

- El marketing es fundamental para el crecimiento y la sostenibilidad de cualquier negocio.
- Los especialistas en marketing pueden ayudar a desarrollar la marca de la empresa, atraer clientes, aumentar los ingresos y aumentar las ventas.



- **Uno de los puntos críticos para los especialistas en marketing es conocer a sus clientes e identificar sus necesidades.**
- **Al comprender al cliente, los especialistas en marketing pueden lanzar una campaña de marketing dirigida que se adapte a necesidades específicas.**
- **Si los datos sobre los clientes están disponibles, la ciencia de datos se puede aplicar para realizar la segmentación del mercado.**



- **En este estudio de caso, te han contratado como consultor de un banco en la ciudad de Nueva York.**
- **El banco tiene datos extensos sobre sus clientes durante los últimos 6 meses.**
- **El equipo de marketing del banco desea lanzar una campaña de marketing publicitaria dirigida dividiendo a sus clientes en al menos 3 grupos distintos.**



- **# CUSTID:** Identificación del titular de la tarjeta de crédito
- **# BALANCE:** Cantidad de saldo que queda en la cuenta del cliente para hacer compras
- **# BALANCE_FREQUENCY:** Frecuencia de la actualización del saldo, puntuación entre 0 y 1 (1 = actualizado con frecuencia, 0 = no actualizado con frecuencia)
- **# PURCHASES:** Cantidad de compras realizadas desde la cuenta
- **# ONEOFFPURCHASES:** Importe máximo de compra realizado en una sola vez
- **# INSTALLMENTS_PURCHASES:** Importe de la compra realizada en cuotas
- **# CASH_ADVANCE:** Anticipo otorgado al usuario
- **# PURCHASES_FREQUENCY:** frecuencia con la que se realizan las compras, puntuación entre 0 y 1 (1 = compras frecuentes, 0 = compras no frecuentes)



- **# PURCHASES_FREQUENCY:** Frecuencia de las Compras se están realizando, puntuación entre 0 y 1 (1 = compra con frecuencia, 0 = no compra con frecuencia)
- **# ONEOFF_PURCHASES_FREQUENCY:** Con qué frecuencia las compras se realizan de una sola vez (1 = compra con frecuencia, 0 = no compra con frecuencia)
- **# PURCHASES_INSTALLMENTS_FREQUENCY:** Con qué frecuencia se realizan las compras a plazos (1 = se realizan con frecuencia, 0 = no se realizan con frecuencia)
- **# CASH_ADVANCE_FREQUENCY:** con qué frecuencia el gasto se paga por adelantado

