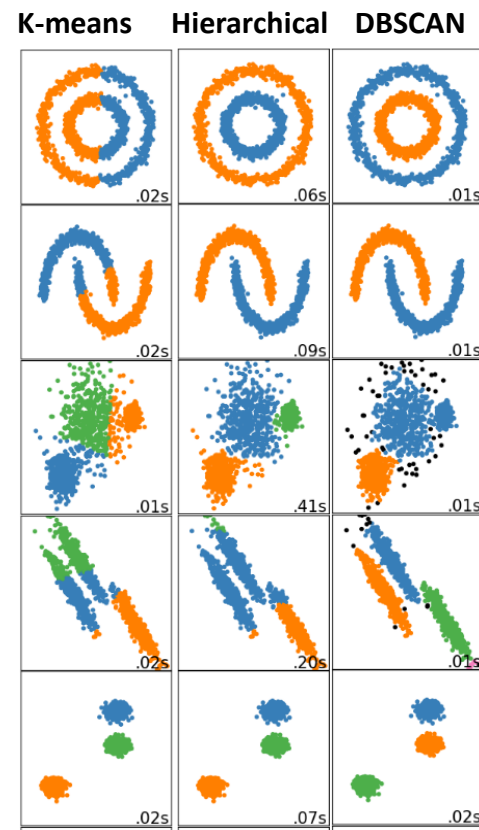
**core** (1 cluster)   **border** (frontera)   **noise** (ruido)





## Comparación de algoritmos de clustering

MÉTODO CLUSTERING	APLICACIÓN	MÉTRICA UTILIZADA
<b>K-means</b>	<ul style="list-style-type: none"> <li>• Uso general</li> <li>• Tamaño de cluster uniforme y de geometría plana</li> <li>• No demasiados clusters.</li> </ul>	Distancia entre puntos
<b>Hierarchical clustering</b>	<ul style="list-style-type: none"> <li>• Muchos clusters</li> <li>• Distancias no euclidianas</li> </ul>	Cualquier distancia por pares
<b>DBSCAN</b>	<ul style="list-style-type: none"> <li>• Geometría no plana</li> <li>• Tamaños desiguales de clusters</li> <li>• Detección de outliers</li> </ul>	Las distancias entre los puntos más cercanos



<https://scikit-learn.org/stable/modules/clustering.html#clustering>





Diferenciar el aprendizaje Supervisado y No Supervisado.

Algoritmos utilizados en Aprendizaje No Supervisado.

**Métricas para encontrar el número óptimo de clusters.**

**Aplicaciones** del clustering.

Introducción a **Reinforcement Learning** y sus aplicaciones.



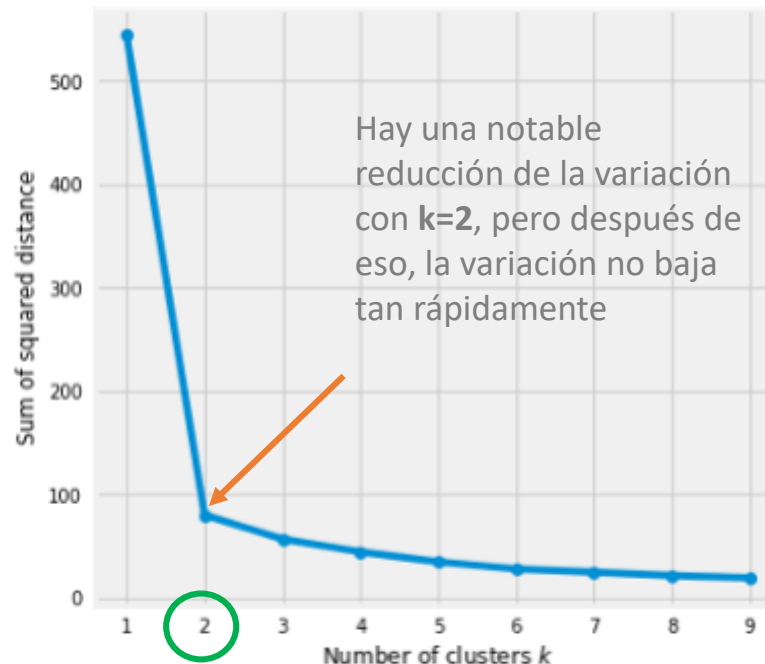
*En el aprendizaje no supervisado no existe una cuantificación del error como tal entre el resultado del modelo de clustering y el valor verdadero, ya **que el valor verdadero es desconocido**. Se utilizarán métodos para estimar el número óptimo  $k$  de clusters.*



## Métrica de evaluación

### Método Elbow

- El método Elbow nos da una idea del número óptimo de **K** de clusters.
- Se basa en la suma de la distancia al cuadrado entre los puntos de datos y los **centroides** de los clústers asignados.



$$Inercia = \sum_{i=0}^N \|x_i - \mu\|^2$$



## Métrica de evaluación

### Coeficiente Silhouette

- Esta métrica de evaluación puede utilizarse para determinar el grado de separación entre los clusters/grupos.
  - Calcular la distancia promedio de todos los puntos de datos en el mismo cúmulo **a(i)**.
  - Calcular la distancia promedio de todos los puntos de datos en el cúmulo más cercano **b(i)**.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

