



Inteligencia Artificial & Machine Learning

Aplicaciones en movilidad

Engineering The logo for Engineering, consisting of the word "Engineering" in a bold, dark blue sans-serif font next to a stylized "X" symbol made of vertical lines.

Founded by the Royal Academy of Engineering
and Lloyd's Register Foundation

Dr. Iván S. Razo Zapata



Aprendizaje No Supervisado

Unsupervised Learning



Aprendizaje No Supervisado

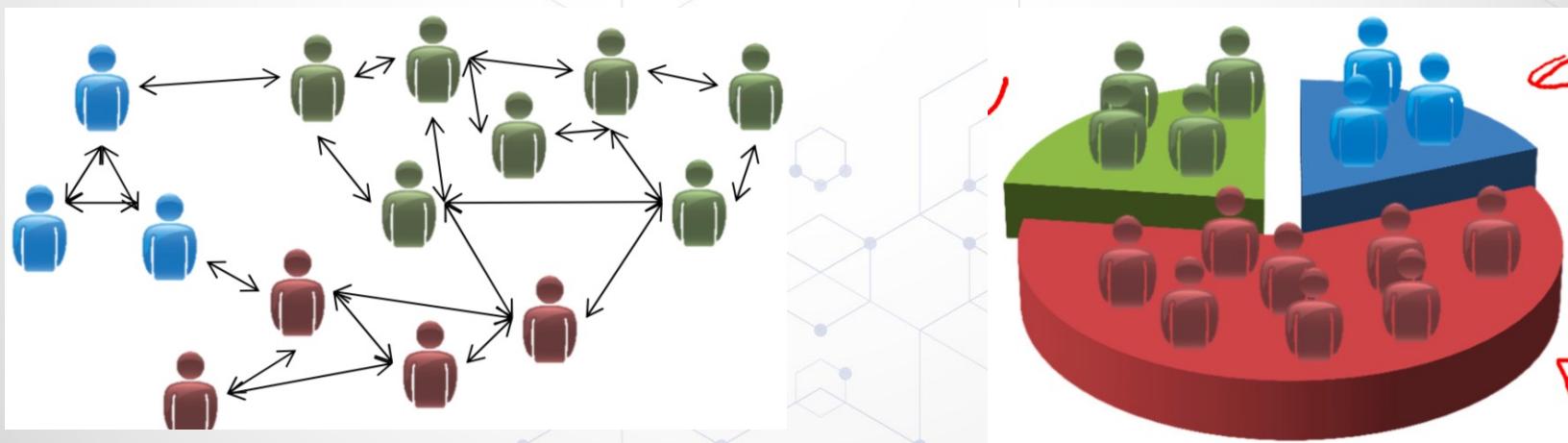
- Aprendizaje sin supervisor / maestro
- N observaciones (x_1, x_2, \dots, x_N)
- Las cuales “se asume” siguen un distribución $Pr(X)$
- La meta consiste en **inferir/aprender** las propiedades de una **densidad de probabilidad** sin la ayuda de un supervisor o maestro que proporcione las respuestas correctas

Aprendizaje No Supervisado - Algunos retos

- La dimension de X es usualmente mayor que en el caso supervisado
- Las propiedades también suelen ser más complicadas
- Un reto más
 - No hay medidas de éxito como en el caso del aprendizaje supervisado
 - Usualmente se pueden emplear heurísticas (dependientes del problema)

Aprendizaje No Supervisado

- Algunas aplicaciones:
- Análisis de redes sociales (social network analysis)
- Segmentación de clientes con base en patrones de compra



Aprendizaje No Supervisado

- Algunas aplicaciones:

Top stories

- For you
- Following
- Saved searches

- COVID-19
- U.S.
- World
- Your local news
- Business
- Technology
- Entertainment
- Sports
- Science
- Health

Science

This American astronaut voted from space. Here's how she did it

CNN · 7 hours ago

- NASA Astronaut Inspired by Planet Earth | Sneak Peek The ISS Experience

TIME · 13 hours ago

[View Full Coverage](#)

Scientists Peer Inside an Asteroid – Is Bennu in the Process of Spinning Itself Into Pieces?

SciTechDaily · 10 hours ago

- NASA's OSIRIS-REx collects science treasure from asteroid Bennu | TheHill

The Hill · 11 hours ago

[View Full Coverage](#)

NASA to Announce "Exciting New Discovery" About the Moon Made by Airborne Observatory

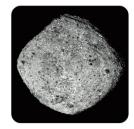
SciTechDaily · 11 hours ago

- NASA will share an 'exciting new discovery' about the moon on Monday. Listen to the announcement live.

msnNOW · 15 hours ago

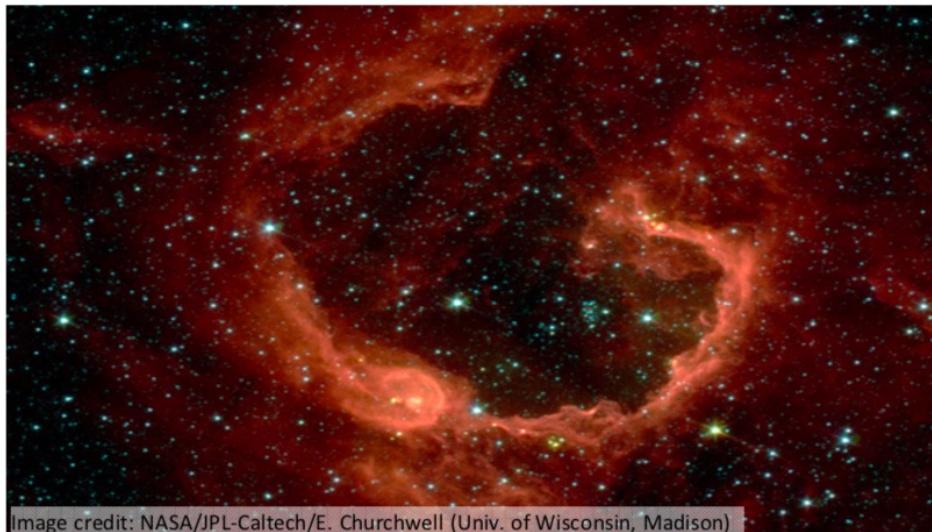
[View Full Coverage](#)

More Science



Aprendizaje No Supervisado

- Algunas aplicaciones:
- Grupos de galaxias



Principales técnicas

- Agrupamiento
- K-medias
- Expectation Maximization
- Por densidad
- Reglas de asociación
- Algoritmo Apriori

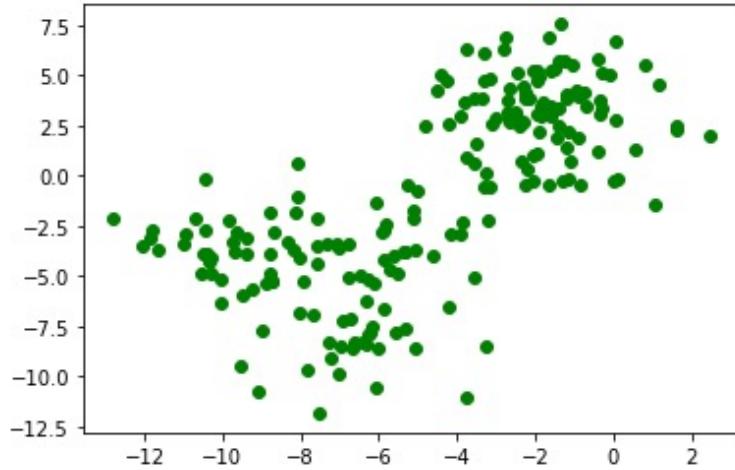


Agrupamiento



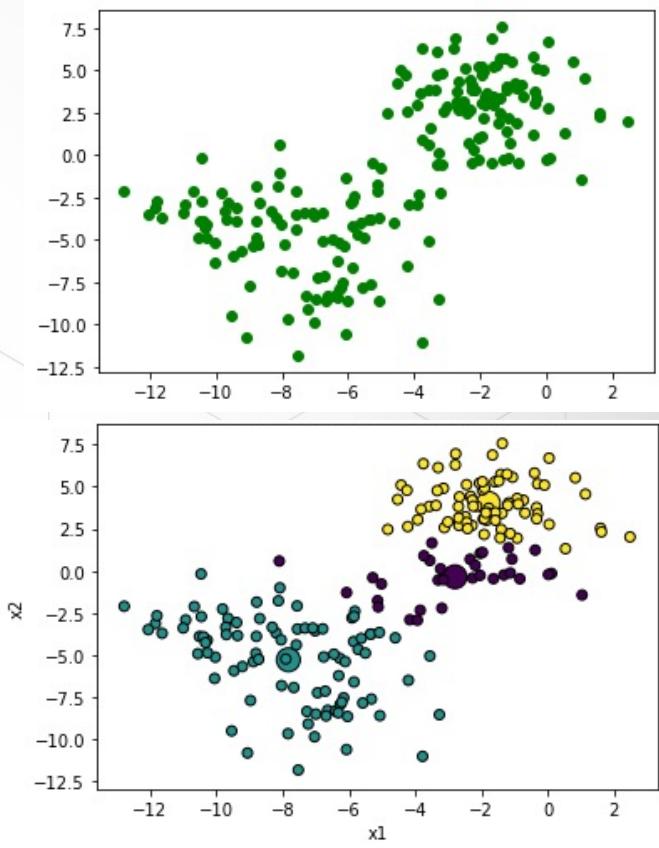
Clustering / Agrupamiento

- Es el proceso de agrupar datos en “clases” o clusters
 - Los objetos de un cluster tengan **similitud**
 - alta entre ellos, y
 - baja (sean muy diferentes) con objetos de otros clusters
- La medida de similitud está basada en los atributos que describen a los objetos
- Los grupos pueden ser exclusivos, con traslapes, jerárquicos
- Puede ser aplicado, por ejemplo, para caracterizar clientes, formar taxonomías, clasificar documentos, etc.