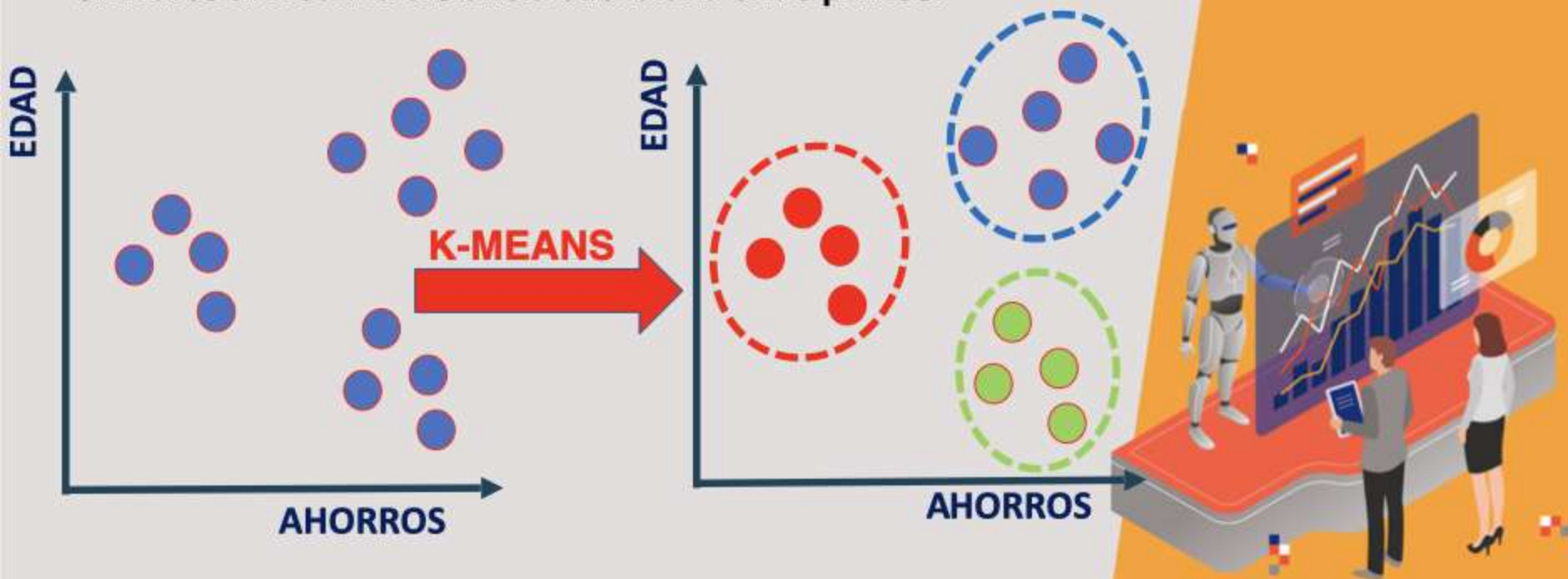


- **# CASH_ADVANCE_TRX:** número de transacciones realizadas con "Efectivo por adelantado"
- **# PURCHASES_TRX:** número de transacciones de compras realizadas
- **# CREDIT_LIMIT:** límite de tarjeta de crédito para el usuario
- **# PAYMENTS:** Número de pagos realizados por el usuario
- **# MINIMUM_PAYMENTS:** cantidad mínima de pagos realizados por el usuario
- **# PRC_FULL_PAYMENT:** porcentaje del pago total pagado por el usuario
- **# TENURE:** Permanencia del servicio de tarjeta de crédito para el usuario



INTUICIÓN DETRÁS DE K-MEANS

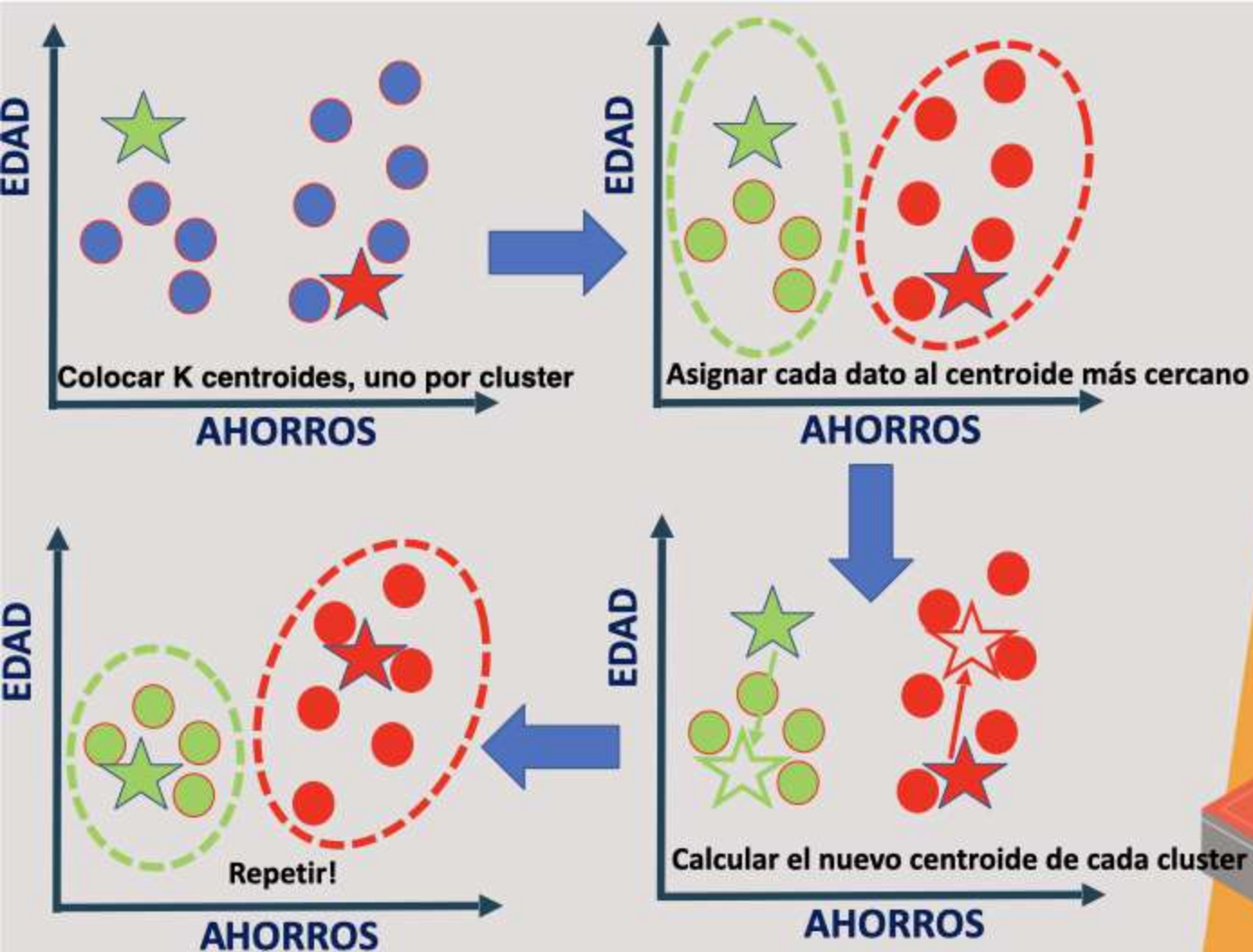
- K-means es un algoritmo de aprendizaje no supervisado (clustering).
- K-means funciona agrupando algunos puntos de datos (clustering) de forma no supervisada.
- El algoritmo agrupa las observaciones con valores de atributos similares al medir la distancia euclidiana entre puntos.