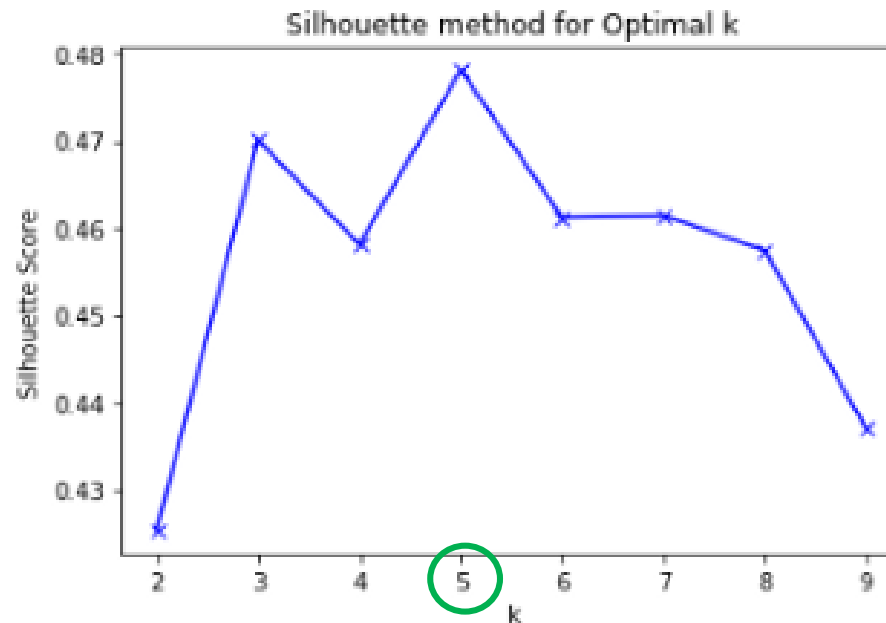
- El coeficiente puede tomar valores en el intervalo [-1, 1].
  - **0**: la muestra está muy cerca de los clusters vecinos.
  - **1**: la muestra está muy lejos de los clusters vecinos.
  - **-1**: la muestra está asignada a los clusters equivocados.



## Métrica de evaluación

### Coeficiente Silhouette

- Queremos que los coeficientes sean lo más grandes posibles y que estén cerca de 1 para tener un buen cluster.



Número óptimo de clusters: **k=5**



Diferenciar el aprendizaje Supervisado y No Supervisado.

Algoritmos utilizados en Aprendizaje No Supervisado.

Métricas para encontrar el número óptimo de clusters.

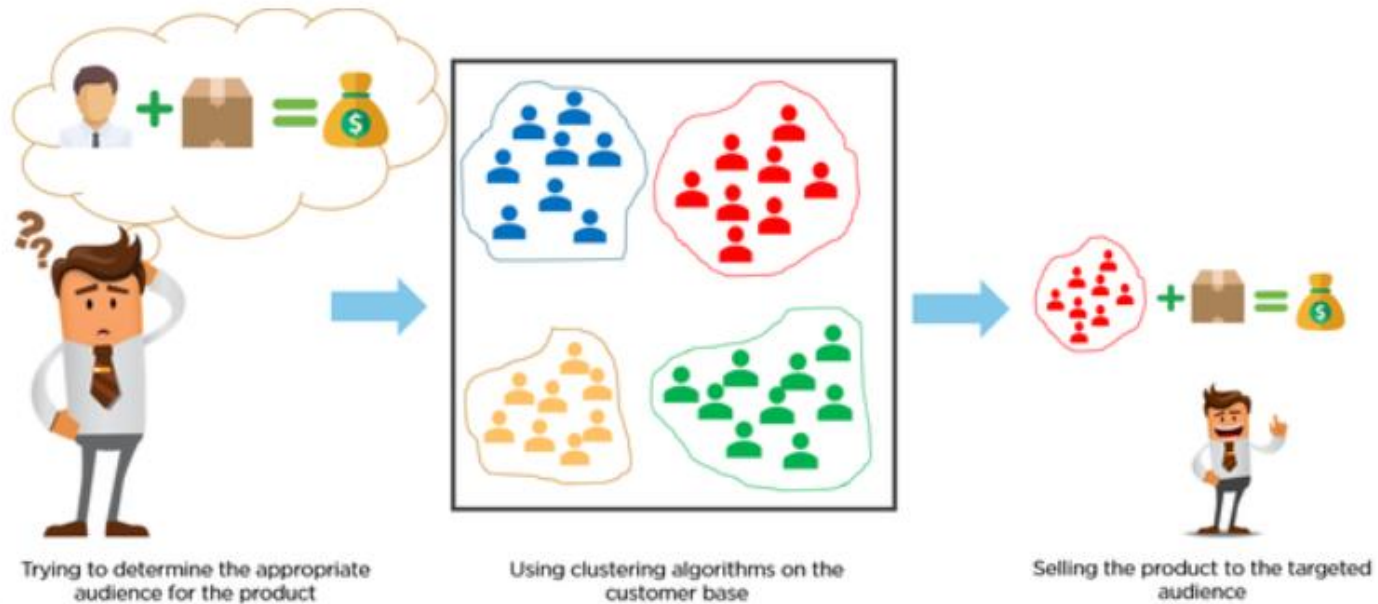
## **Aplicaciones del clustering.**

Introducción a **Reinforcement Learning** y sus aplicaciones.



