

# Aprendizaje automático no supervisado

Marta Caro Martínez

Adaptado de Enrique Martín y Javier Arroyo

# Principales técnicas

Reglas de asociación

Clustering

# Reglas de asociación

- En datos no etiquetados no existe un atributo especial (clase), por lo que el objetivo de las reglas es encontrar relaciones entre distintos atributos.
- En principio nos interesa cualquier regla que muestre relación entre atributos.
- Las reglas tienen un aspecto más complejo que las de clasificación: pueden tener varias condiciones tanto en el lado izquierdo como en el derecho.
- En general podemos generar muchas reglas, pero estaremos interesados únicamente en aquellas con una calidad mínima (**confianza, soporte y completitud**)

# Reglas de asociación: confianza

- La confianza de una regla  $r = si\ l\ entonces\ r$  es la proporción de instancias correctamente predichas de todas las que encaja l:

$$confianza(r) = \frac{\#instancias\ cumplen\ l\ y\ r}{\#instancias\ cumplen\ l}$$

- Si hay muchas instancias que cumplen l pero no r, la confianza será baja (cercana a 0)
- Si casi todas las instancias que cumplen l cumplen también r, entonces la confianza será alta (cercana a 1)

# Reglas de asociación: soporte

- El soporte de una regla  $r = si \text{ / entonces } r$  es la proporción de instancias a las que se aplica correctamente la regla:

$$soporte(r) = \frac{\#instancias \text{ cumplen } l \text{ y } r}{\#instancias \text{ total}}$$

- Si la regla predice pocas instancias del conjunto total, el soporte será bajo (cerca a 0)
- Si muchas instancias del conjunto total son predichas por la regla su soporte será alto (cercano a 1)

# Reglas de asociación: completitud

- La completitud de una regla  $r = si \text{ / entonces } r$  es la proporción de instancias correctamente predichas de todas las que encaja  $r$ :

$$completitud(r) = \frac{\#instancias \text{ cumplen } l \text{ y } r}{\#instancias \text{ cumplen } r}$$

- Si pocas instancias que cumplen  $r$  cumplen también  $l$ , la completitud será baja (cerca a 0)
- Si casi todas las instancias que cumplen  $r$  cumplen también  $l$ , la completitud serán muy alta (cercana a 1)

# Ejemplo de reglas de asociación

- Análisis de datos de una lista de la compra:

Transacción 1: Pan, Leche, Huevos

Transacción 2: Pan, Leche

Transacción 3: Leche, Pañales, Cerveza

Transacción 4: Pan, Huevos, Pañales

Transacción 5: Leche, Pan, Huevos, Pañales

- Algoritmo aplicado: Apriori

Ejemplo de reglas obtenidas

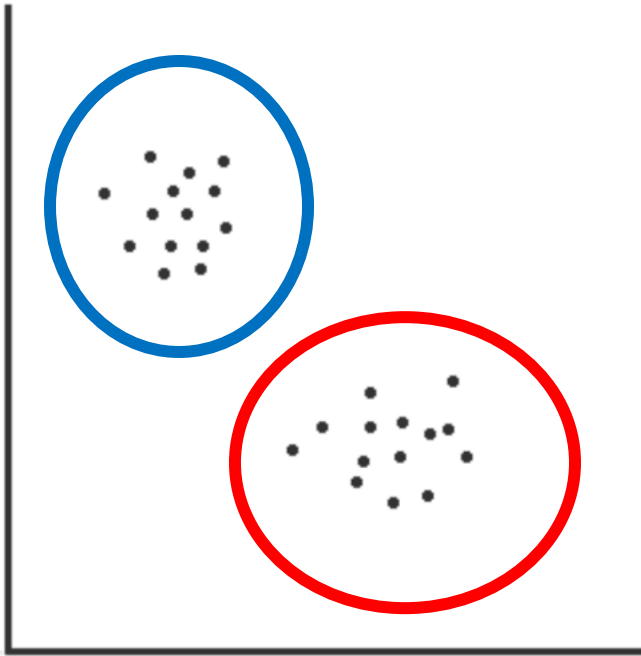
$\{\text{Pan, Leche}\} \Rightarrow \{\text{Huevos}\}$

Si un cliente compra pan y leche, es probable que también compre huevos.

$\{\text{Leche, Pañales}\} \Rightarrow \{\text{Cerveza}\}$

Si un cliente compra leche y pañales, es probable que también compre cerveza.

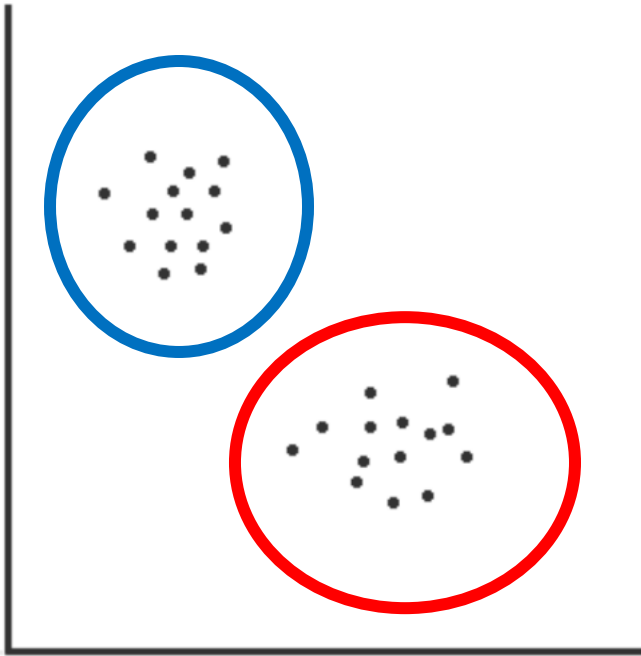
# Clustering



- Considerando unos datos no etiquetados, la tarea de clustering (agrupado) trata de encontrar conjuntos de instancias similares.
- En este caso tenemos 2 conjuntos claramente separados