PASOS DEL ALGORITMO DE K-MEANS

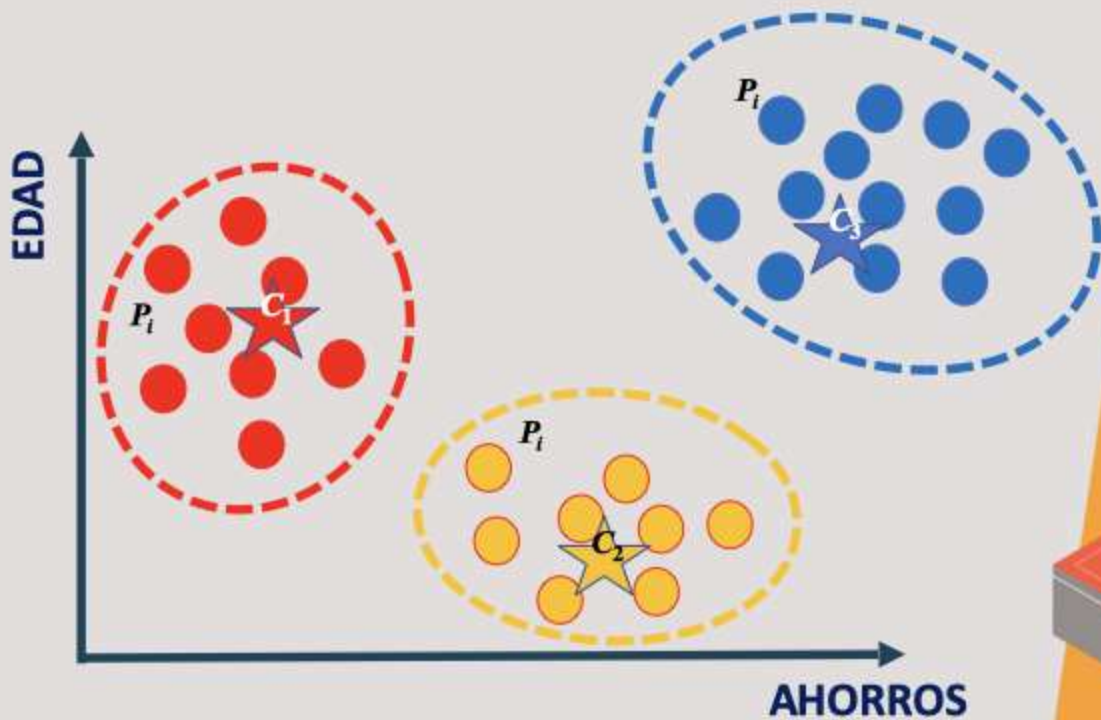
1. Elegir el número de clústers "K"
2. Seleccionar K puntos aleatorios que serán los centroides de cada cluster
3. Asignar cada punto del dataset al centroide más cercano, hacerlo nos permitirá crear un número "K" de clústers con dichos puntos
4. Calcular un nuevo centroide para cada cluster
5. Reasignar cada punto de datos al nuevo centroide más cercano
6. Ir al paso 4 y repetir.