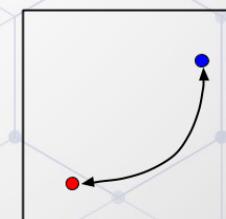
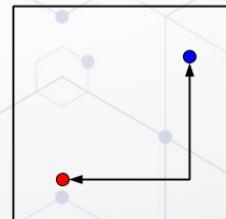
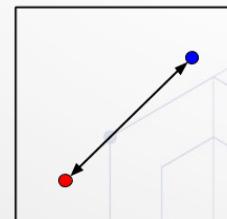
Clustering / Agrupamiento

- Es el proceso de agrupar datos en “clases” o clusters
 - Los objetos de un cluster tengan **similitud**
 - alta entre ellos, y
 - baja (sean muy diferentes) con objetos de otros clusters
- La medida de similitud está basada en los atributos que describen a los objetos
- Los grupos pueden ser exclusivos, con traslapes, jerárquicos
- Puede ser aplicado, por ejemplo, para caracterizar clientes, formar taxonomías, clasificar documentos, etc.



K - means

- K - medias
- Es de los más conocidos y usados
- Toma como parámetro k que es el número de clusters que forma
- Los objetos de un cluster tengan similitud alta entre ellos



K - means

- K - medias
- Es de los más conocidos y usados
- Toma como parámetro **k** que es el número de clusters que forma
- Los objetos de un cluster tengan similitud
 - alta entre ellos, y
 - baja (sean muy diferentes) con objetos de otros clusters

Un “buen” cluster tiene poca variación interna

$W(C_k)$ variación dentro del cluster

$$\min_{C_1, \dots, C_K} \left\{ \sum_{k=1}^K W(C_k) \right\}$$

K - means

$$\min_{C_1, \dots, C_K} \left\{ \sum_{k=1}^K W(C_k) \right\}$$

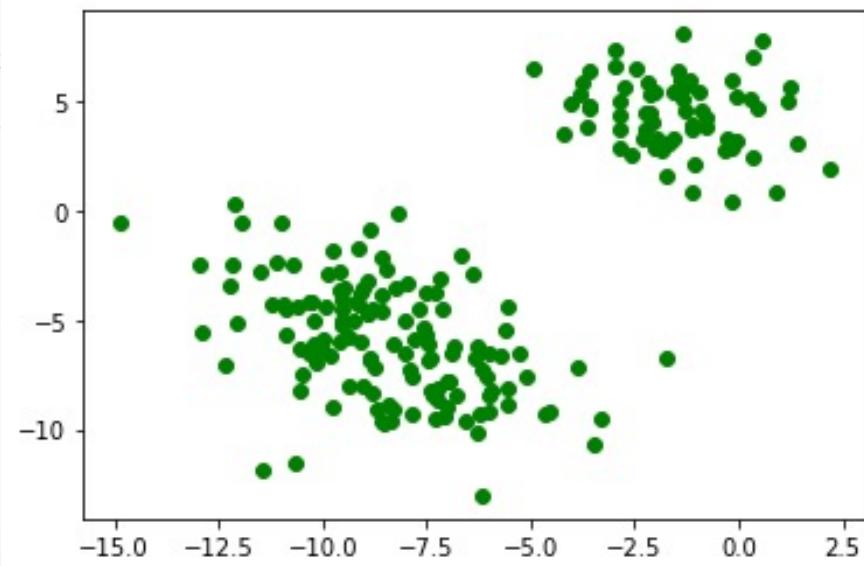
$$W(C_k) = \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2$$

$$\min_{C_1, \dots, C_K} \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i, i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\}$$

within-cluster sum of squares (WSS)

Ejemplo

- ¿Cuántos grupos / clusters?



K - means

- K - medias
- Pasos:
 1. Selecciona k elementos aleatoriamente, los cuales representan “temporalmente” el centro (centroide) o media de cada cluster
 2. A cada objeto restante se le asigna el cluster con el cual más se parece, basándose en una distancia entre el objeto y la media del cluster
 3. Después calcula la nueva media del cluster e itera hasta no cambiar de medias

K - means

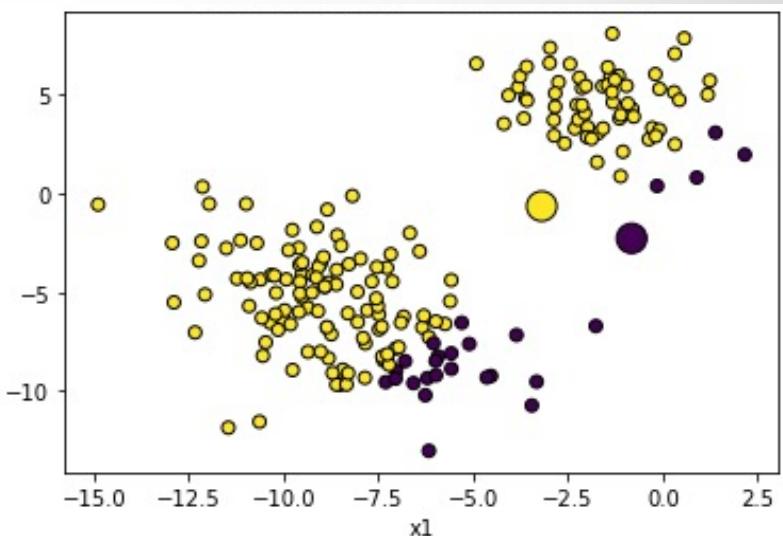
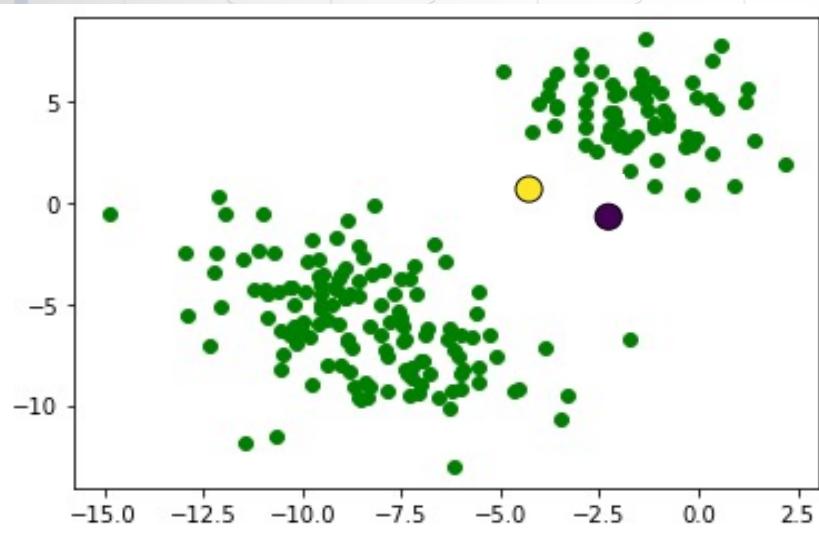
Selecciona k objetos aleatoriamente
repetir

- Re(asigna) cada objeto al cluster más similar
- Actualiza el valor de las medias de los clusters

hasta que no haya cambio en las medias

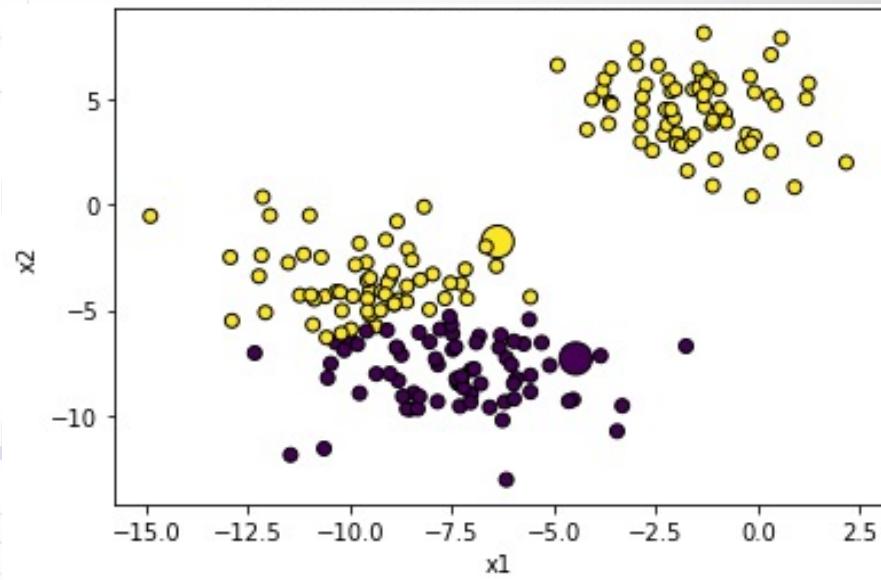
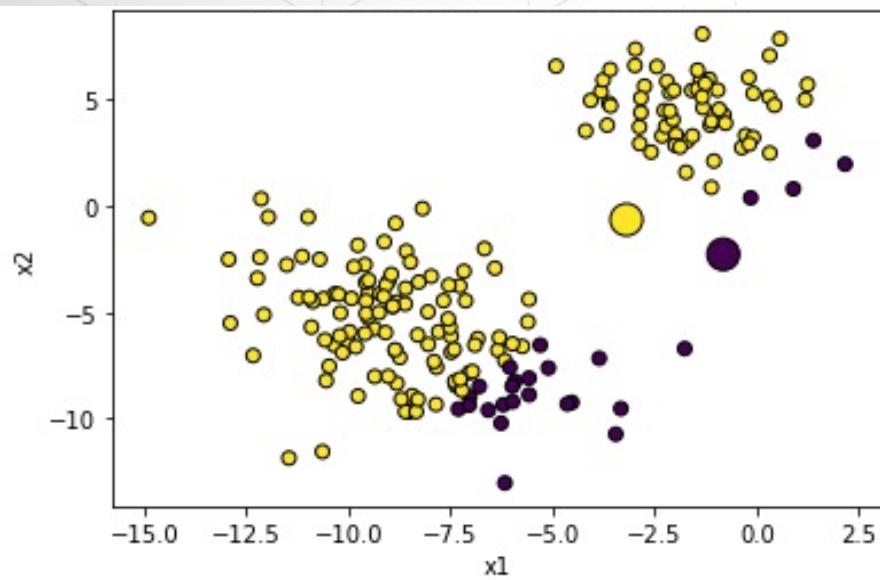
Ejemplo