# Clustering



- Objetivo:
  - Encontrar grupos con instancias similares
  - Diferenciar los grupos al máximo
- Podremos estudiar los grupos:
  - Cuáles son los que tienen más instancias y menos,
  - Los que son más homogéneos y más dispersos, o incluso tienen outliers

# Tipos de clustering

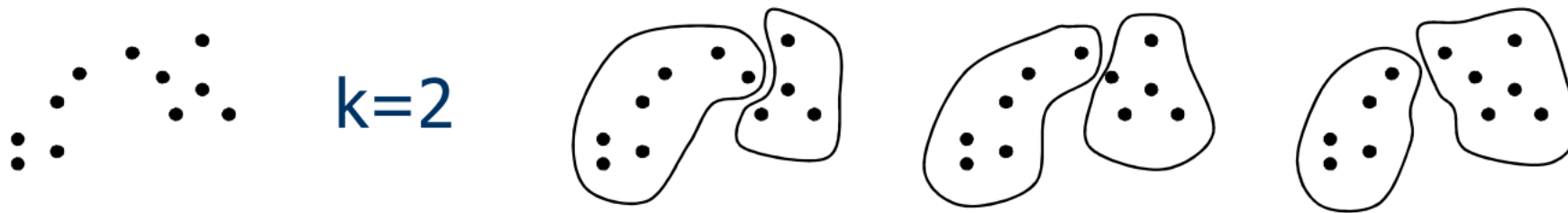
Algoritmos de **clustering basados en particiones**

Algoritmos de **clustering jerárquico**

# Clustering basado en particiones

## Algoritmos de **clustering basados en particiones**

- Los conjuntos están ya creados y las instancias van cambiando de un cluster a otro
- K (número de clusters se fija desde el principio)
- Ejemplo: K-Means



# K-Means

## **Inicialización:**

- Tenemos  $k$  fijado
- Creamos centroides aleatoriamente

## **Proceso iterativo:**

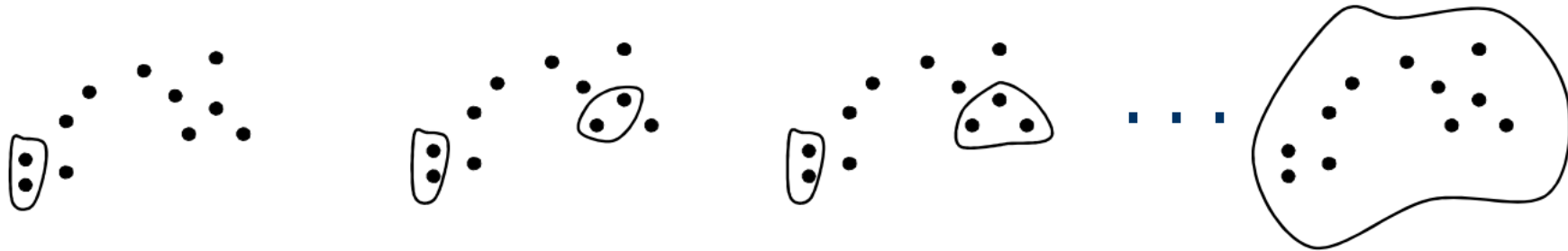
Mientras que los centroides cambien:

- Calcular las distancias de todos los puntos a los  $k$  centroides.
- Formar  $k$  grupos, asignando cada punto al centroide más cercano.
- Recalcular los nuevos centroides (puntos centrales de los  $k$  grupos).

# Clustering jerárquico

## Algoritmos de **clustering jerárquico**

- Iterativamente va construyendo los conjuntos, añadiendo instancias



# Clustering jerárquico aglomerativo

- **Estrategia general:**

1. Cada individuo empieza siendo un cluster, es decir, hay  $n$  clusters, tantos como individuos
2. En cada iteración fusiona los dos clusters que estén más cercanos
3. Termina cuando todos se tiene todos los grupos aglomerados en un solo cluster.