¿CÓMO SELECCIONAR EL NÚMERO ÓPTIMO DE CLÚSTERS (K)? "MÉTODO DE CODO"

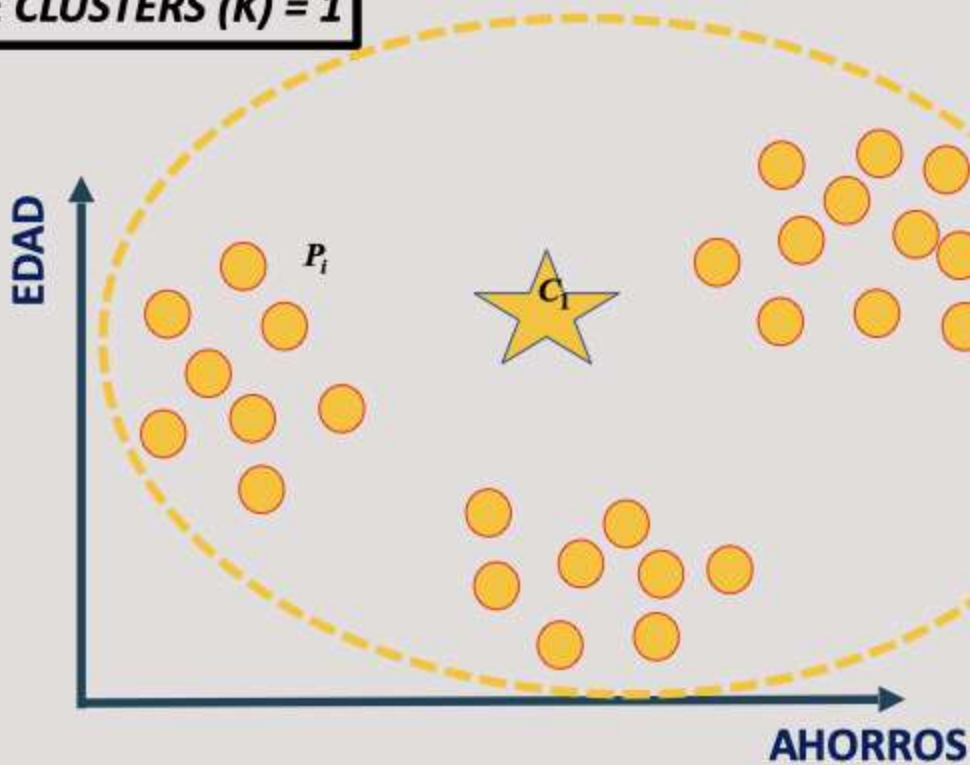
$$\text{Within Cluster Sum of Squares (WCSS)} = \sum_{P_i \text{ in Cluster 1}} \text{distance}(P_i, C_1)^2 + \sum_{P_i \text{ in Cluster 2}} \text{distance}(P_i, C_2)^2 + \sum_{P_i \text{ in Cluster 3}} \text{distance}(P_i, C_3)^2$$



¿CÓMO SELECCIONAR EL NÚMERO ÓPTIMO DE CLÚSTERS (K)? "MÉTODO DE CODO"

$$\text{Within Cluster Sum of Squares (WCSS)} = \sum_{P_i \text{ in Cluster } 1} \text{distance}(P_i, C_1)^2$$