## Aplicaciones del Clustering

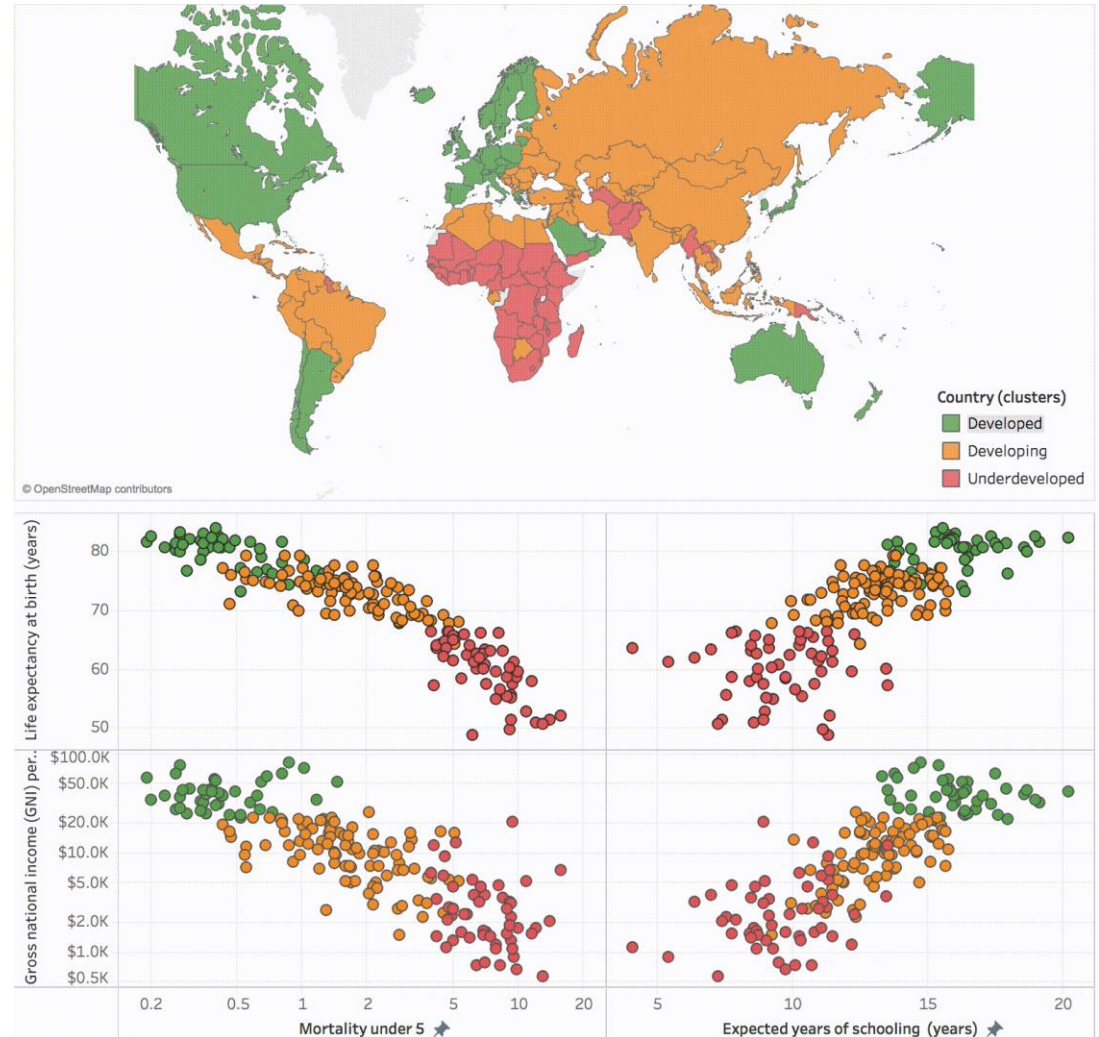
**Márketing:** Segmentación de clientes





## Aplicaciones del Clustering

### Elaboración de gráficos

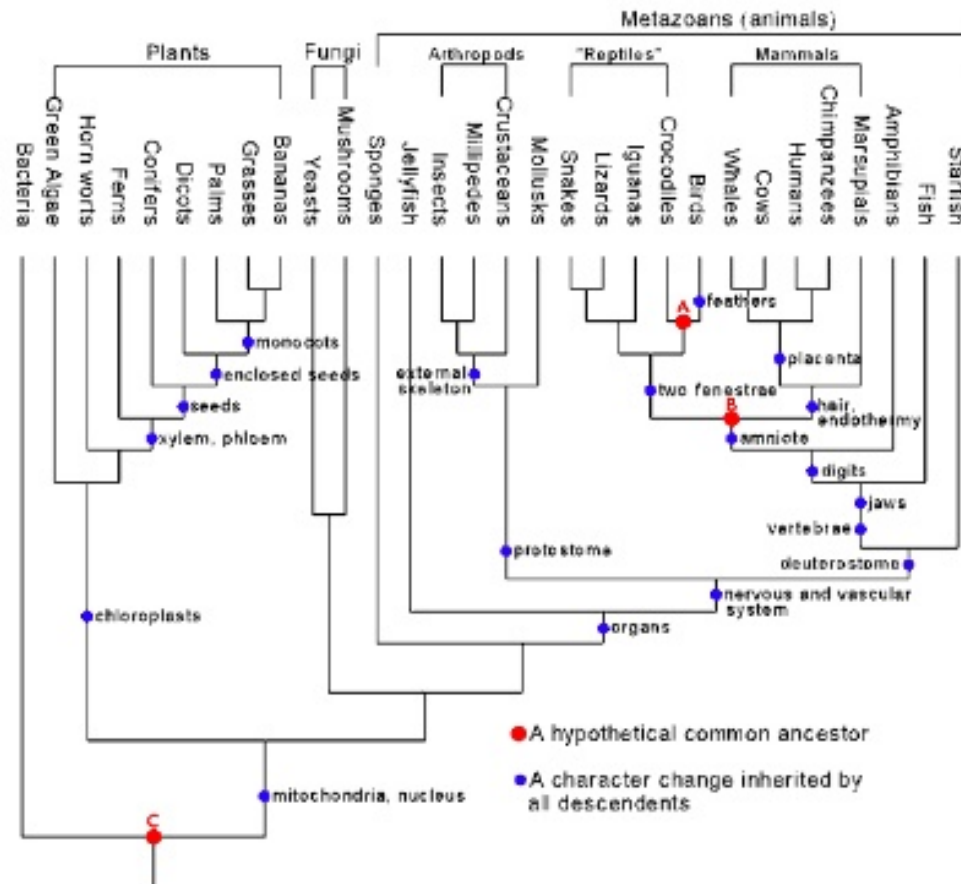






## Aplicaciones del Clustering

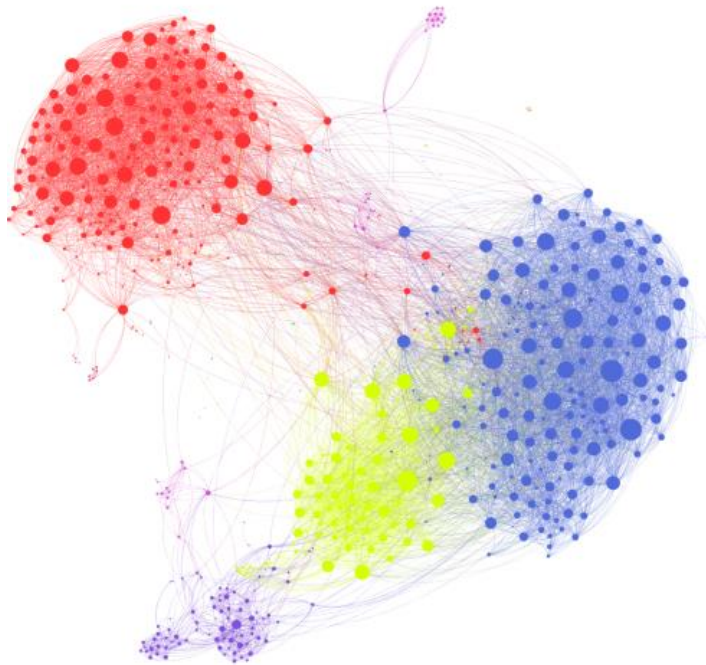
**Biología:** clasificación de plantas y animales según sus características.