- $K = 2$



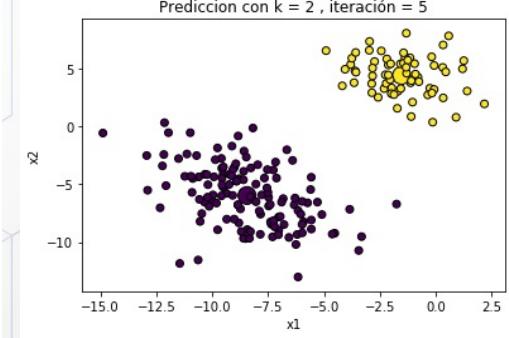
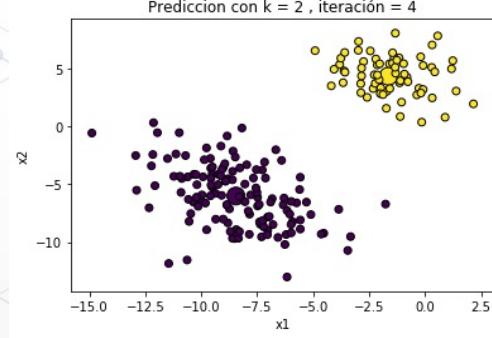
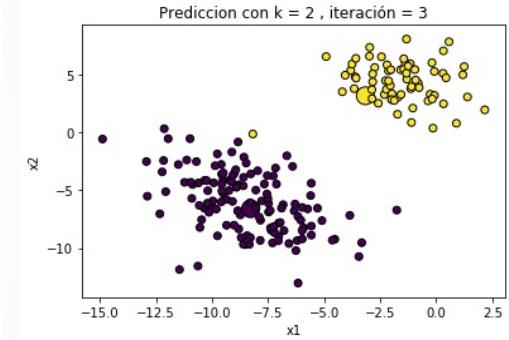
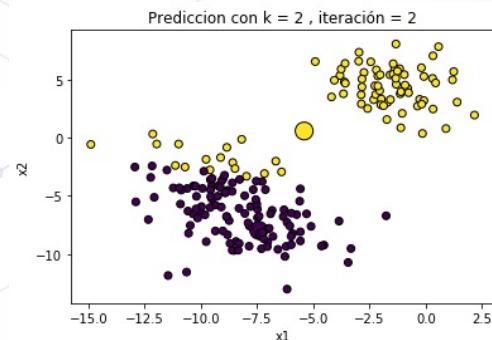
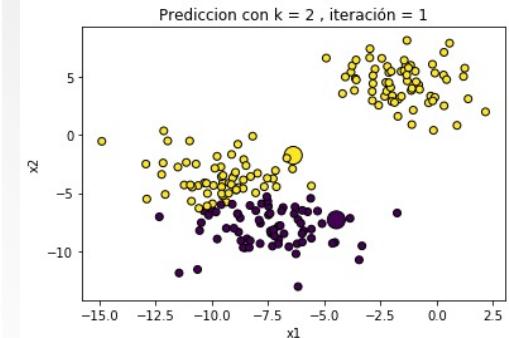
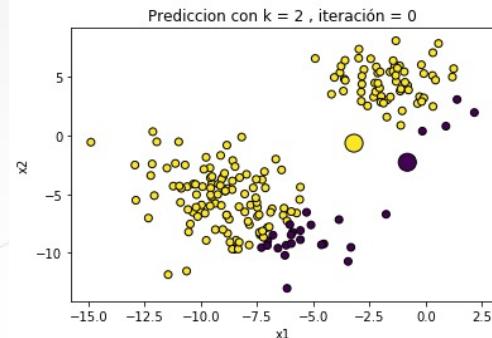
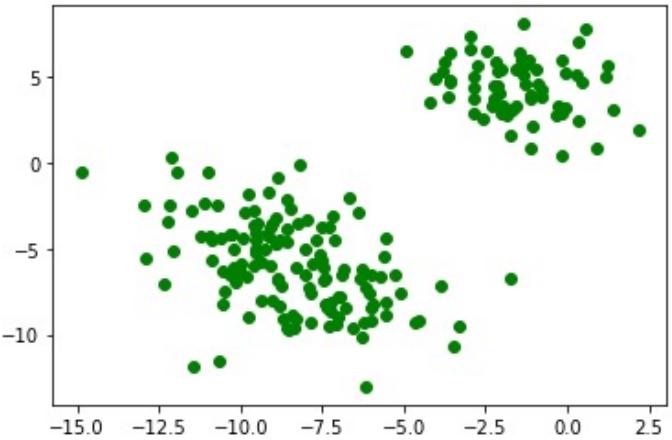
Ejemplo

- $K = 2$



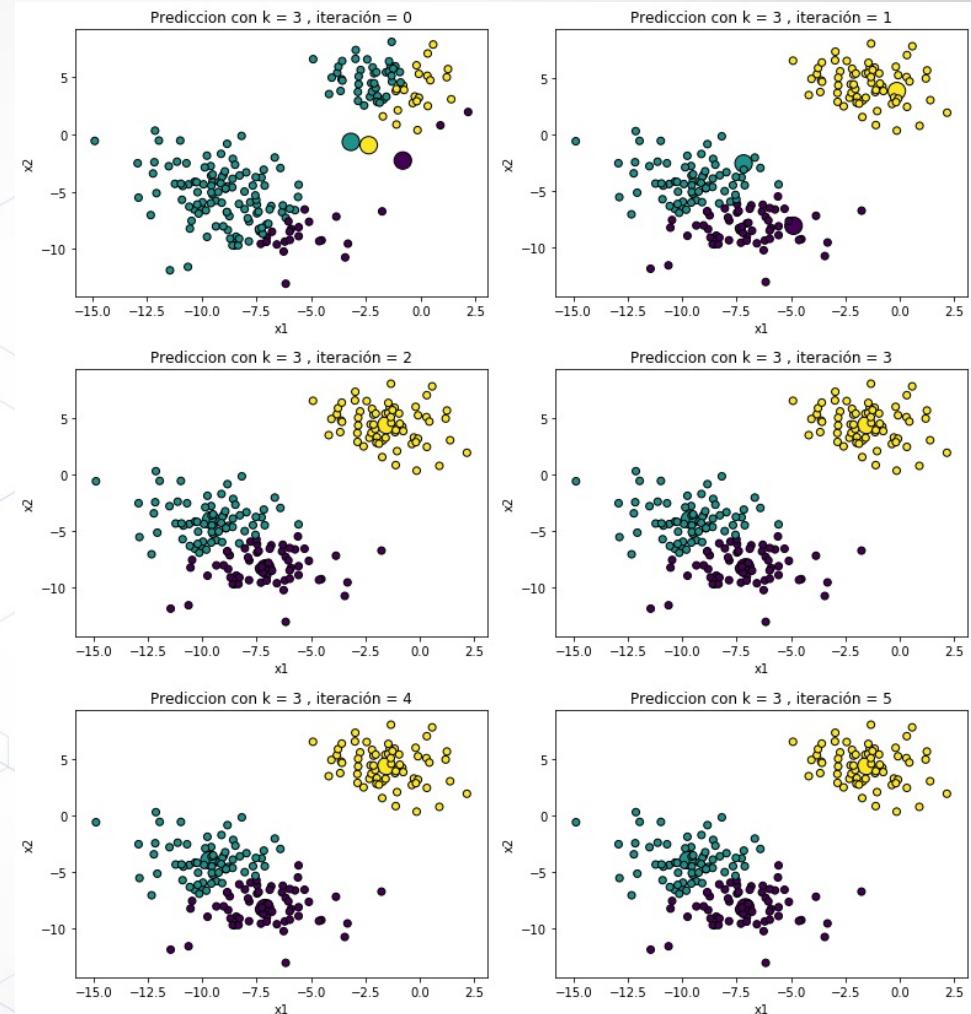
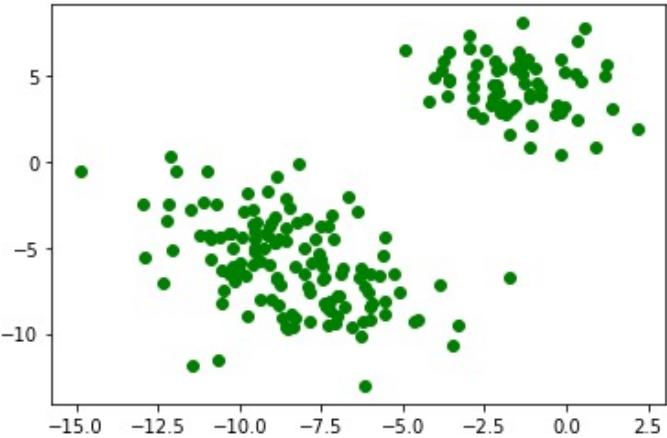
Ejemplo