# Clustering jerárquico aglomerativo

chỗ C cộng cộng để đặt

, A giống chỗ để C ở sự phân biệt

, Các lần là phân biệt nên chỗ phân

Élégis loại độ phân sự chia phân biệt và  
nhiều lần

- Các lần là phân biệt nên nên phân biệt và nên  
sự

- Rẻ rẻ loại rẻ và nên phân biệt và đặt lần  
phân biệt nên nên phân biệt

# Ejemplo de CJA

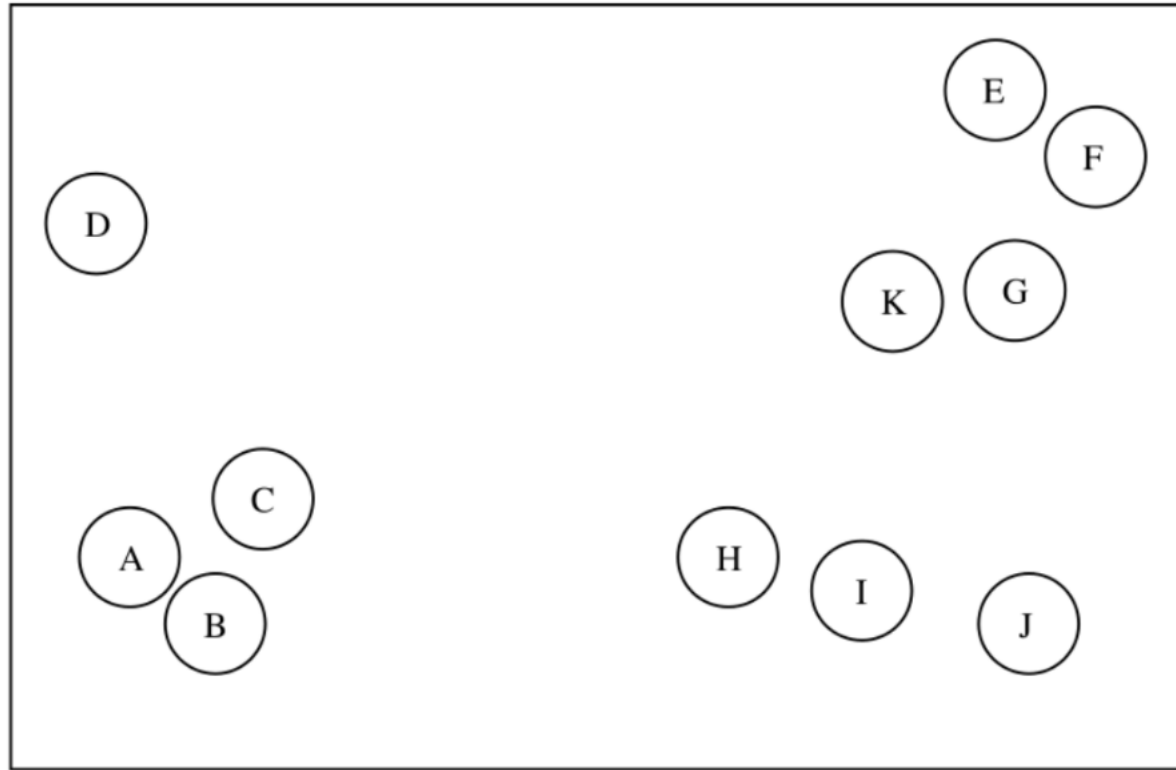


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.