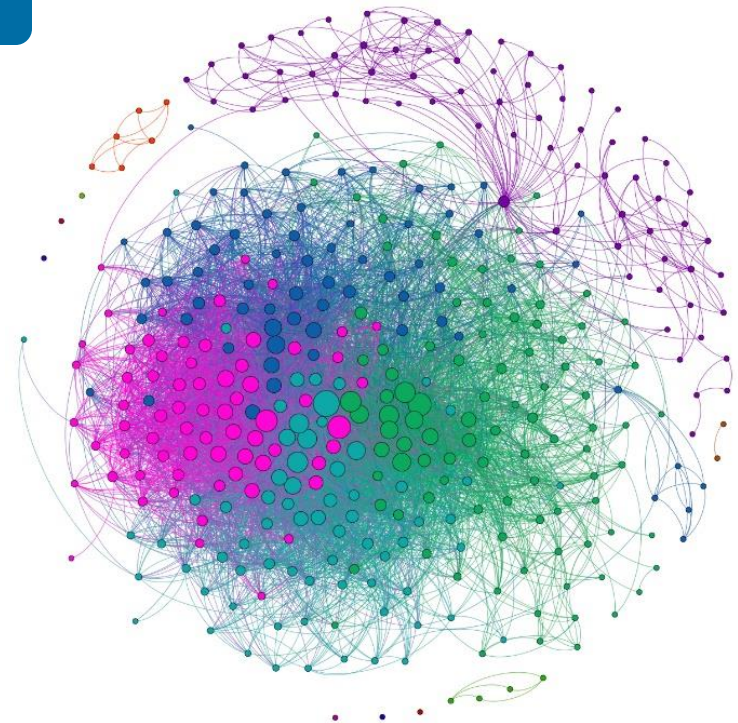


## Aplicaciones del Clustering

### Comunidades de redes sociales



<https://griffgraphs.wordpress.com/tag/social-network/>

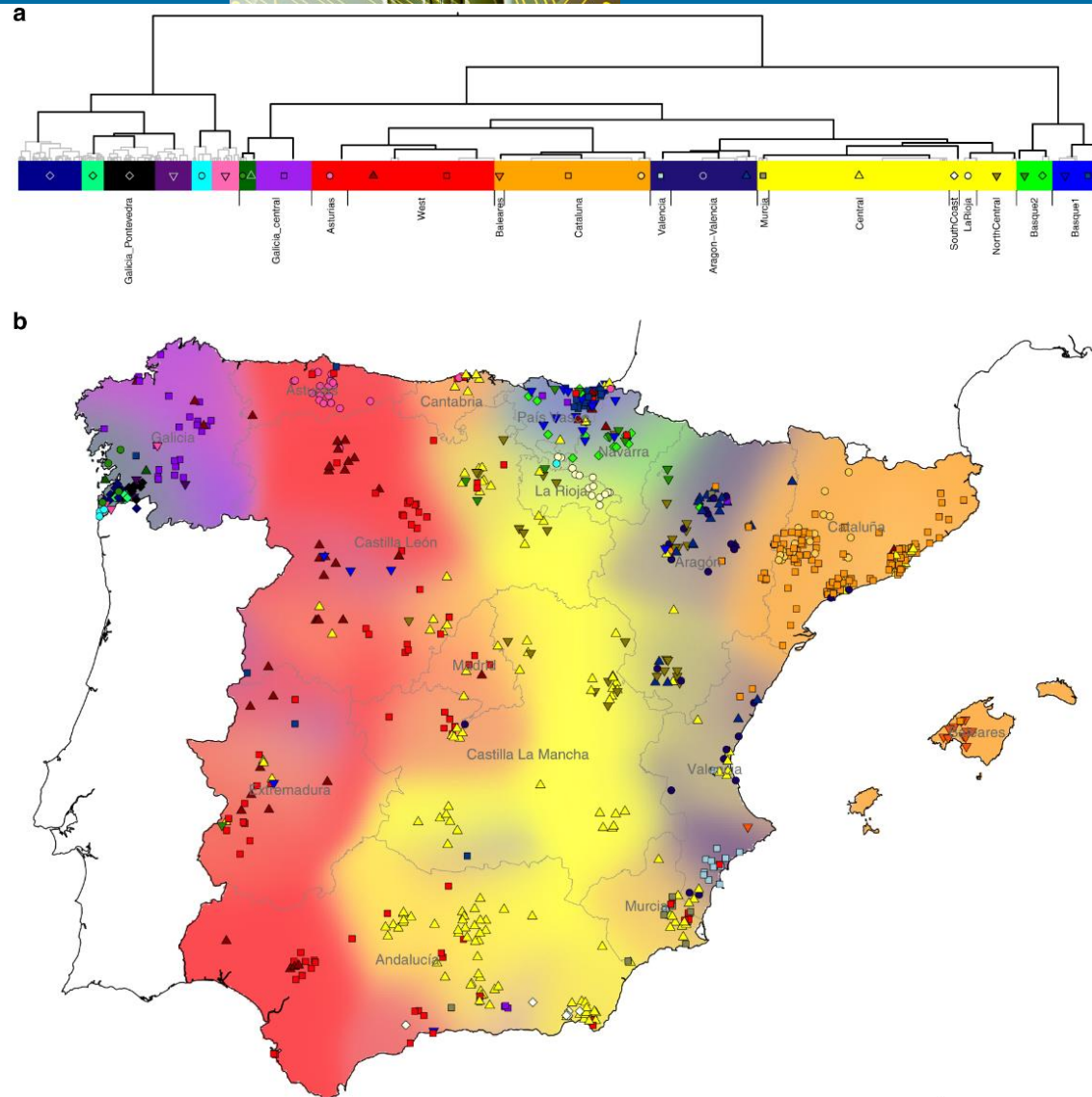


<http://social-dynamics.org/gephi-facebook/>



## Aplicaciones

### Genética





## Introducción al aprendizaje no supervisado

¿Qué abarca el aprendizaje no supervisado?