- $K = 2$



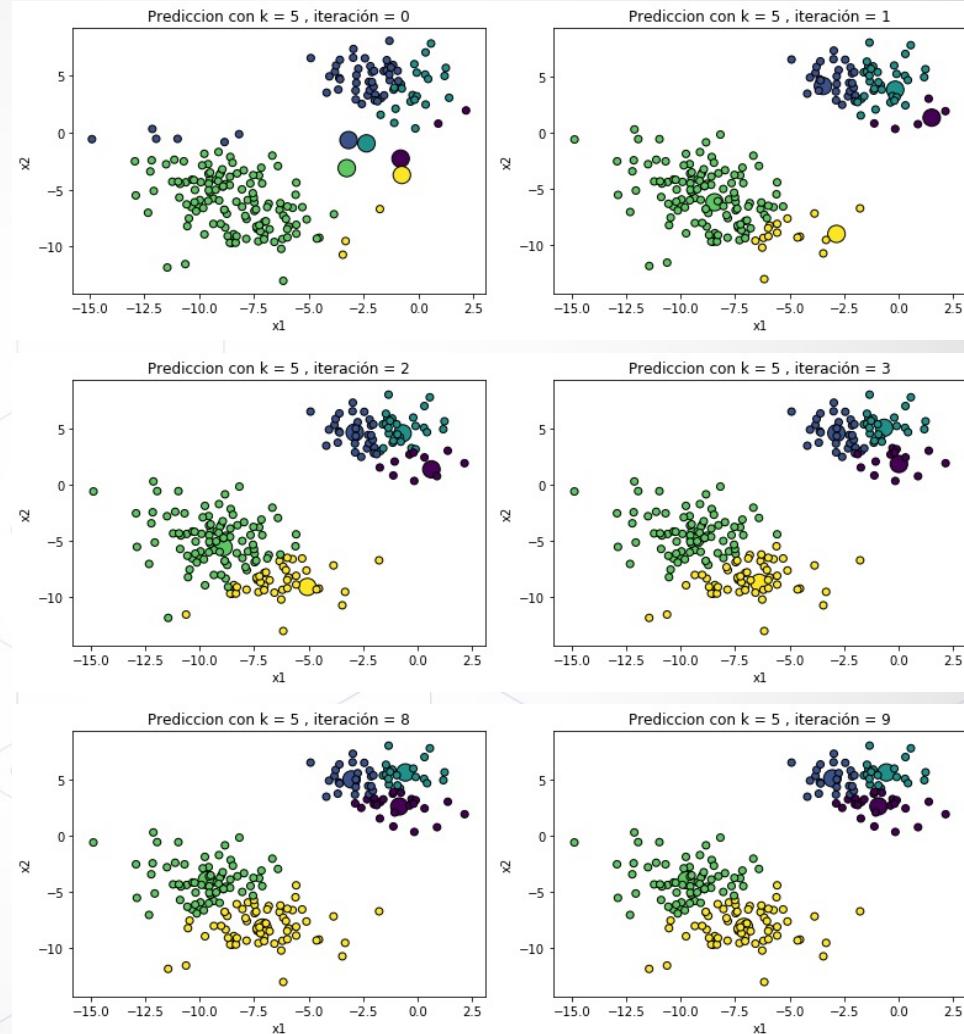
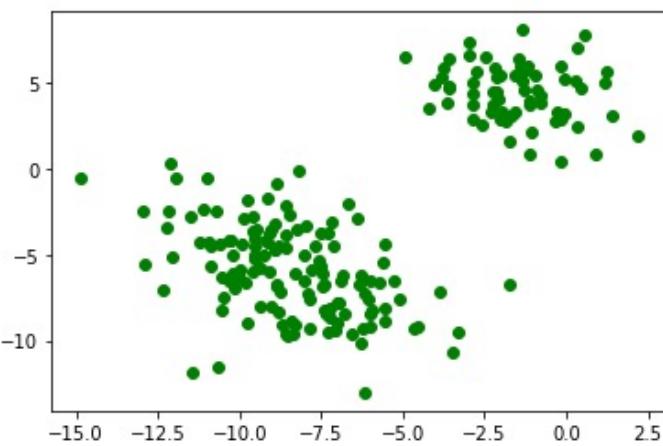
Ejemplo