# Ejemplo de CJA

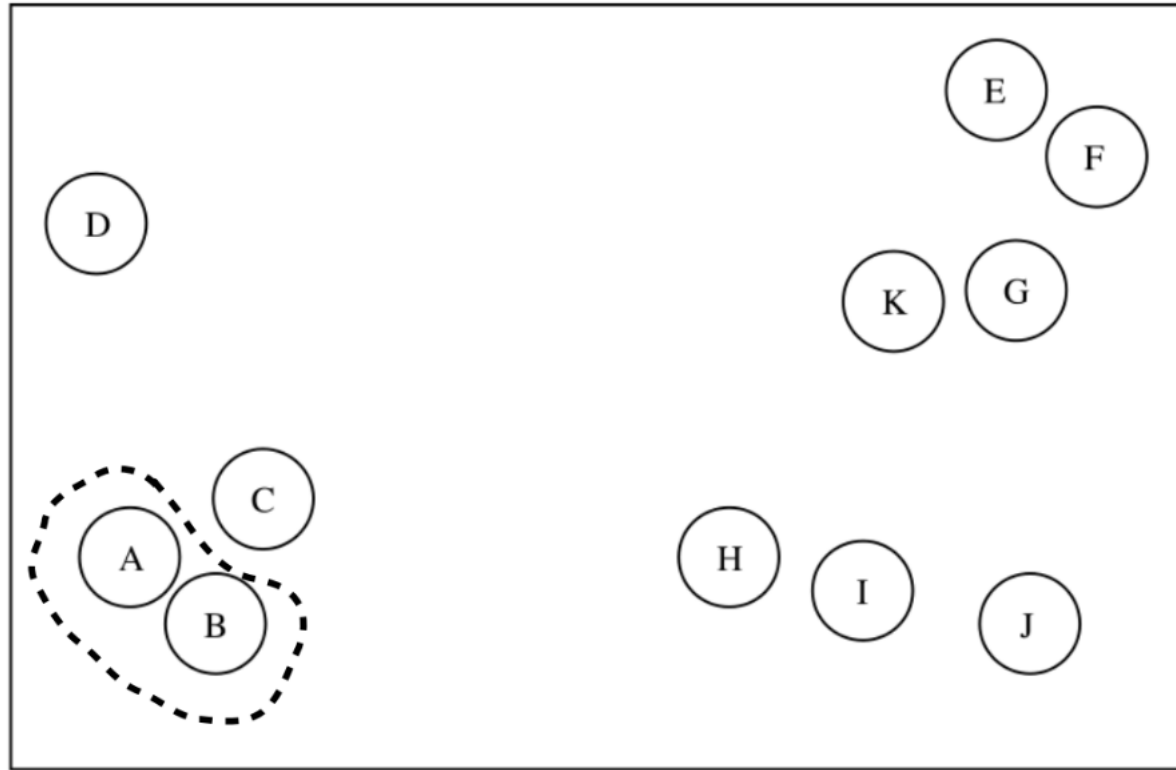


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

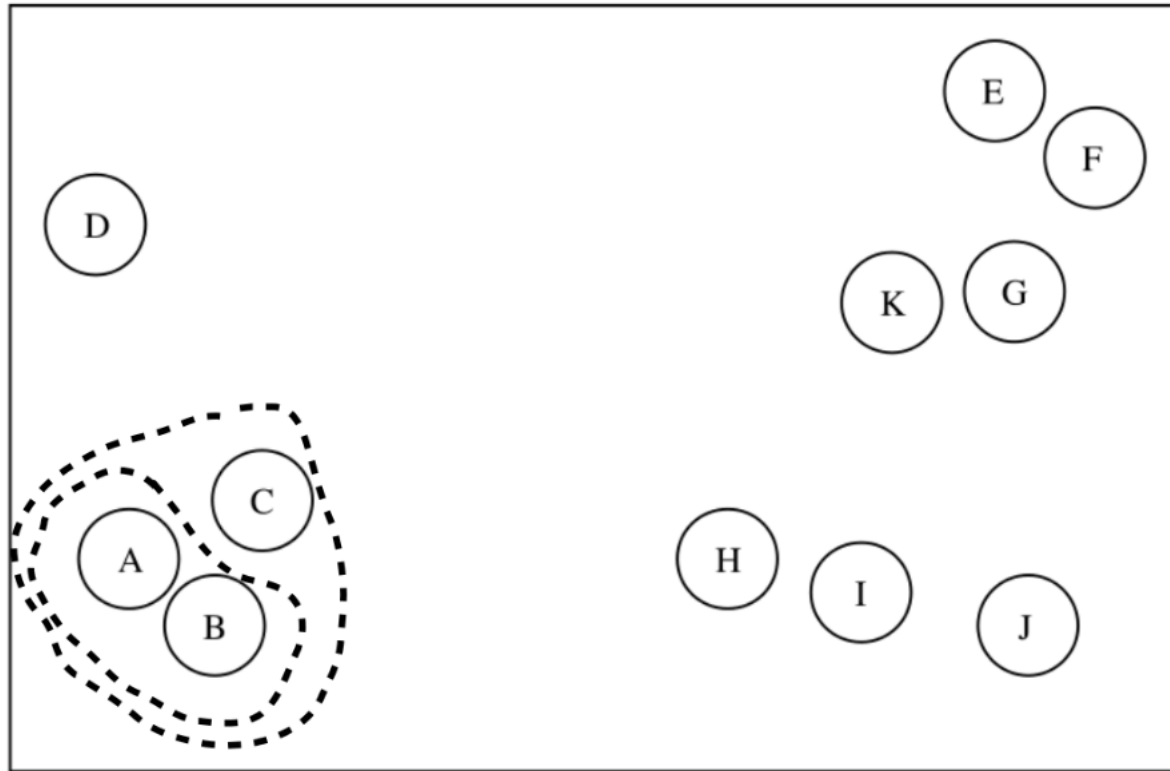


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