### Clustering

- Algoritmos de partición (K-means).
- Algoritmos basados en la jerarquía (Hierarchical).
- Algoritmos basados en la densidad (density).

**Reducción de dimensiones (Análisis de componentes principales (PCA))**



## Introducción al aprendizaje no supervisado

¿Qué abarca el aprendizaje no supervisado?

### Clustering

- Algoritmos de partición (K-means).
- Algoritmos basados en la jerarquía (Hierarchical).
- Algoritmos basados en la densidad (density).

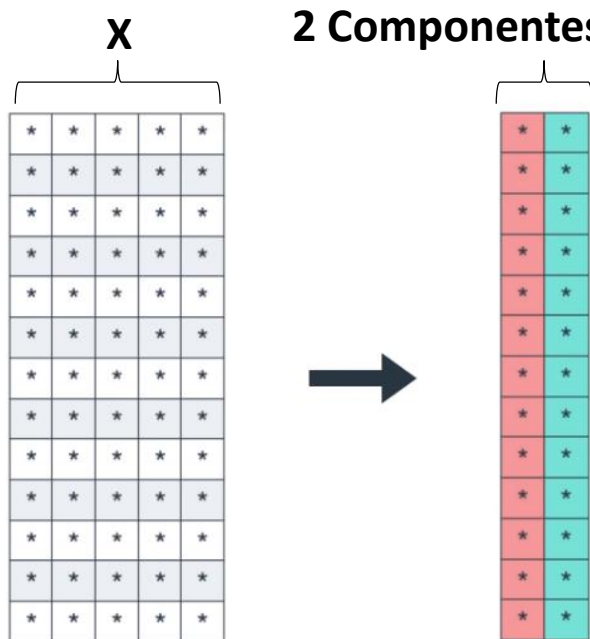
**Reducción de dimensiones (Análisis de componentes principales (PCA))**





## Reducción de dimensiones: principal component analysis (PCA)

- El **análisis de componentes principales** es uno de los algoritmos de **selección de características** (feature selection) más habituales.
- Se usa para reducir la “dimensión” de los datos. Con dimensión nos referimos al “número de variables”.



- Pasamos de 5 características a 2.
- Las 2 columnas capturan la mayoría de la información de las 5 columnas anteriores.
- El resultado indica que no se pierde mucha información y el dataset (matriz X) se simplifica bastante.

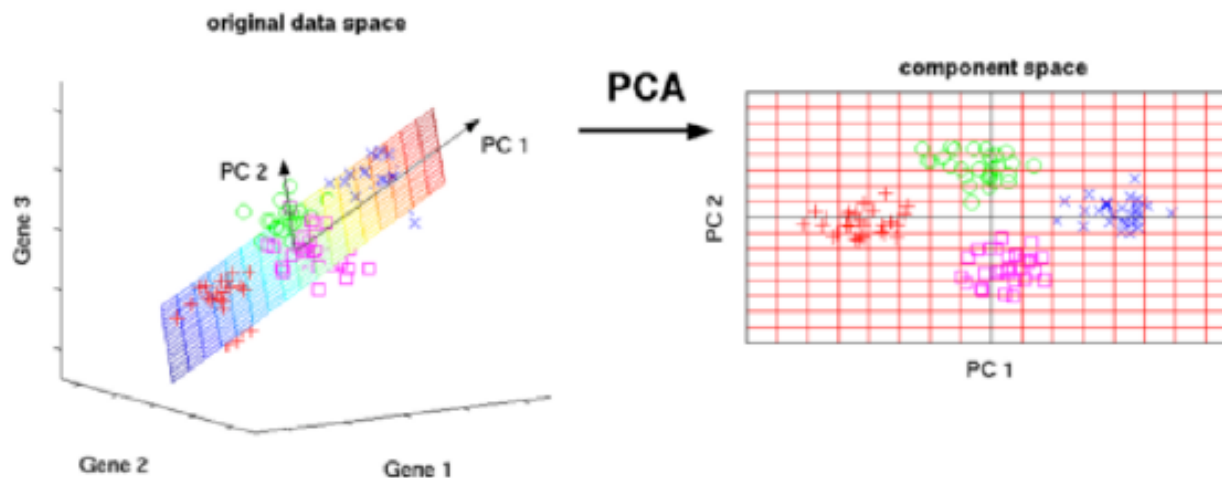


## Reducción de dimensiones: principal component analysis (PCA)

**Ejemplo.** Pasamos de 3 dimensiones a 2 dimensiones utilizando PCA.

```
from sklearn.decomposition import PCA
```

```
>>> import numpy as np
>>> from sklearn.decomposition import PCA
>>> X = np.array([[ -1, -1], [-2, -1], [-3, -2], [ 1,  1], [ 2,  1], [ 3,  2]])
>>> pca = PCA(n_components=2)
>>> pca.fit(X)
```





# Ejercicio de Clustering







Diferenciar el aprendizaje Supervisado y No Supervisado.

Algoritmos utilizados en Aprendizaje No Supervisado.

Métricas para encontrar el número óptimo de clusters.

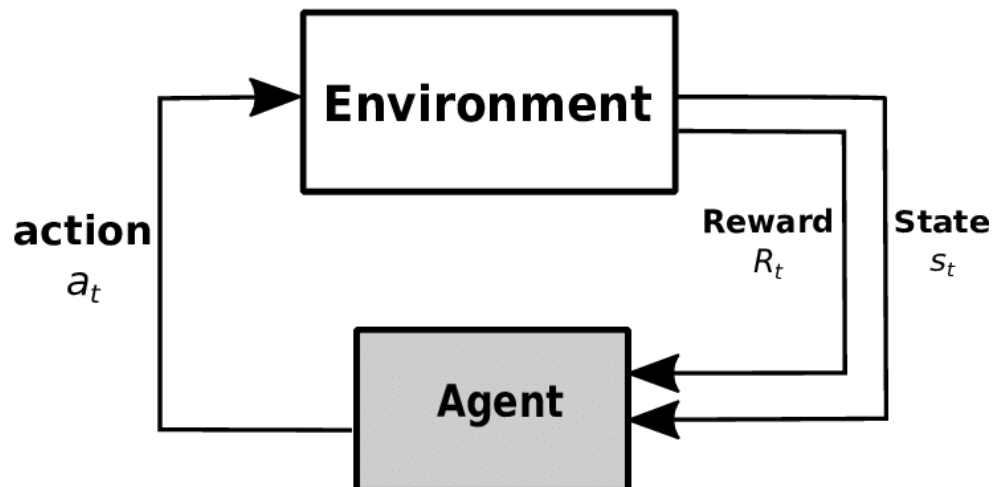
Aplicaciones del clustering.