- ¿Cuántos grupos / clusters?
- K = 3



Ejemplo

- ¿Cuántos grupos / clusters?
- K = 5

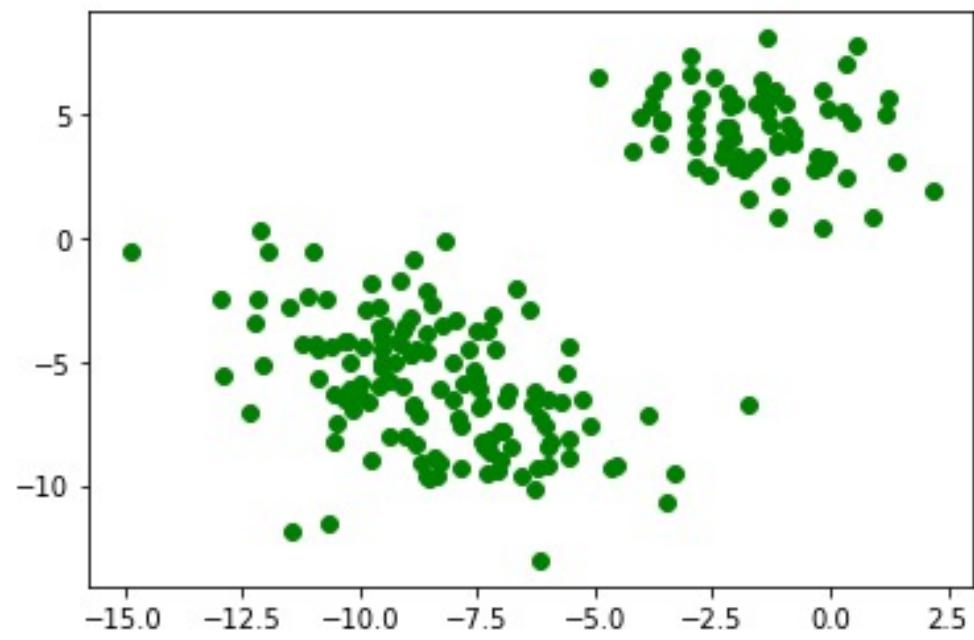


Código



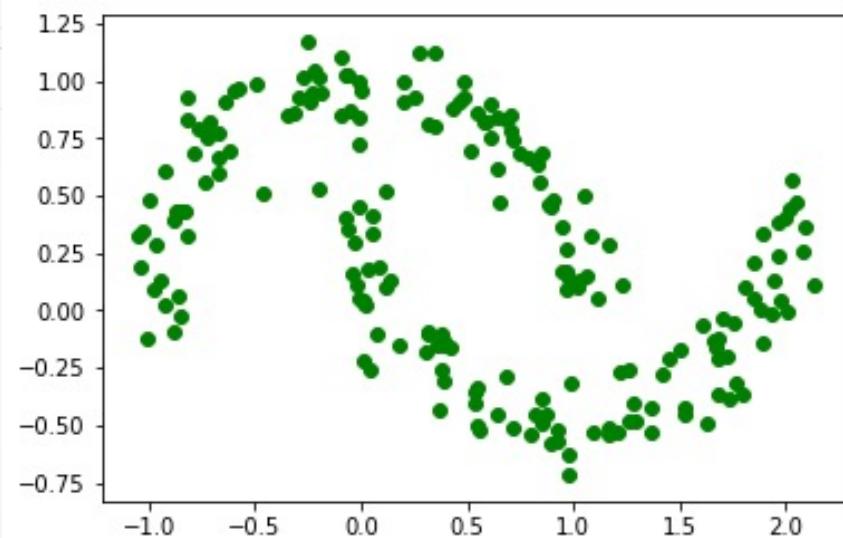
Ejemplo

- ¿Iniciando con diferentes centroides?
- $K = \{2,3,4,5\}$



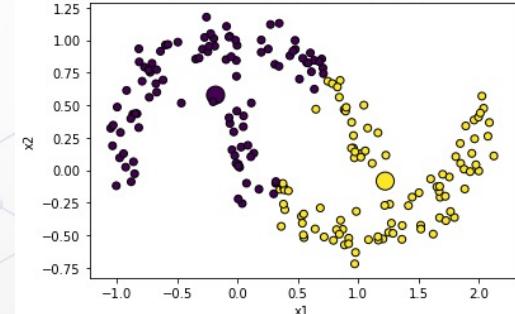
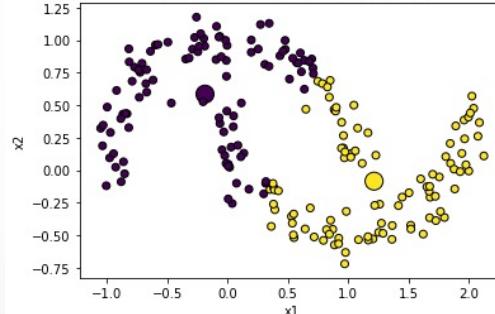
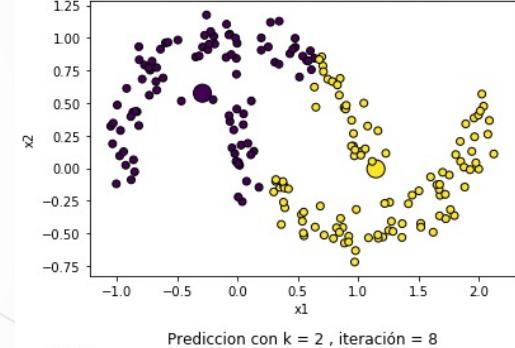
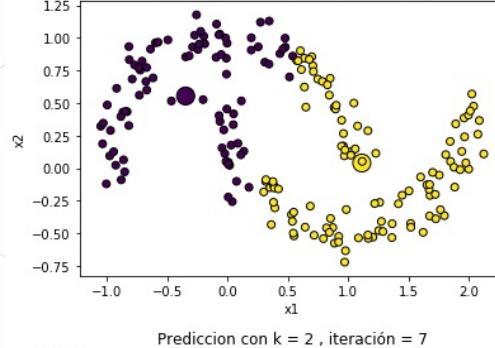
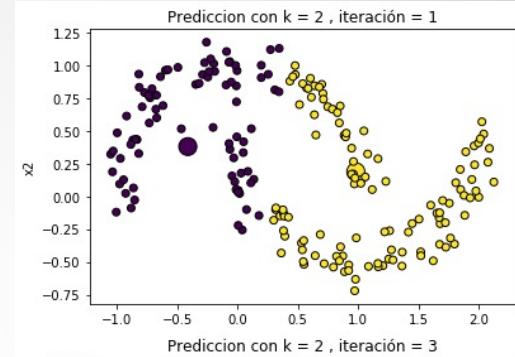
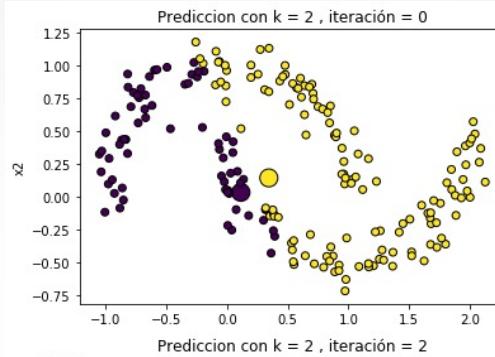
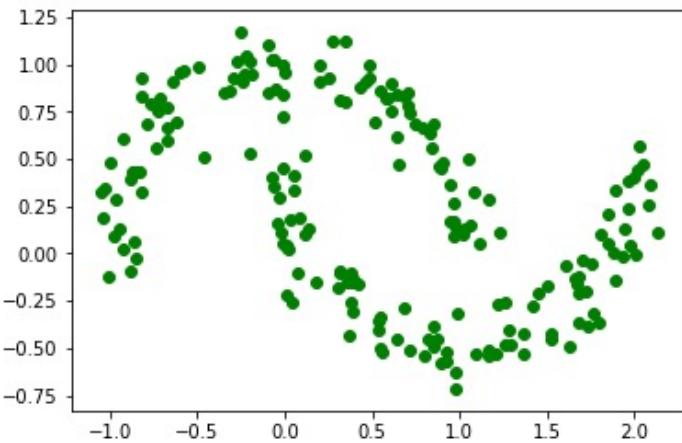
Ejemplo 2

- ¿Cuántos grupos / clusters?



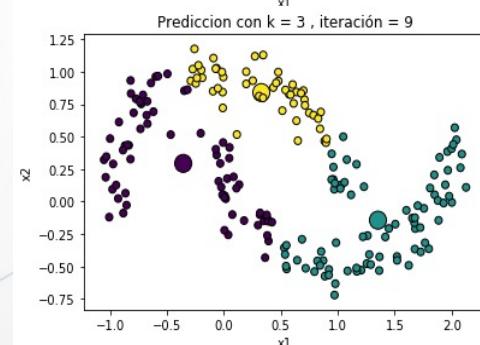
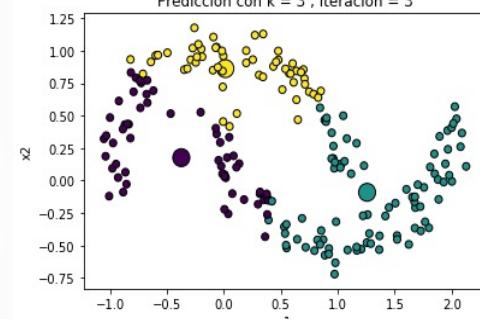
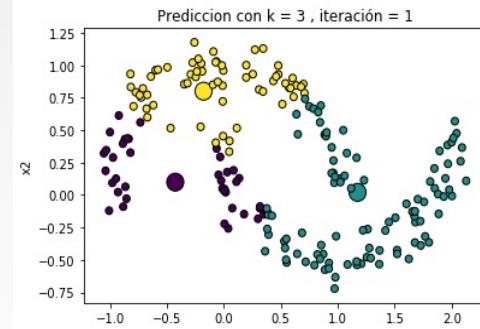
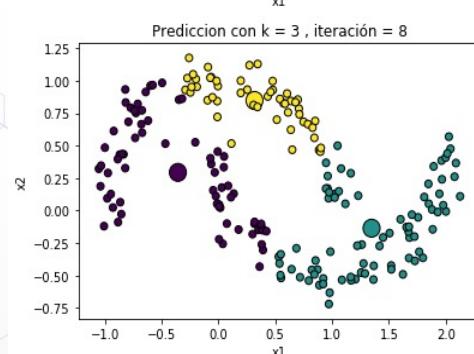
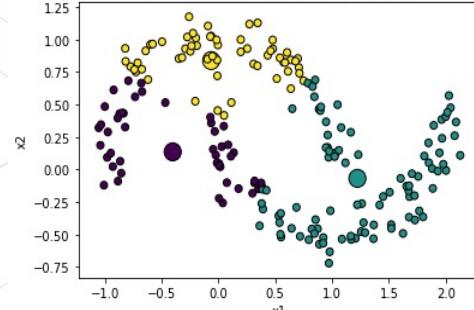
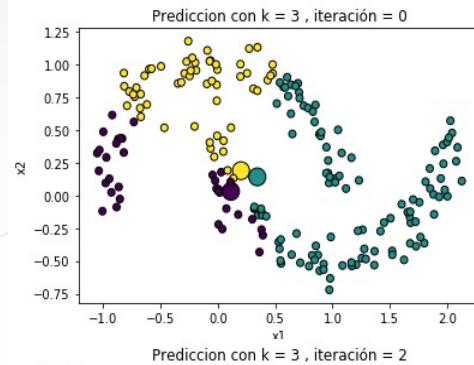
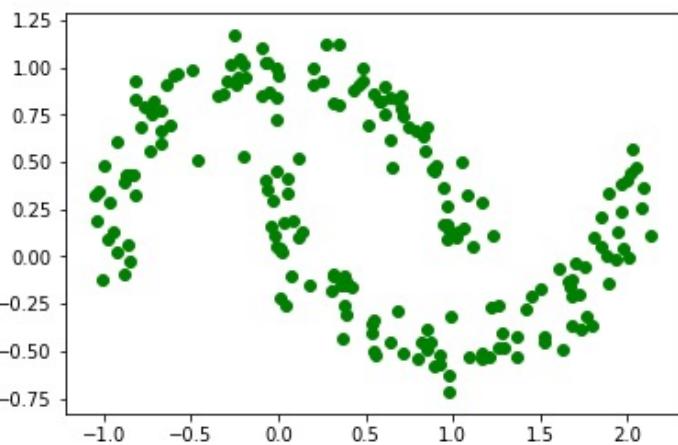
Ejemplo 2

- ¿Cuántos grupos / clusters?
- K = 2



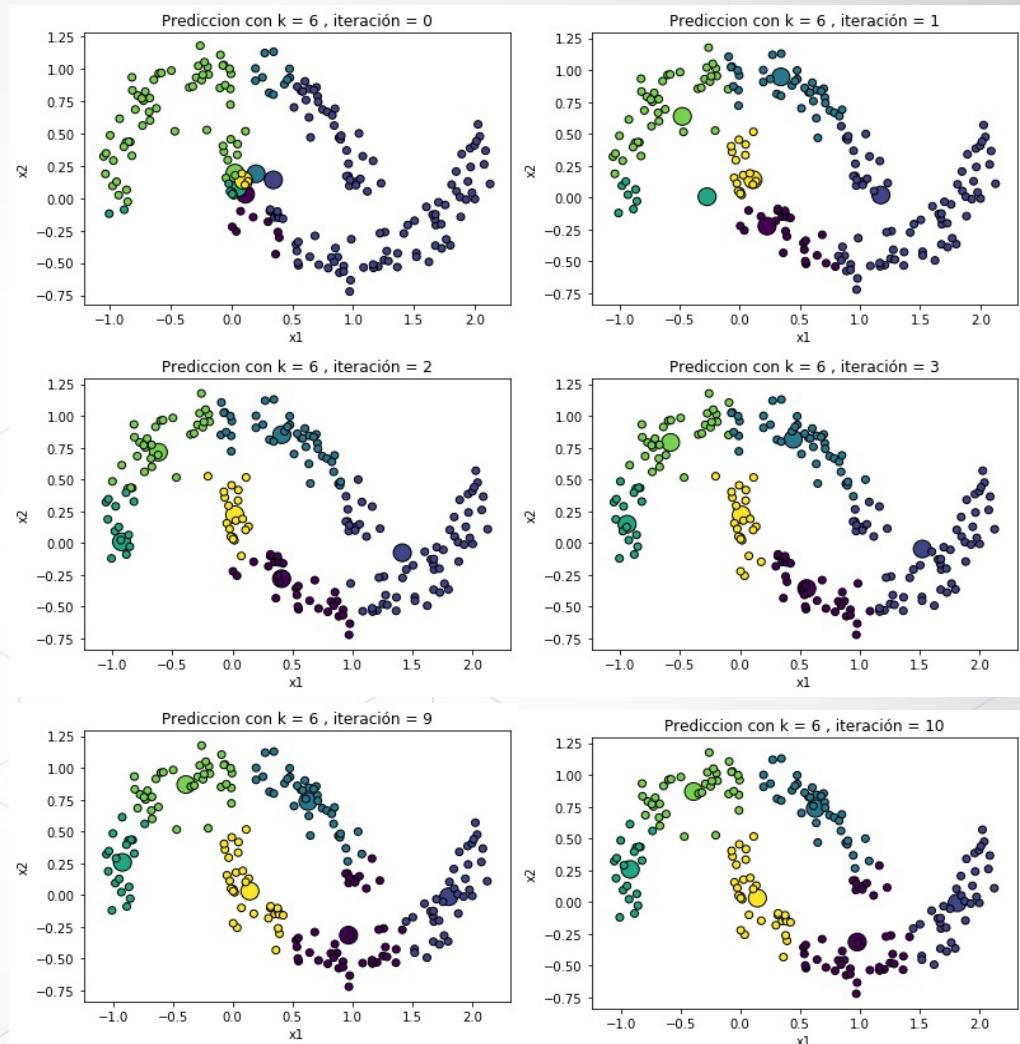
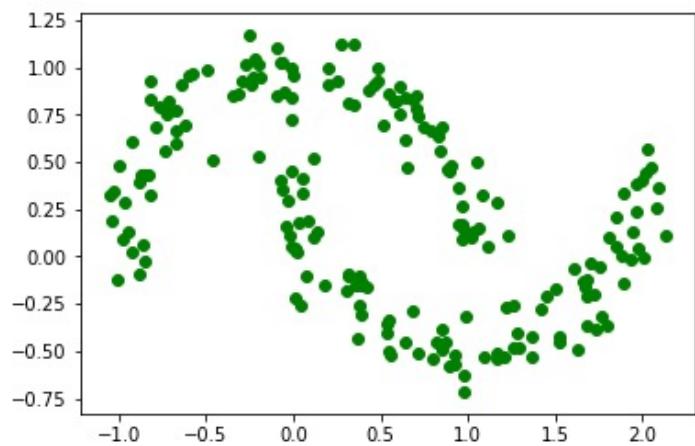
Ejemplo 2