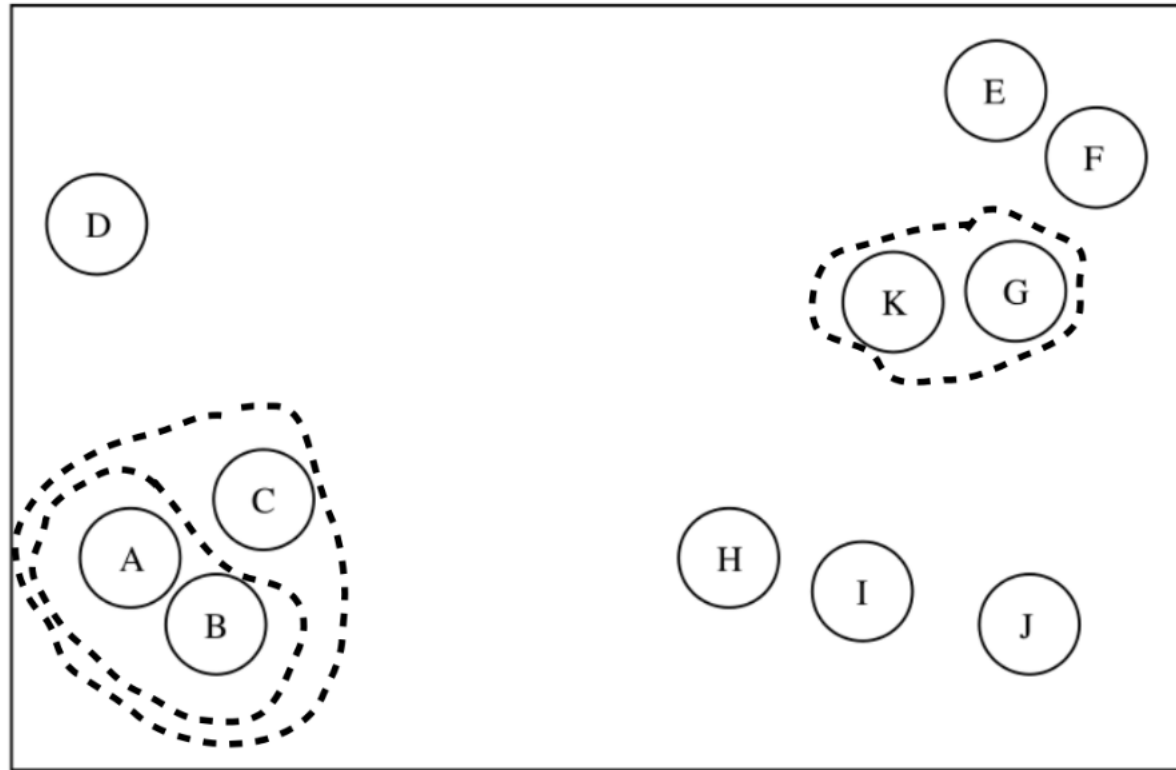


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

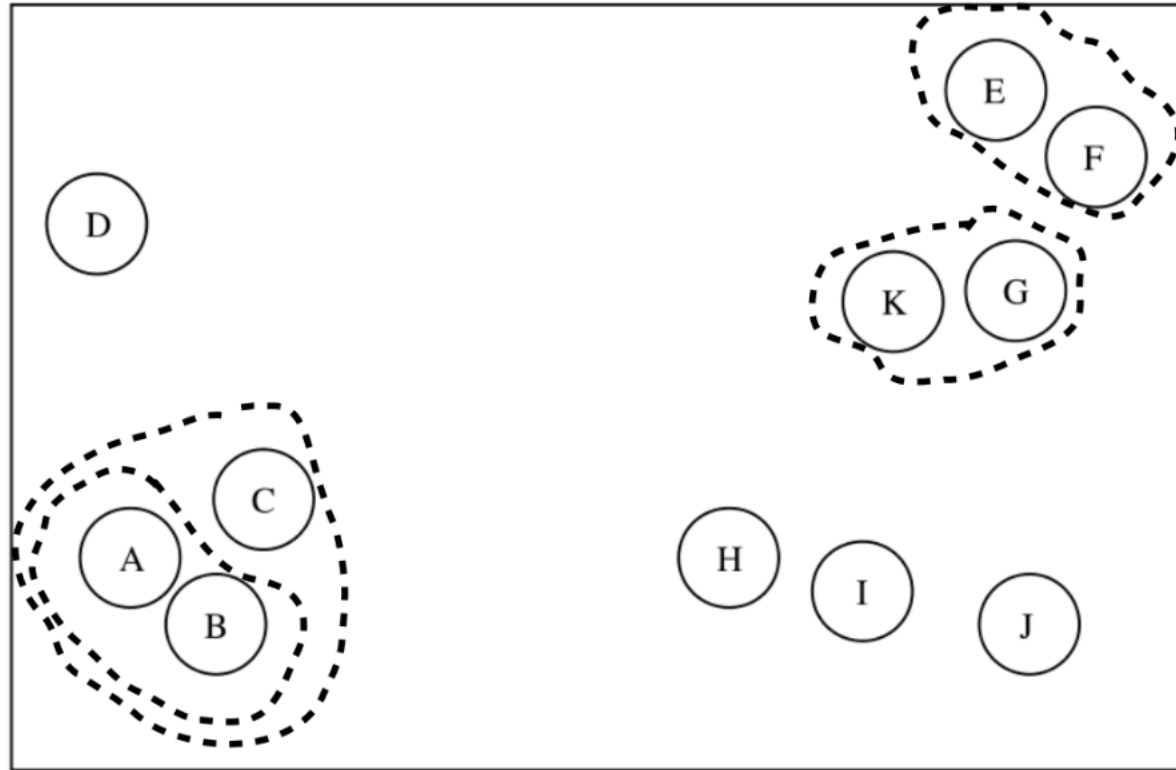


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