**Introducción a Reinforcement Learning y sus aplicaciones.**



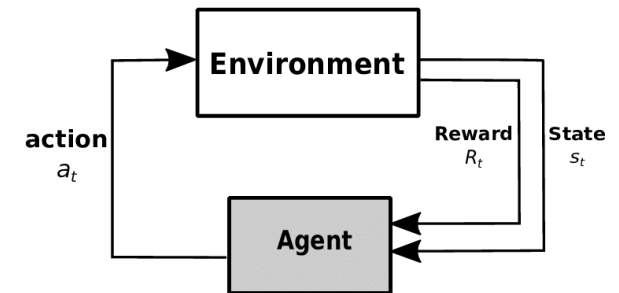
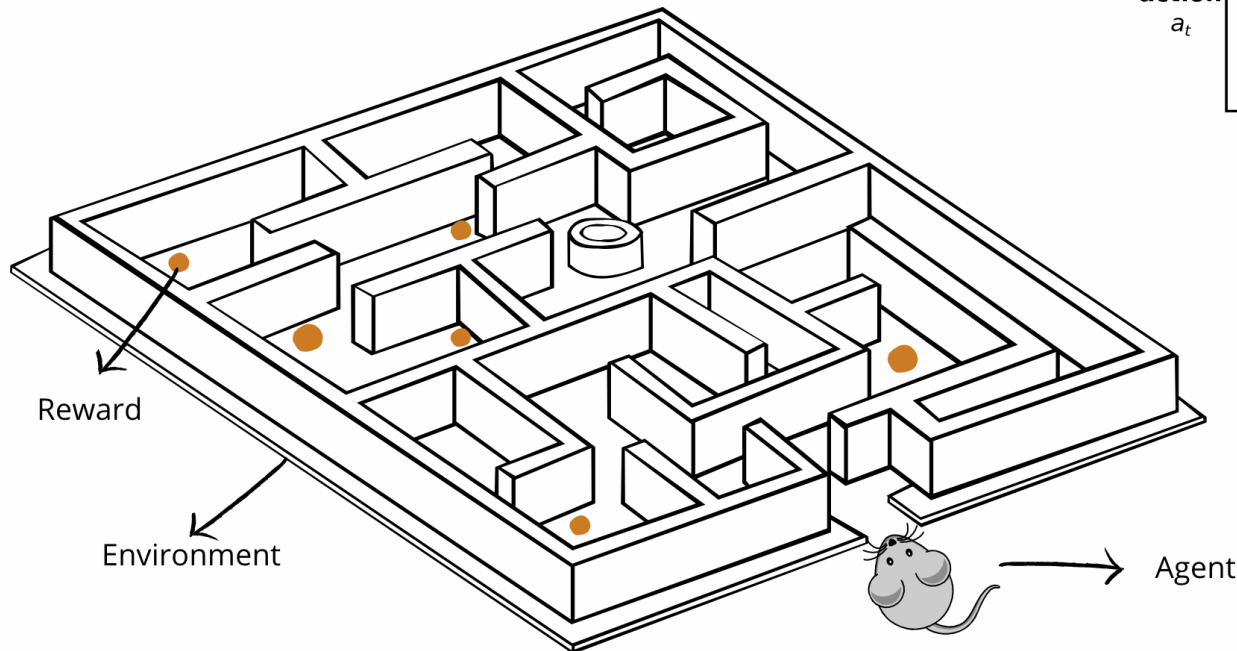
## Reinforcement Learning

- Reinforcement Learning buscar determinar qué acciones debe escoger un agente de software en un entorno dado con el fin de **maximizar** alguna noción de "**recompensa**" o premio acumulado.
- Con el tiempo, aprende a hacer buenas secuencias de decisiones bajo incertidumbre.





## Reinforcement Learning

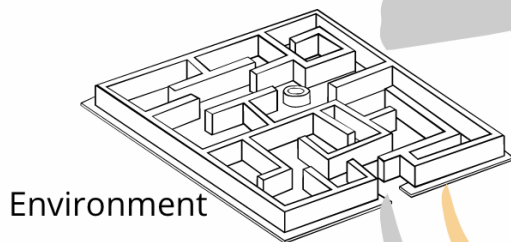
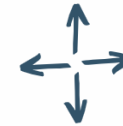