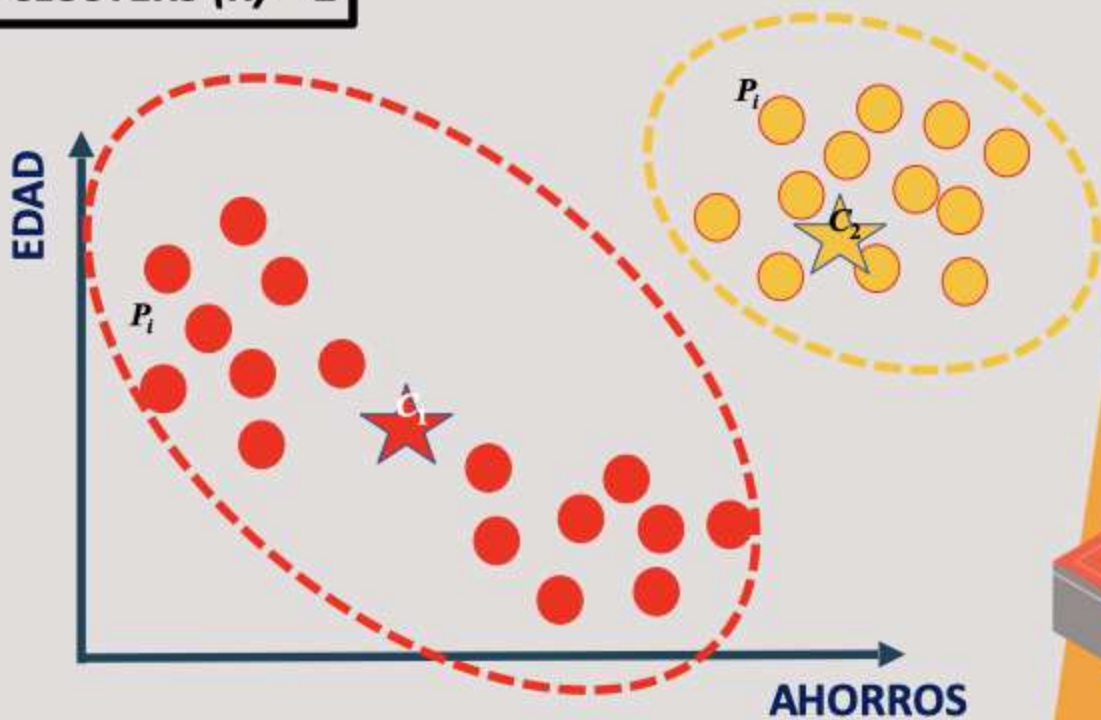
NÚMERO DE CLUSTERS (K) = 1



¿CÓMO SELECCIONAR EL NÚMERO ÓPTIMO DE CLÚSTERS (K)? "MÉTODO DE CODO"

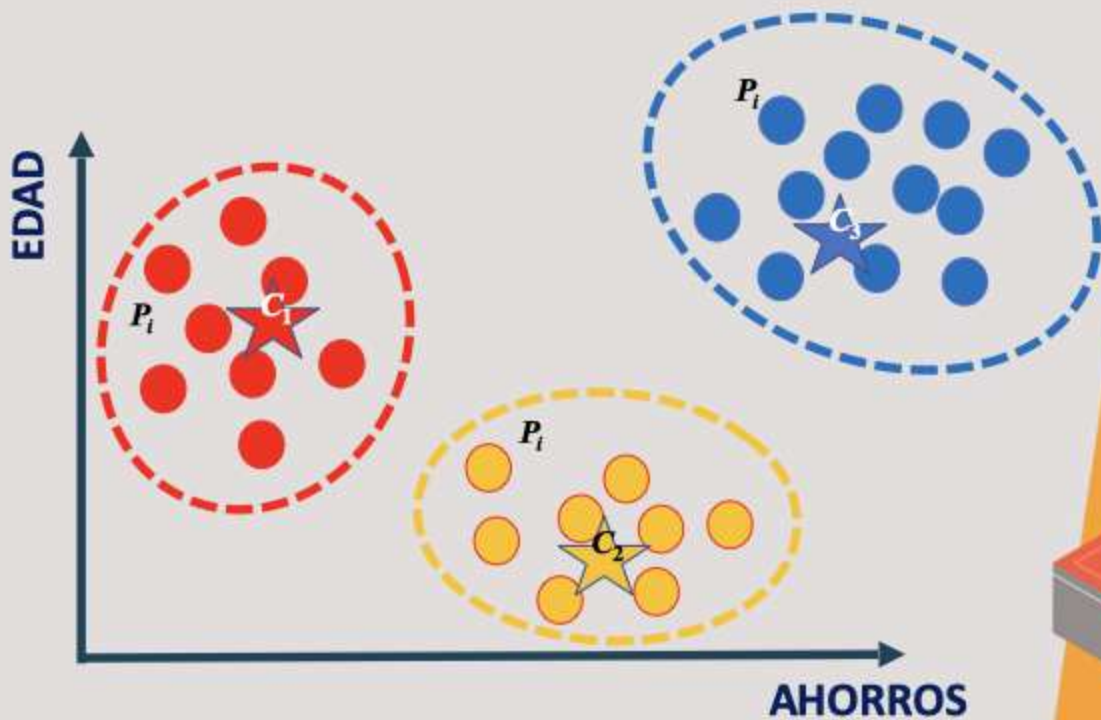
$$\text{Within Cluster Sum of Squares (WCSS)} = \sum_{P_i \text{ in Cluster 1}} \text{distance}(P_i, C_1)^2 + \sum_{P_i \text{ in Cluster 2}} \text{distance}(P_i, C_2)^2$$