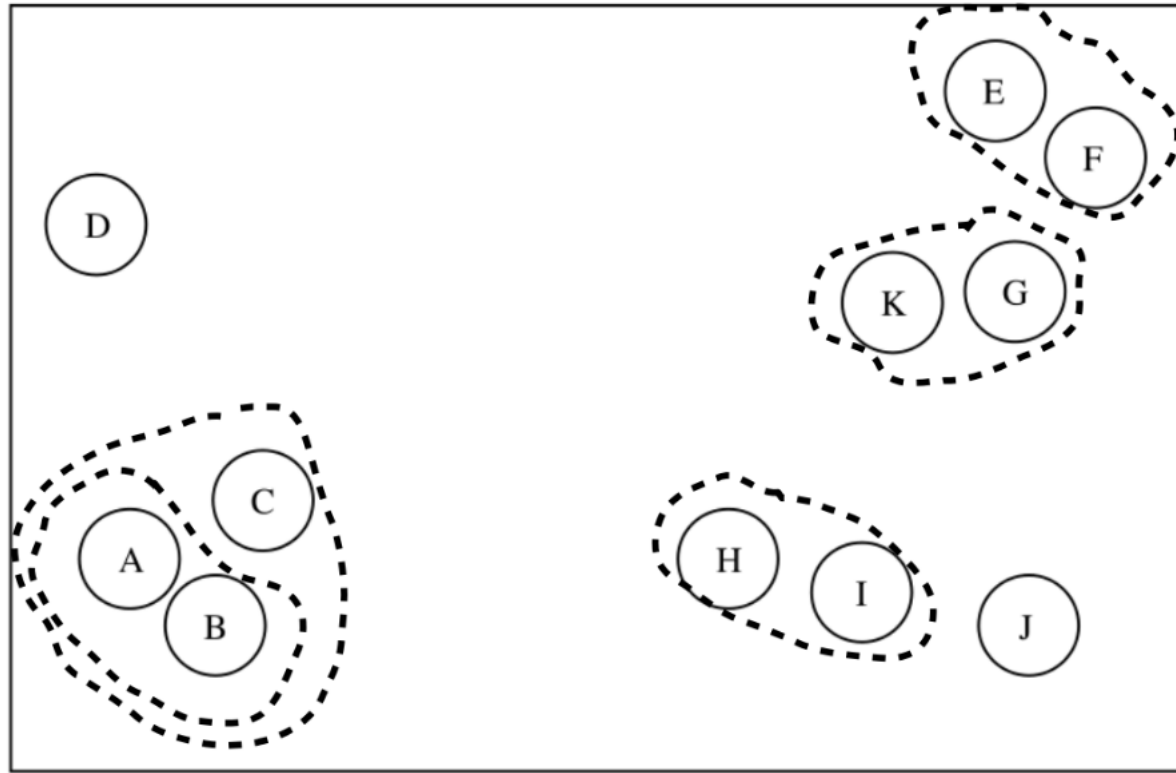


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

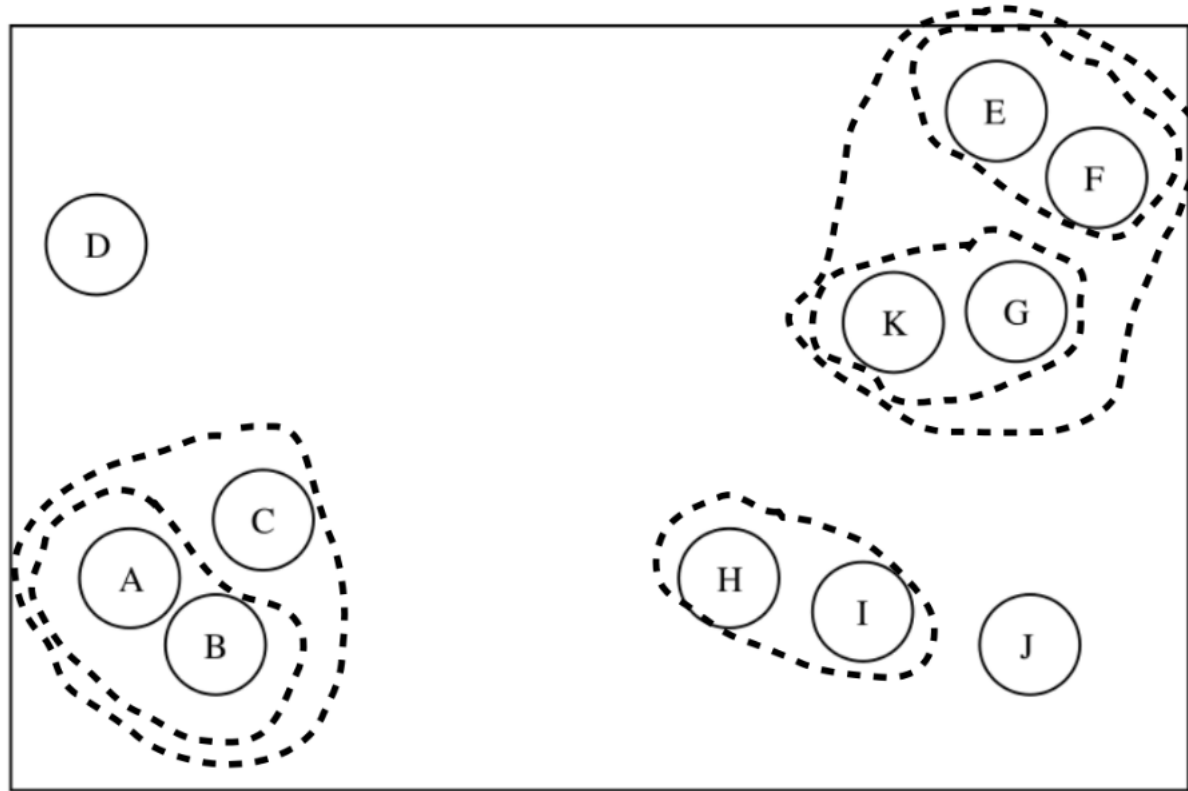


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