## Reinforcement Learning

### Mouse's Markov Decision Process

Action taken in state 1 = Moving in 4 directions



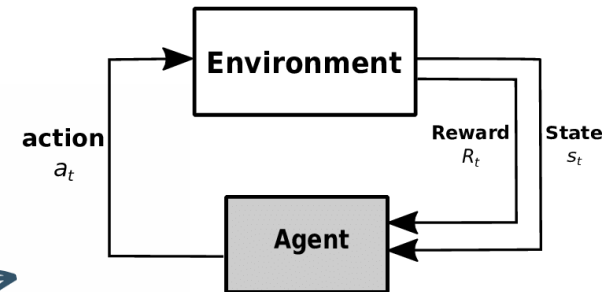
Environment

State 1

State 2



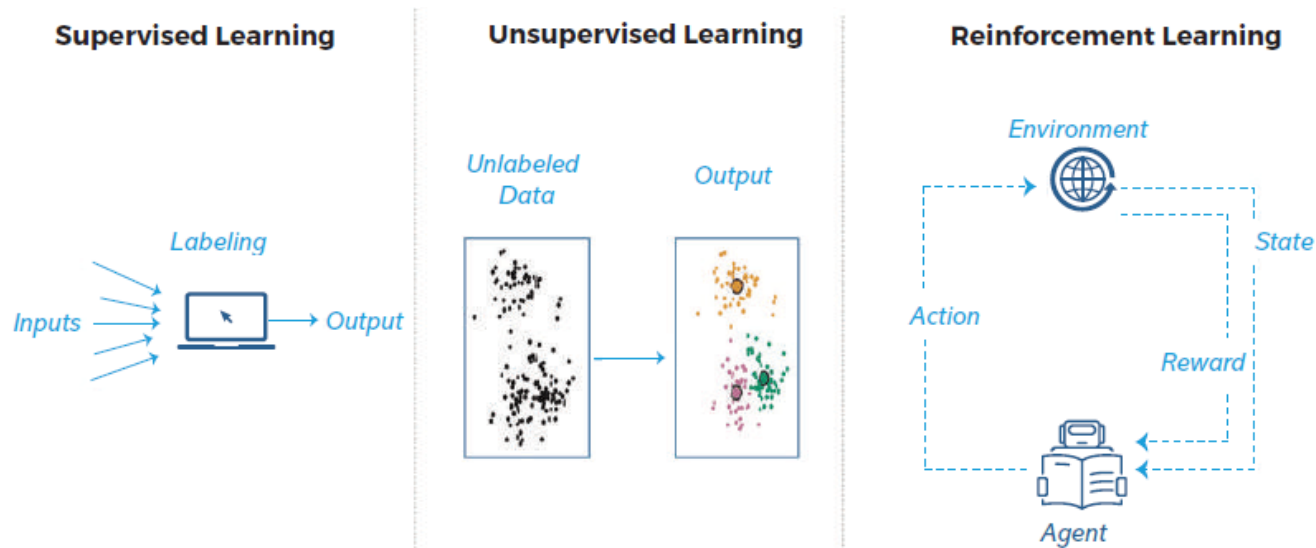
Reward for action taken in state 1 = Peanut butter





## Reinforcement Learning

- En contraste con modelos supervisados, los algoritmos de Reinforcement Learning (RL) están diseñados para **aprender continuamente de la experiencia** y no sólo de los datos.
- Esta capacidad de auto-aprendizaje de los algoritmos RL es crítica para lograr un comportamiento más humano en los robots y otras aplicaciones de la IA.

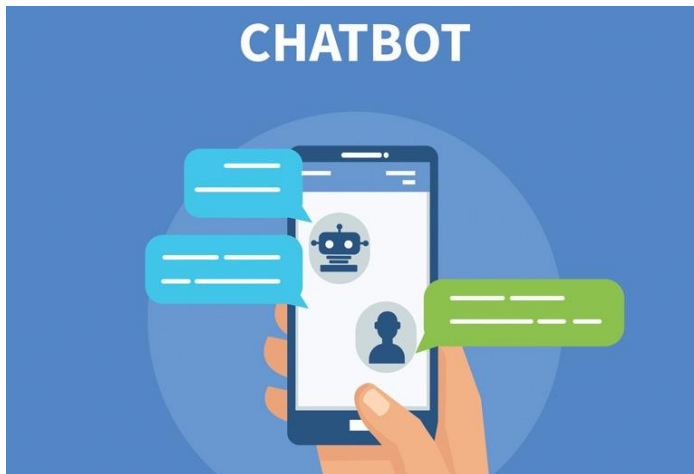


*Diferencia entre aprendizaje supervisado, no supervisado y reinforcement learning.*



## Aplicaciones de Reinforcement Learning

### Chat boots



Pero... **cuidado!**

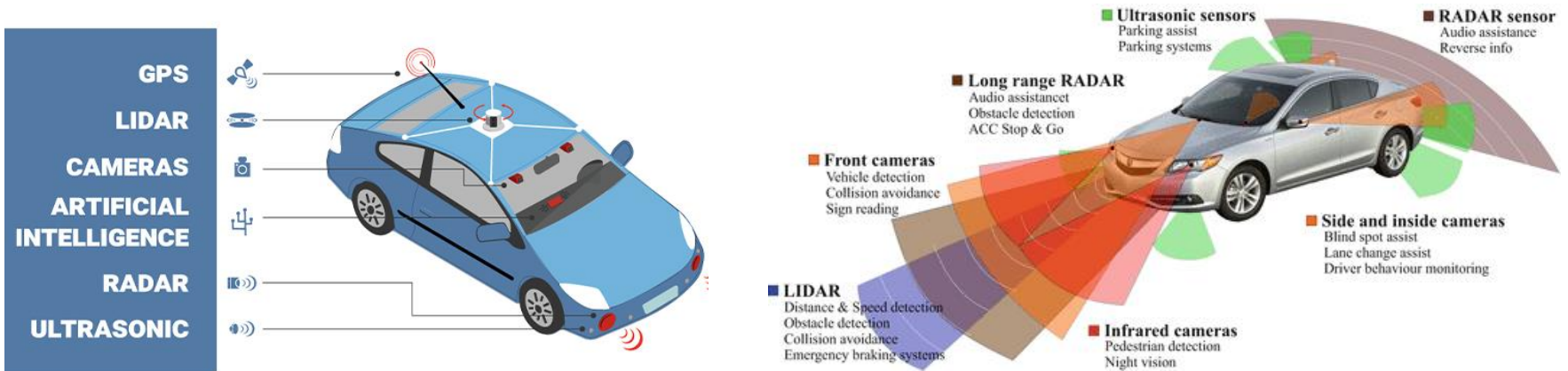




## Aplicaciones de Reinforcement Learning

### Robótica

### Conducción autónoma vehículos



[Learning to drive in a day!](#)

La explicación [aquí](#)





## Aplicaciones de Reinforcement Learning

### Servicios financieros

#### Portfolio Optimization



Aims to optimize asset allocation for portfolios on a continual basis to help clients achieve their financial goals. AI Portfolio Manager complements asset managers by building, monitoring & automatically rebalancing a portfolio. Virtusa formulated a Deep reinforcement learning solution with a structured approach, the solution recommended asset allocation considering financial and demographic aspects of the client. AI Portfolio Manager helps enhance both efficiency and effectiveness of human managers.