- ¿Cuántos grupos / clusters?
- K = 3



Ejemplo 2

- ¿Cuántos grupos / clusters?
- K = 6



Remarks

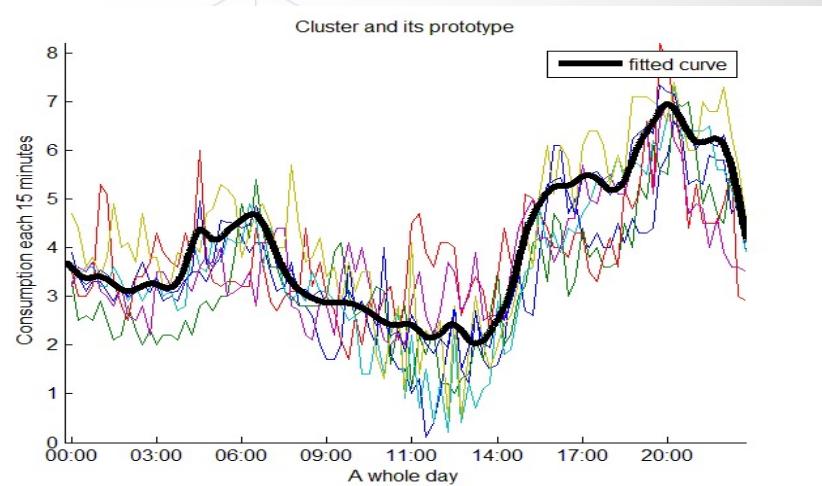
- ¿A cuántos grupos / clusters puede pertenecer una muestra a la vez?
- ¿Cómo es el traslape de clusters?

Remarks

- ¿K?

Aplicaciones

- Subgrupos de usuarios de servicios de movilidad
- Subgrupos de pacientes con algún padecimiento
- Subgrupos de películas más vistas en una plataforma
- Subgrupos de consumidores de electricidad



Código





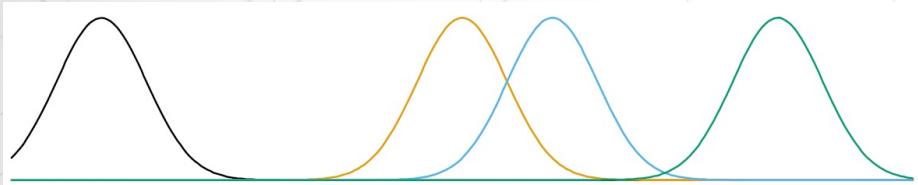
Expectation Maximization

Agrupamiento

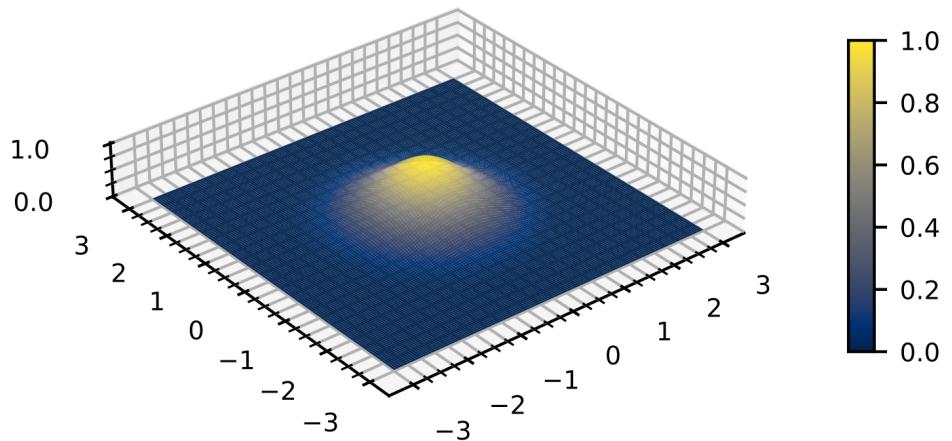


Recap

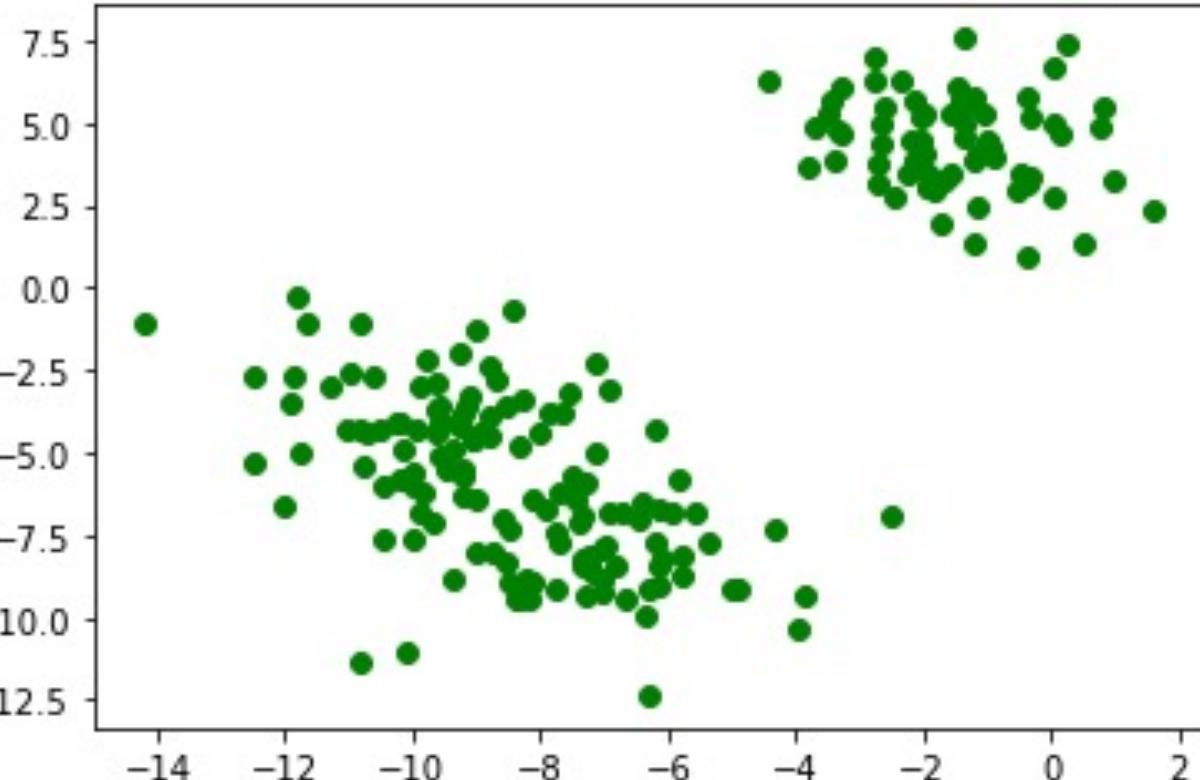
¿Elementos de una distribución gaussiana?