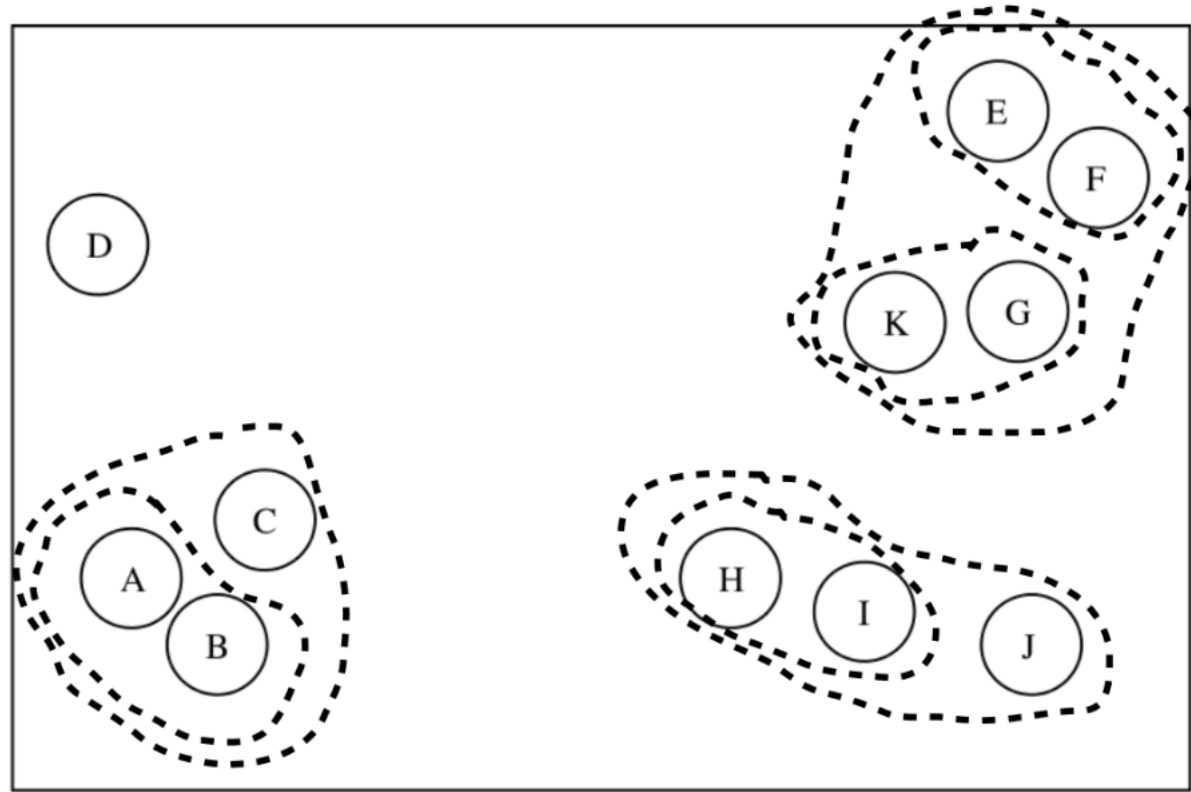


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

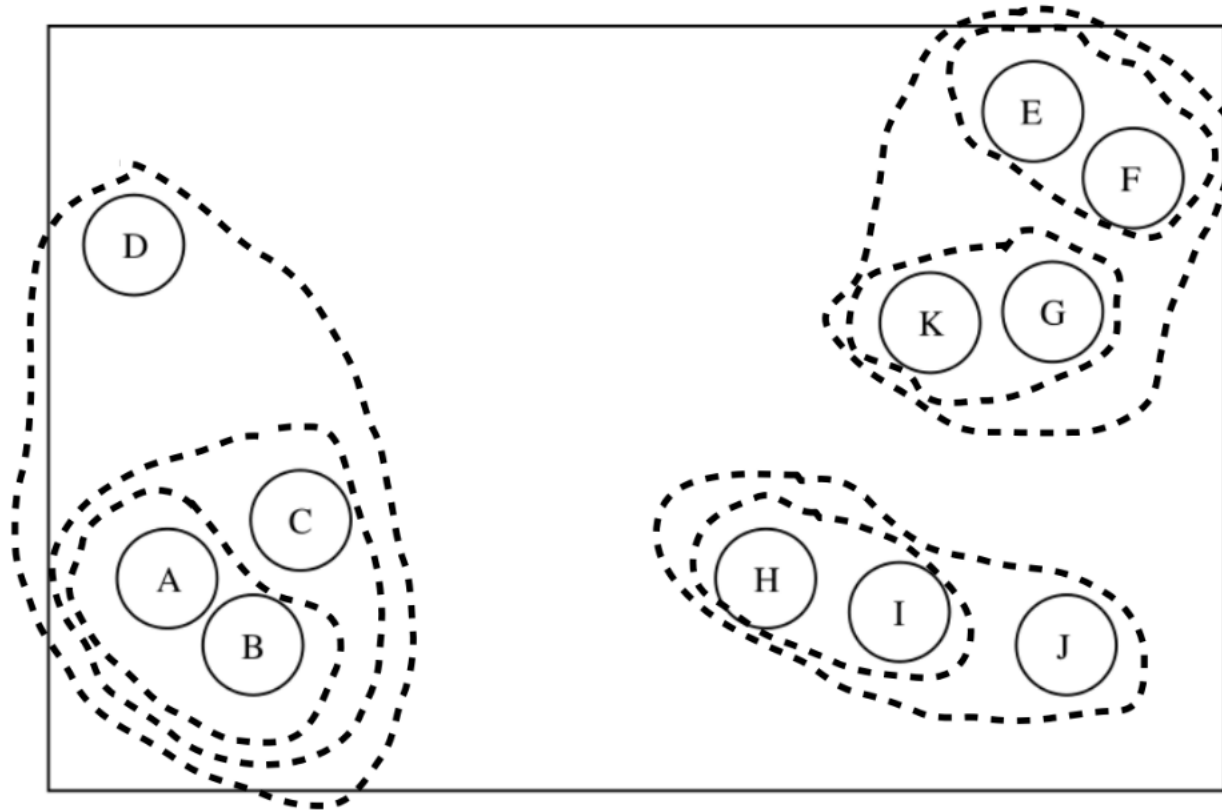


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.