NÚMERO DE CLUSTERS (K) = 2

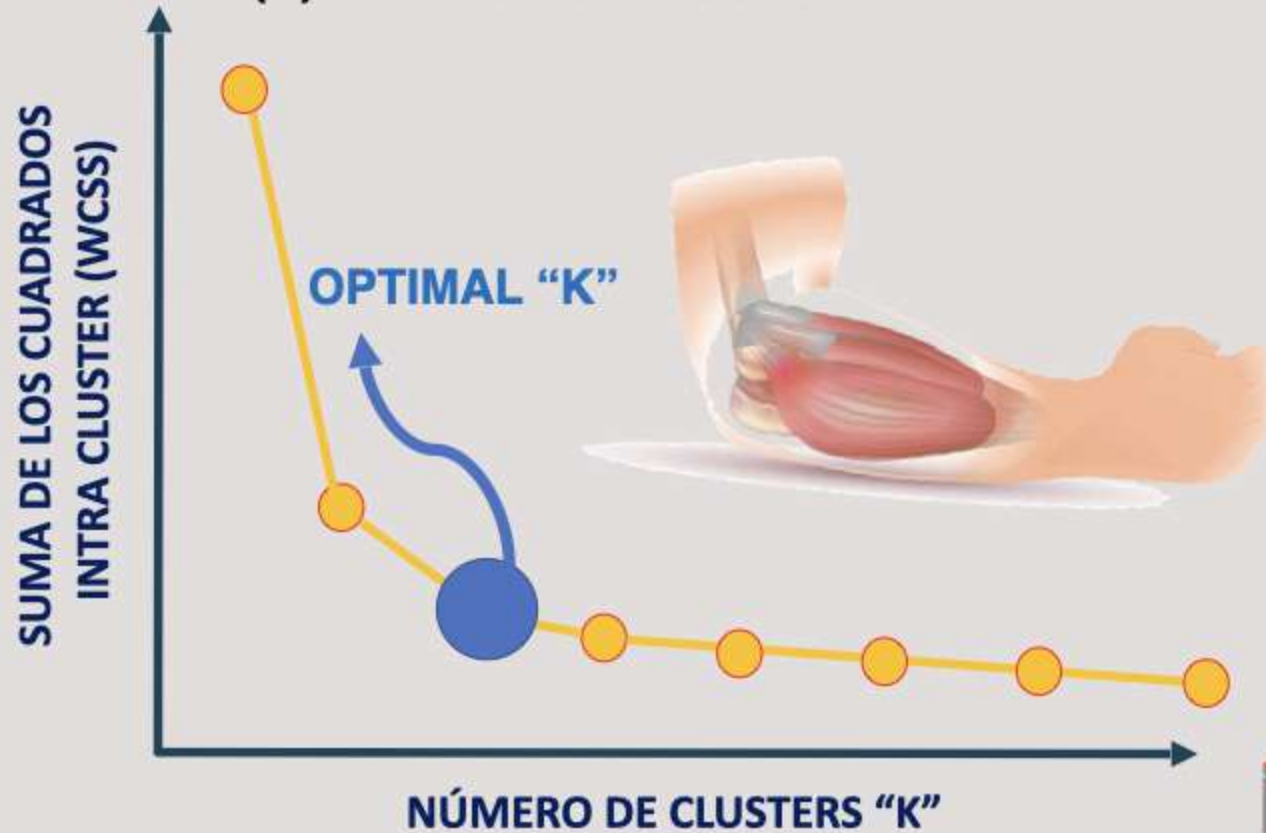


¿CÓMO SELECCIONAR EL NÚMERO ÓPTIMO DE CLÚSTERS (K)? "MÉTODO DE CODO"

$$\text{Within Cluster Sum of Squares (WCSS)} = \sum_{P_i \text{ in Cluster 1}} \text{distance}(P_i, C_1)^2 + \sum_{P_i \text{ in Cluster 2}} \text{distance}(P_i, C_2)^2 + \sum_{P_i \text{ in Cluster 3}} \text{distance}(P_i, C_3)^2$$



¿CÓMO SELECCIONAR EL NÚMERO ÓPTIMO DE CLÚSTERS (K)? "MÉTODO DE CODO"



Source: https://commons.wikimedia.org/wiki/File:Tennis_Elbow_Illustration.jpg



INTUICIÓN DETRÁS DE LOS AUTOENCODERS

- Los autoencoders son un tipo de redes neuronales artificiales que se utilizan para realizar una tarea de codificación de datos (representation learning).
- Los autoencoders utilizan los mismos datos del dataset para la entrada y la salida. Suena loco, ¿verdad?