$$f(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

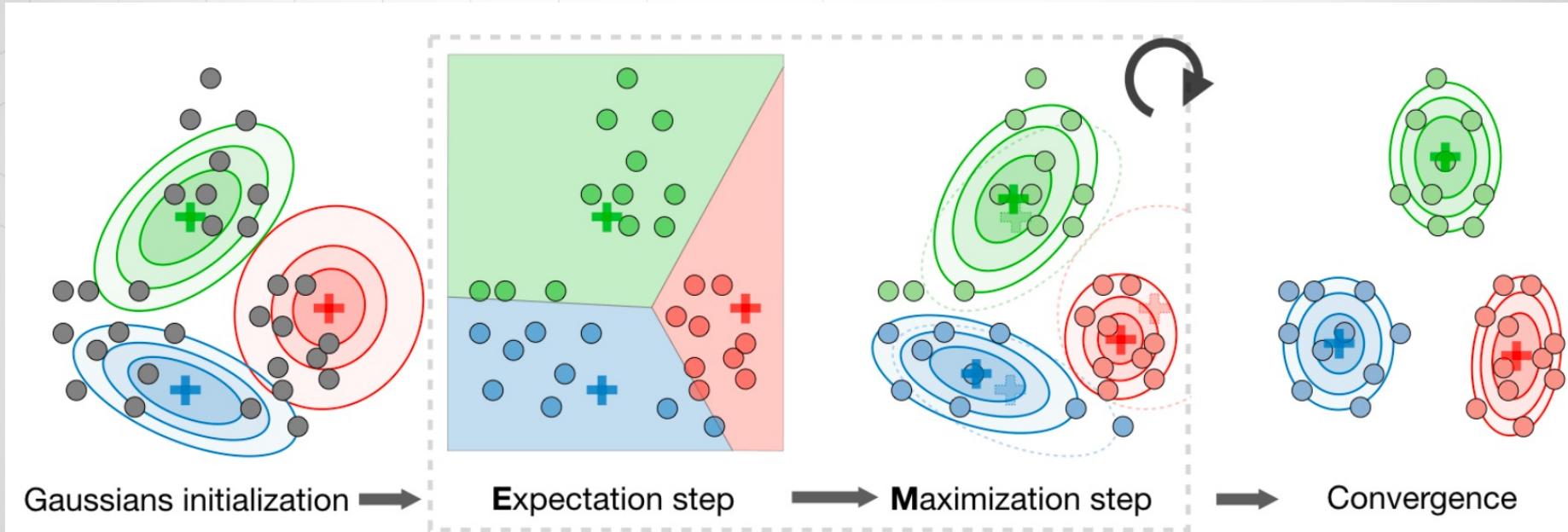


Expectation Maximization – Mezcla de Gaussianas



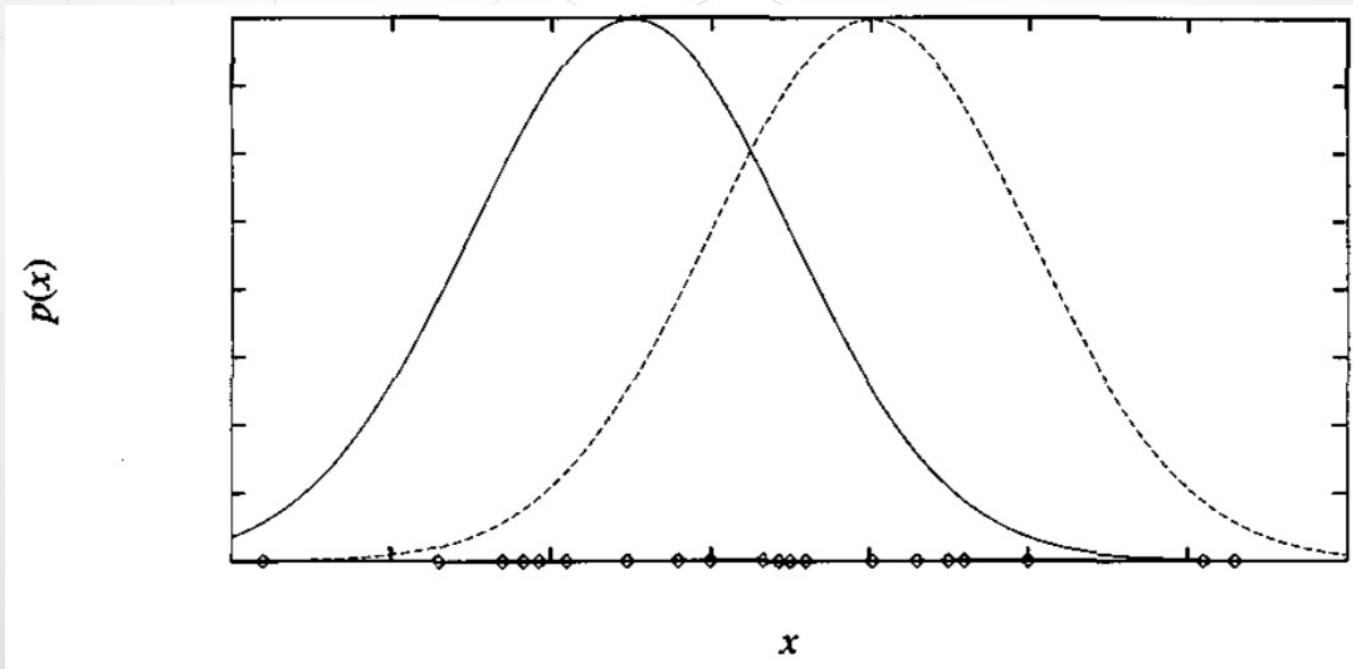
- ¿Cuántas gaussianas?
- Proceso iterativo

Expectation Maximization



Expectation Maximization

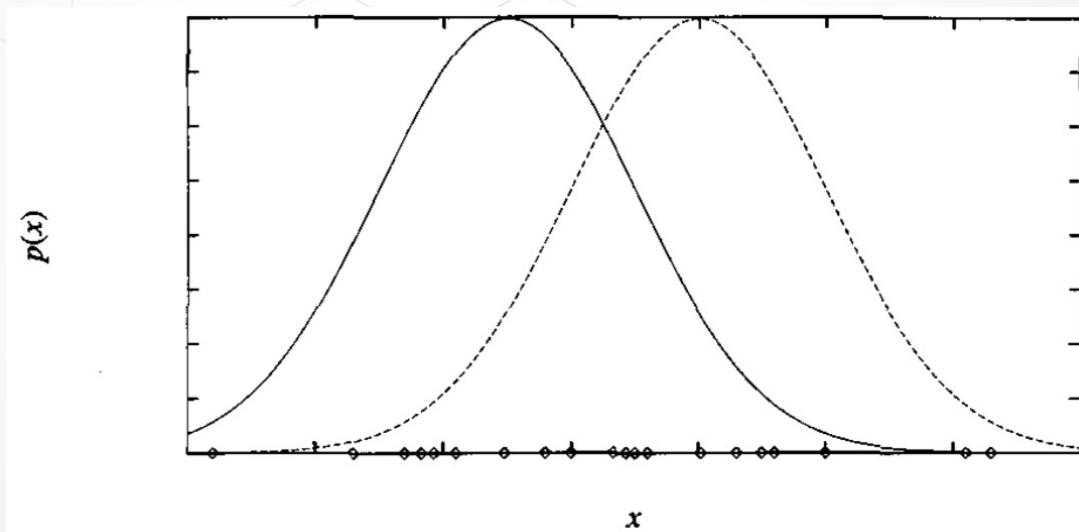
- Asumamos que los datos “provienen” de dos distribuciones



Expectation Maximization

Cada muestra puede ser representada por la tupla

$$\langle x_i, z_{i1}, z_{i2} \rangle$$



- $z_{i1} \approx 1$ Si el dato proviene de la distribución 1
 $z_{i2} \approx 1$ Si el dato proviene de la distribución 2

Expectation Maximization – Asumiendo dos agrupaciones

- Paso 1: Inicia con $M = 2$ distribuciones gaussianas aleatorias
- Paso 2: Calcula la expectativa para cada dato $E[z_{ij}]$ asumiendo que los datos vienen de la distribuciones $h = \langle \mu_1, \mu_2 \rangle$
- Paso 3: Calcula una nueva máxima verosimilitud $h' = \langle \mu'_1, \mu'_2 \rangle$ tomando los valores calculados en el primer paso ($E[z_{ij}]$). Reemplazar $h = \langle \mu_1, \mu_2 \rangle$ con $h' = \langle \mu'_1, \mu'_2 \rangle$
- Paso 4: Repetir paso 2 y 3 hasta convergencia

Expectation Maximization

- Paso 2:

$$E[z_{ij}] = \frac{p(x = x_i | \mu = \mu_j)}{\sum_{n=1}^N p(x = x_i | \mu = \mu_n)}$$

$$= \frac{e^{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}}}{\sum_{n=1}^N e^{-\frac{(x_i - \mu_n)^2}{2\sigma_n^2}}}$$

Probabilidad
de que el dato
 x_i haya sido
generado por
la gaussiana j

- Paso 3:

$$\mu_j = \frac{\sum_{i=1}^m E[z_{ij}] x_i}{\sum_{i=1}^m E[z_{ij}]}$$

$$\sigma_j^2 = \frac{\sum_{i=1}^m E[z_{ij}] (x_i - \mu_j)^2}{\sum_{i=1}^m E[z_{ij}]}$$