#### Option pricing



Reinforcement learning can be used to design a model for derivatives pricing. The classical Black-Scholes-Merton (BSM) model assumes no transaction costs and continuous hedging, both these assumptions make it eliminate any risk. In RL formulation, a risk-adjusted Markov decision process is constructed as a discrete-time version of the BSM model. The model provides a simultaneous and consistent hedging and pricing of a derivative, unlike many other market models.<sup>5</sup>

#### Optimal Trading



Aims at creating a robotic trader that generates profit from trading in financial markets. Innovation and research in Deep reinforcement learning provides a framework for build such an agent. Researchers have proposed an Markov Decision Process (MDP) model suitable for the financial trading using deep recurrent Q-network (DRQN) algorithm<sup>6</sup>. RL agents have been able to learn optimal trading strategies (for synthetic assets) which outperformed the simple Buy & Hold strategy.

#### Recommendation Systems



Recommender systems exists to solve the problem of choosing one among plenty. Traditional Machine learning approaches use Collaborative filtering or content based methods to recommend items. But reinforcement learning approach considers a recommender system as a Markov decision process (MDP) problem. Researchers have formulated recommender systems as a grid world game by using a biclustering technique<sup>3</sup>. Benefits of using RL over traditional system include overcoming cold-start problem, online-learning and explainable recommendations. The performance of this method has been examined to be better than standard recommender techniques<sup>7</sup>.

#### Conversational Chatbots



Conversational chatbots are also built with an intention to make agent capable of conversing with humans on popular topics, such as entertainment, weather, and sports etc. Researchers have leveraged deep learning and reinforcement learning to combine retrieval and generation models<sup>8</sup>. RL based chatbots have shown substantial improvements in A/B testing experiments with real-world users.

#### P2P lending



In P2P setting, Managers have to manage portfolios of loans. Loan portfolios can be optimized using Reinforcement learning. RL helps the lenders decide the fund allocation by optimizing risk-return profile in the context of sequential investment decision-making. The solution can be formulated as convex optimization similar to optimization formulation for Stock portfolio optimization.

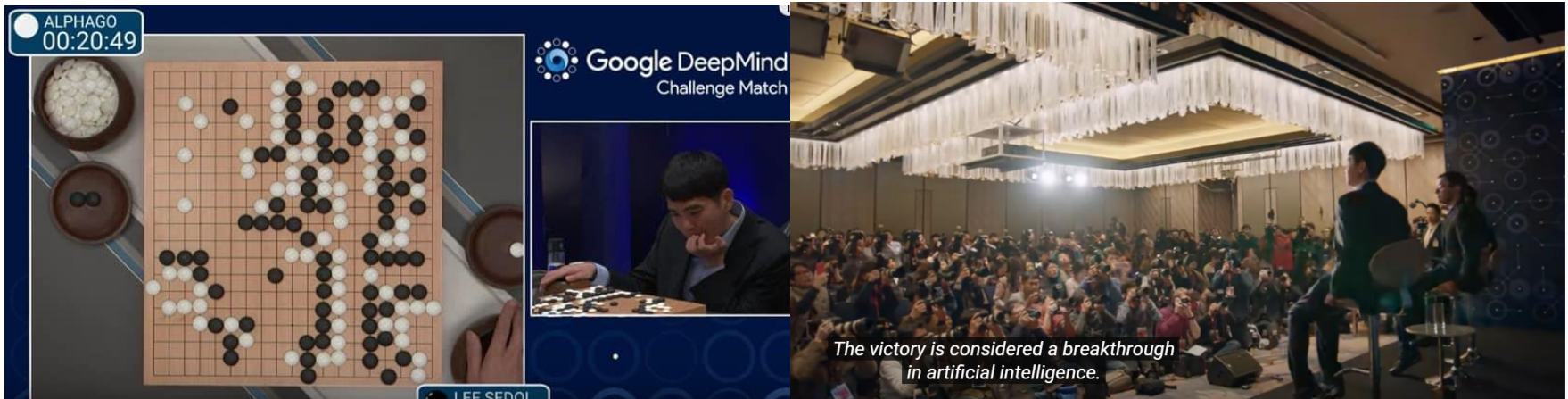
<https://medium.com/analytics-vidhya/reinforcement-learning-next-step-in-ai-for-banking-and-financial-services-part-2-de41b453cdfc>





## Aplicaciones de Reinforcement Learning

### Juegos



**Multi-Agent Hide and Seek**



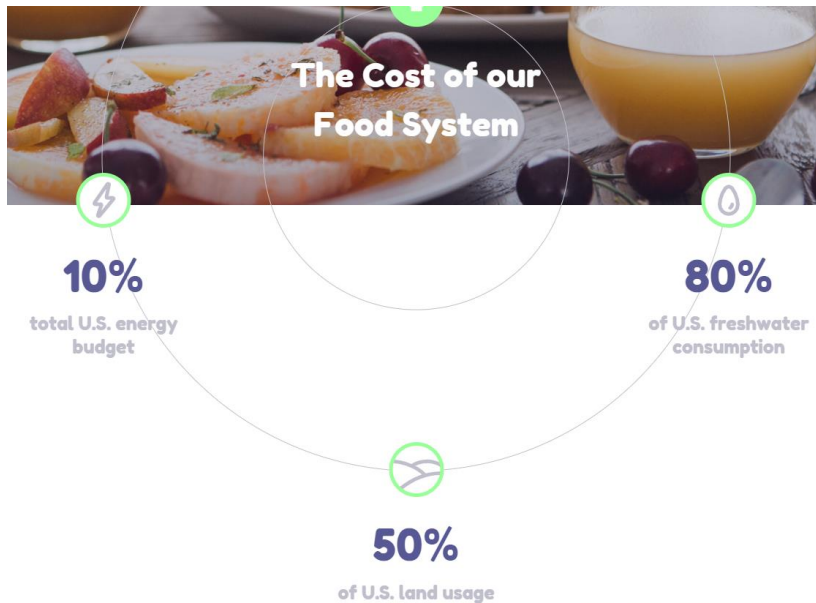
**Documental acerca de DeepMind -- AlphaGO**



## Aplicaciones de Reinforcement Learning

### Alimentación

<https://www.wasteless.com/>



### Expected results



Reduction in food waste

**-50%**



Increase in revenues

**+20%**



Increase in net margins

**+3%**



<https://www.youtube.com/watch?v=BPXr6RWkubA>