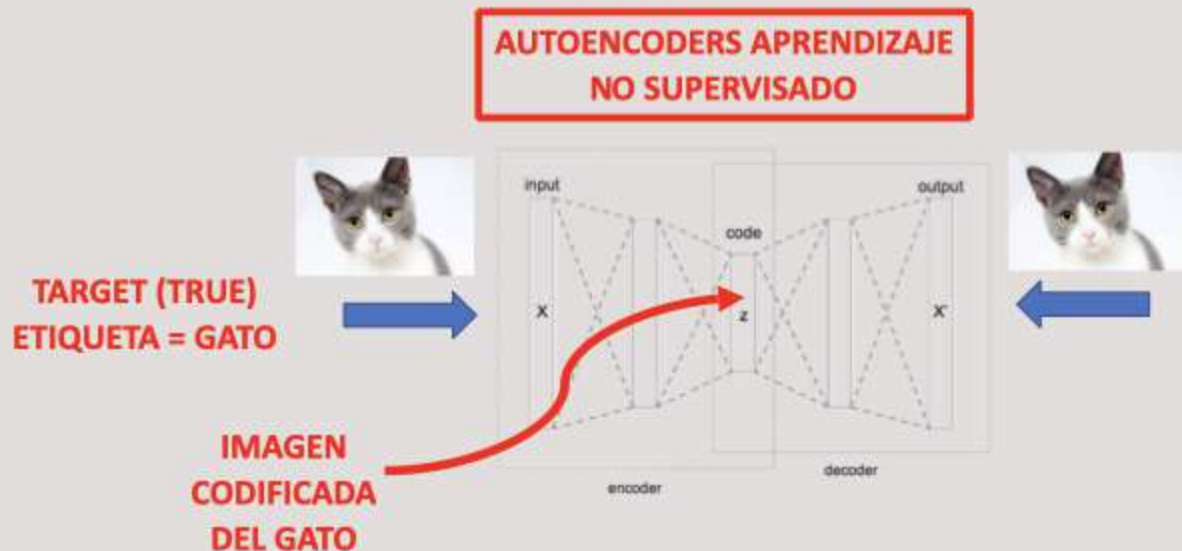


Photo Credit: https://commons.wikimedia.org/wiki/File:Autoencoder_structure.png

Photo Credit: https://commons.wikimedia.org/wiki/File:Artificial_neural_network_imEDAD_recognition.png

Photo Credit: <https://www.pexels.com/photo/grey-and-white-short-fur-cat-104827/>



LA CAPA DE CODIFICACIÓN

- Los autoencoders funcionan agregando un cuello de botella en la red.
- Este cuello de botella obliga a la red a crear una versión comprimida (codificada) de la entrada original.
- Los autoencoders funcionan bien si existen correlaciones entre los datos de entrada (funcionan mal si todos los datos de entrada son independientes)

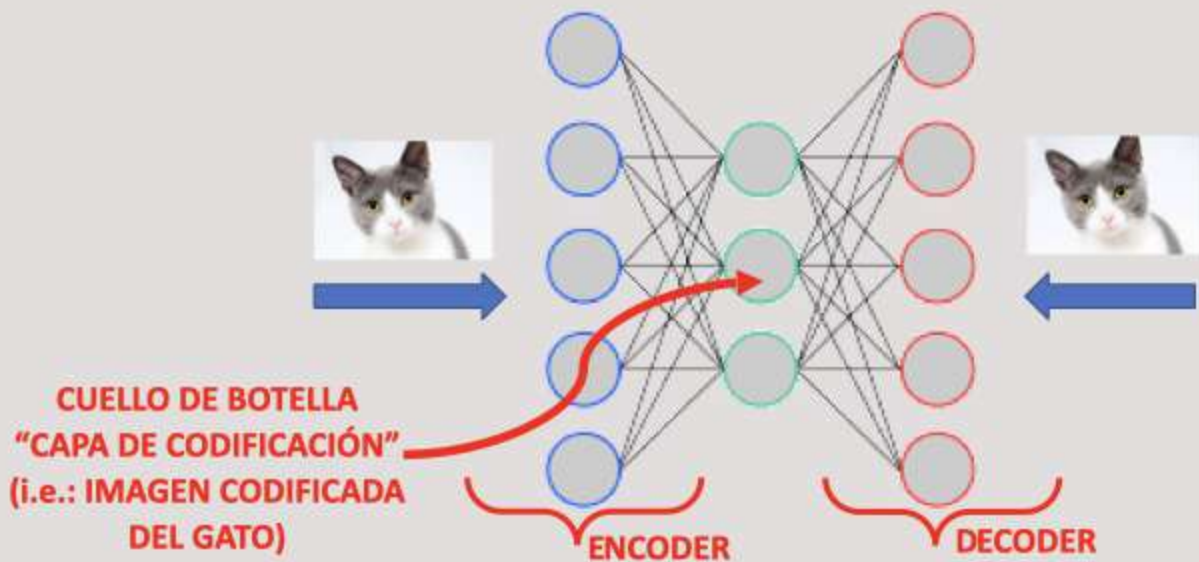


Photo Credit: https://commons.wikimedia.org/wiki/File:Autoencoder_structure.png

Photo Credit: https://commons.wikimedia.org/wiki/File:Artificial_neural_network_imEDAD_recognition.png