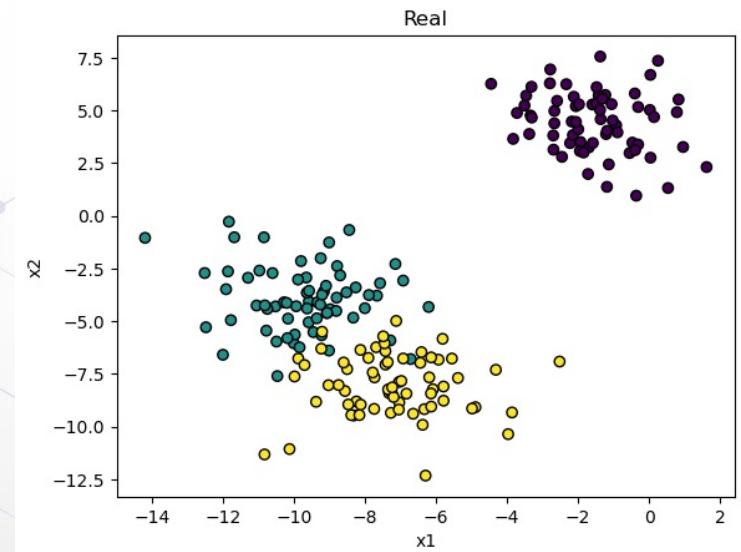
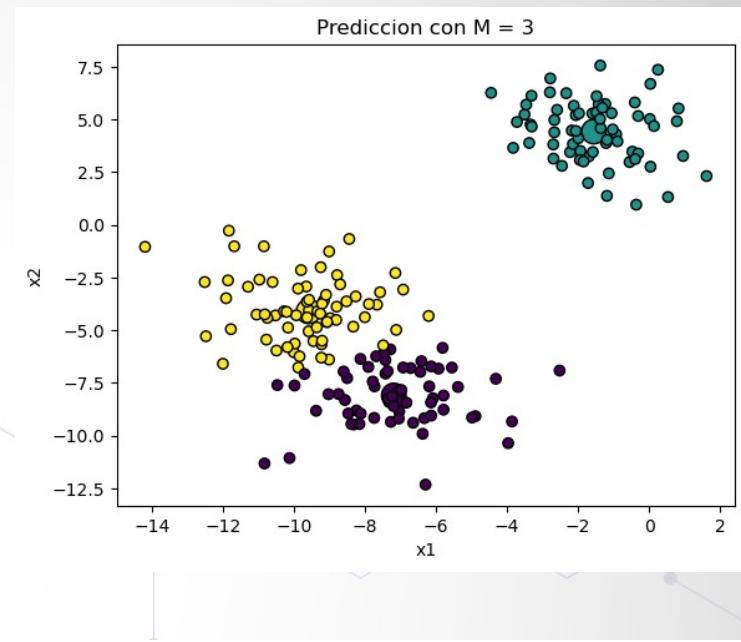
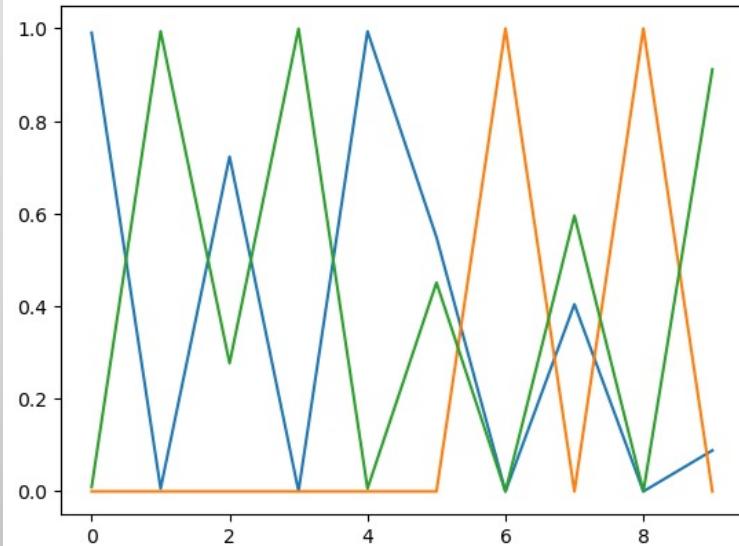
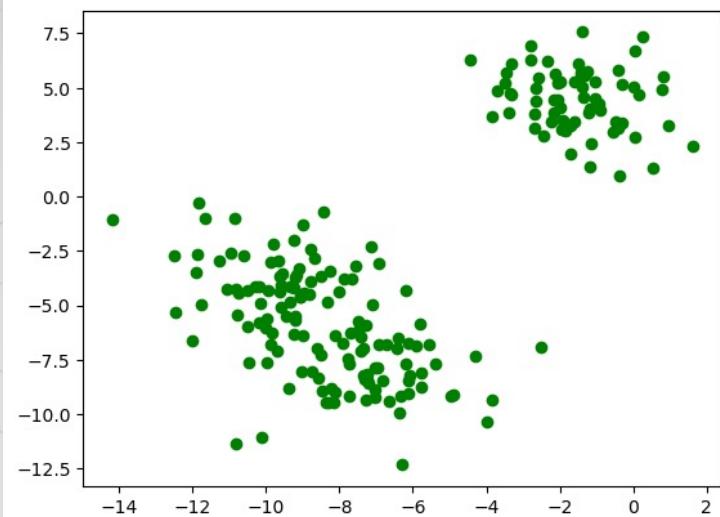
Expectation Maximization

- Se busca el grupo de clusters más probables dados los datos
- Los objetos tienen cierta “probabilidad” de pertenecer a un grupo o cluster
- El algoritmo EM (Expectation Maximization) empieza adivinando los parámetros de las distribuciones y los usa para calcular las probabilidades de que cada objeto pertenezca a un cluster

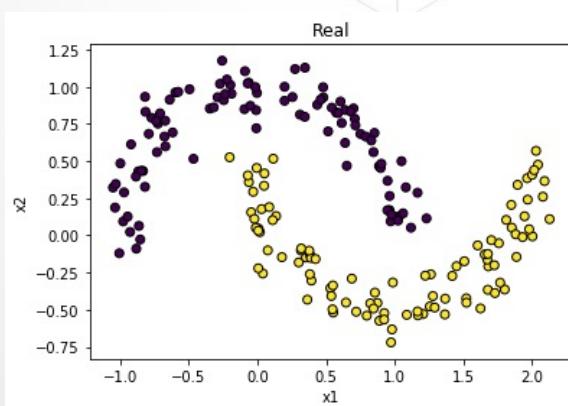
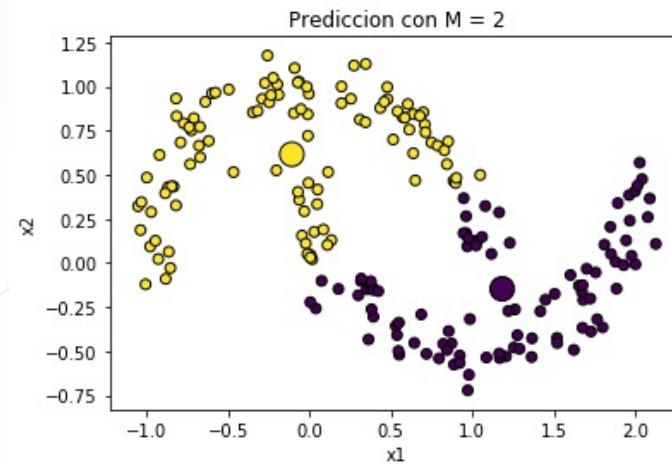
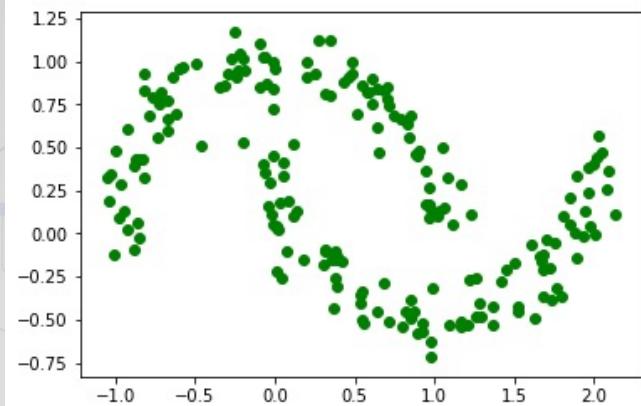
Expectation Maximization

- El cálculo de las **probabilidades de las clases** o los valores esperados de las clases es la parte de **expectation**
- El paso de calcular los valores de los parámetros de las distribuciones, es maximization, **maximar la verosimilitud de las distribuciones** dados los datos
- Para estimar los parámetros, tenemos que considerar que tenemos únicamente las probabilidades de pertenecer a cada cluster y no los clusters en si

Ejemplo 1



Ejemplo 2



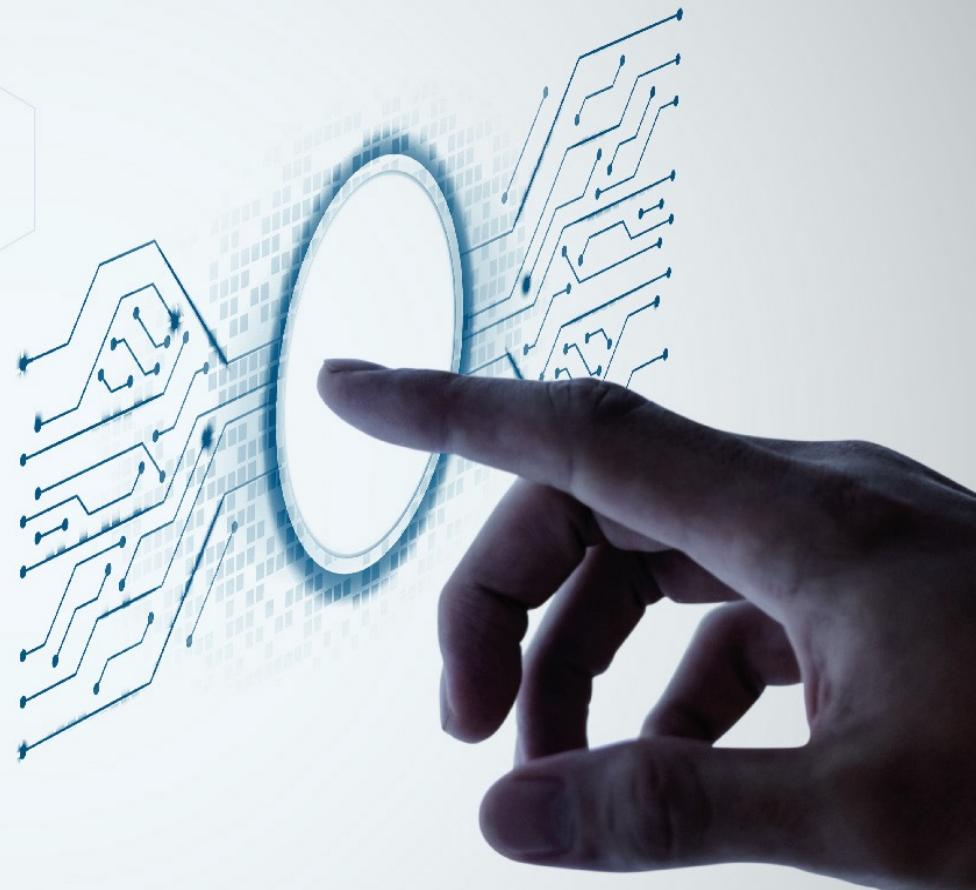
Código



Engineering X

Founded by the Royal Academy of Engineering
and Lloyd's Register Foundation

GRACIAS



<https://hubiq.mx/>

HUBIQURO HUBIQURO HUBIQURO