Photo Credit: <https://www.pexels.com/photo/grey-and-white-short-fur-cat-104827/>

LAS MATEMÁTICAS DEL AUTOENCODER

ENCODER:

$$h(x) = \text{sigmoid}(W * x + b)$$

DECODER:

$$\hat{x} = \text{sigmoid}(W^* * h(x) + c)$$

PESOS LIGADOS:

Los pesos de la entrada a la capa oculta serán iguales a los pesos de la capa oculta a la salida

$$W^* = W^T$$

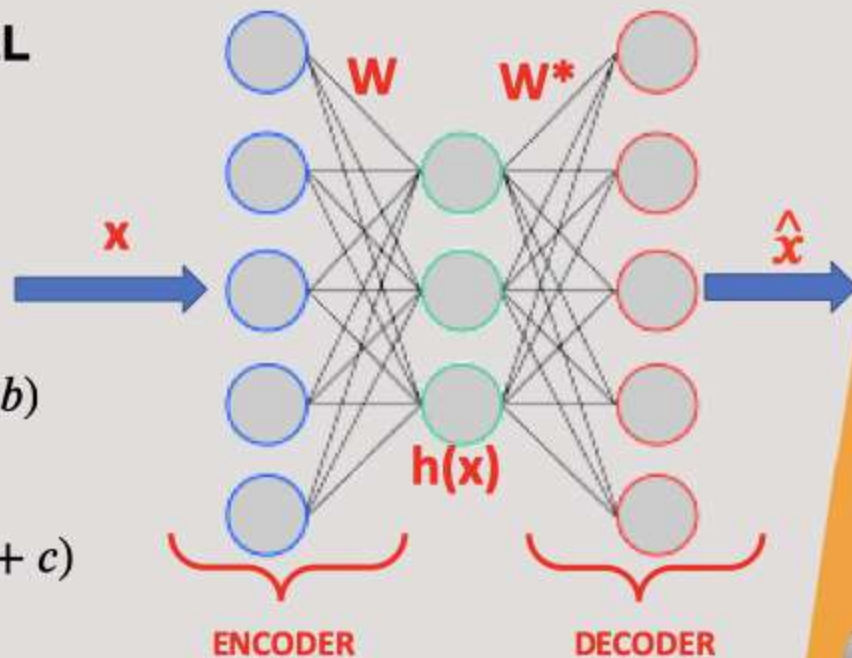


Photo Credit: https://commons.wikimedia.org/wiki/File:Autoencoder_structure.png

Photo Credit: https://commons.wikimedia.org/wiki/File:Artificial_neural_network_imEDAD_recognition.png

Photo Credit: <https://www.pexels.com/photo/grey-and-white-short-fur-cat-104827/>

REPASO DEL ANÁLISIS DE COMPONENTES PRINCIPALES

- PCA es un algoritmo de aprendizaje automático no supervisado.
- PCA realiza reducciones de dimensionalidad mientras intenta mantener la información original sin cambios.
- PCA funciona tratando de encontrar un nuevo conjunto de características llamadas componentes.
- Los componentes son compuestos de las características de entrada o no correlacionadas entre si.

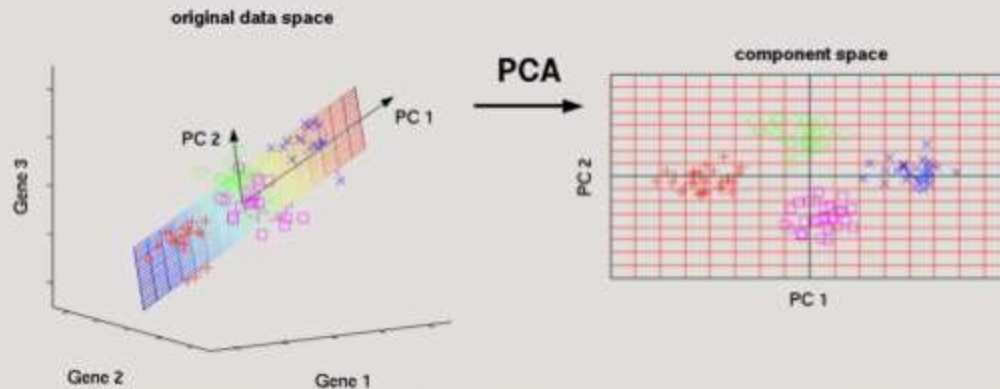


Photo Credit: <http://phdthesis-bioinformatics-maxplanckinstitute-molecularplantphys.matthias-scholz.de/>