# Ejemplo de CJA

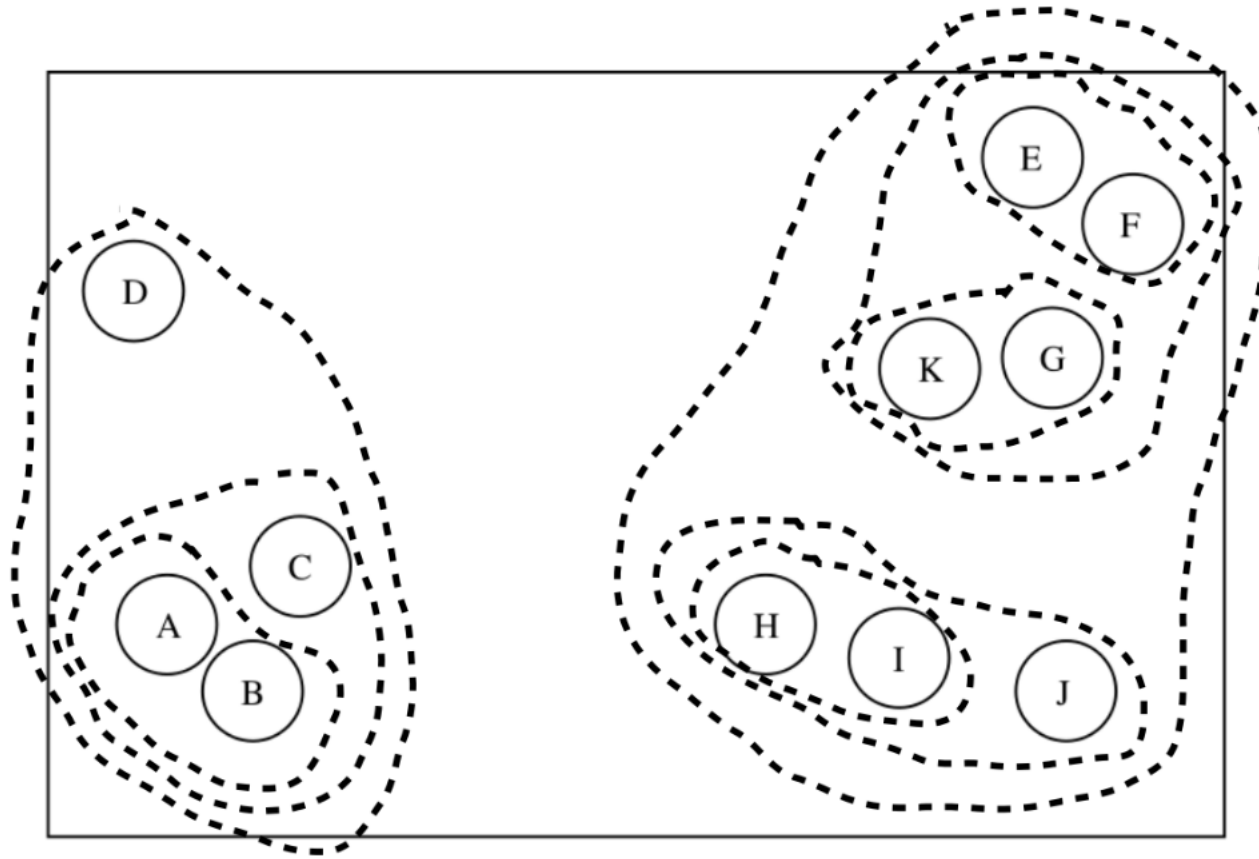


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA

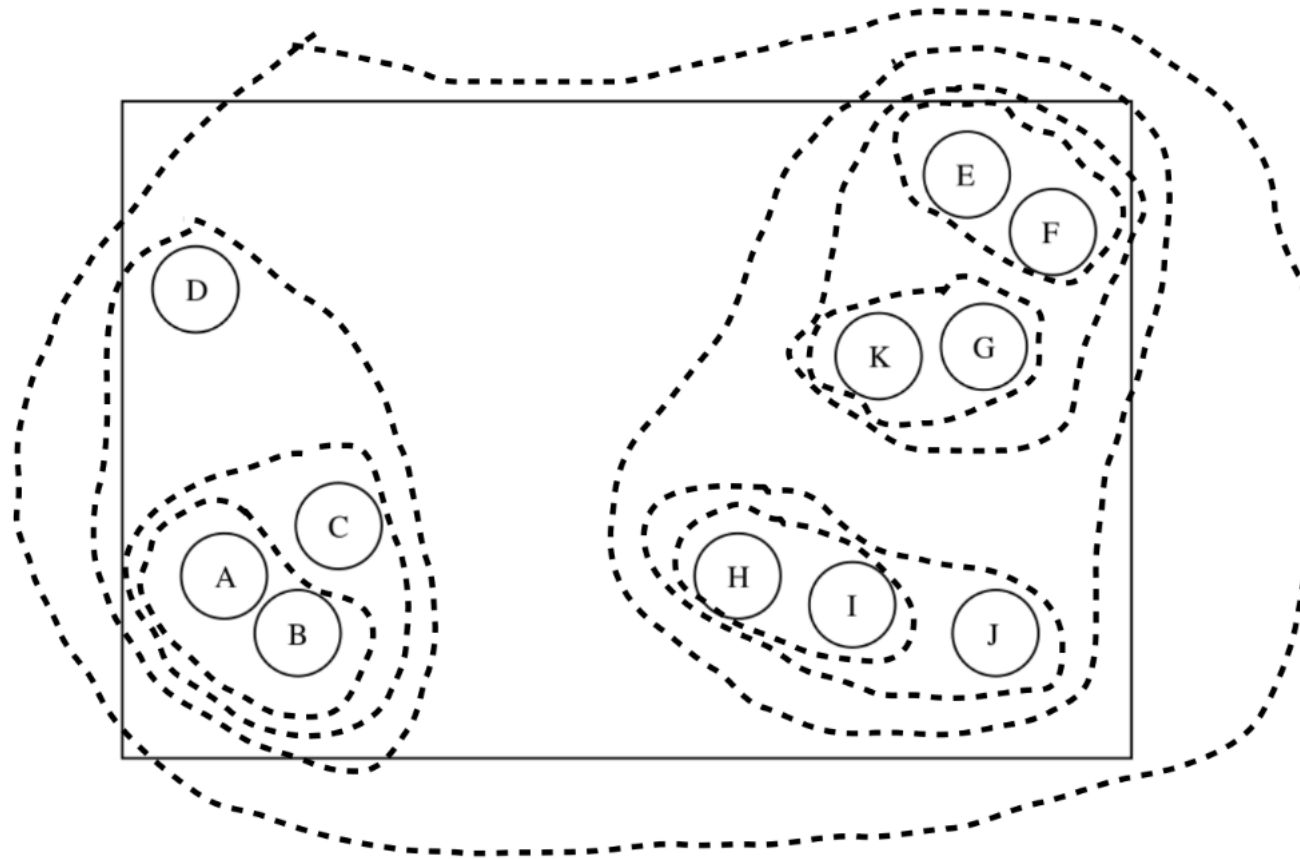


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# Ejemplo de CJA - dendograma

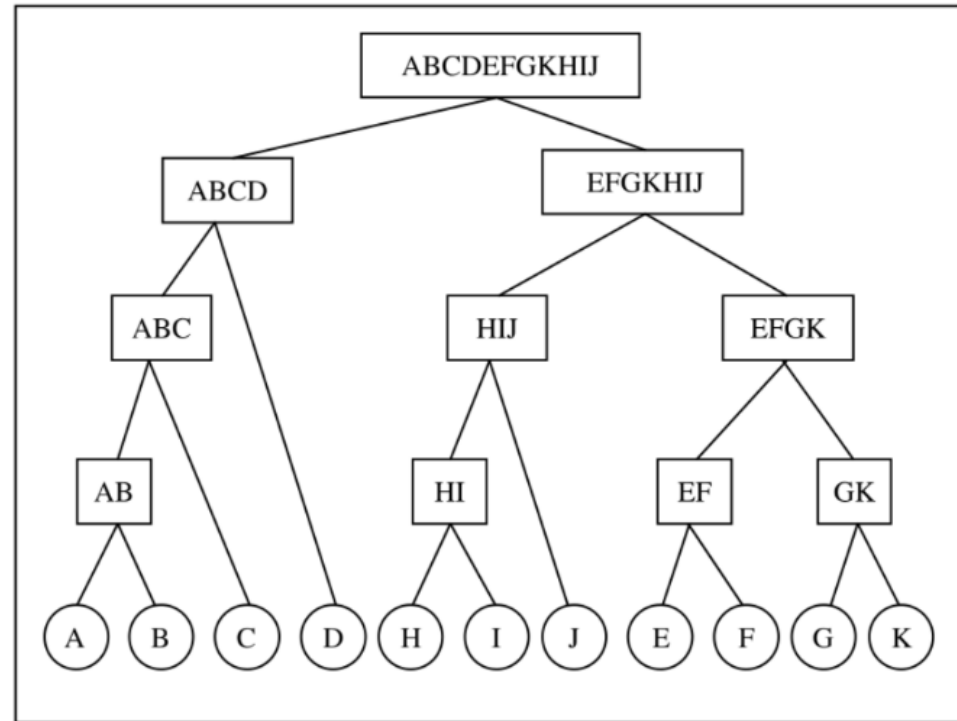


Figura obtenida de Principles of Data Mining, second edition. Max Bramer. Springer 2013.

# CJA resultado: dendograma

- **Dendograma:** representación visual de cómo se ha realizado el proceso.
- Nos ayuda a decidir cuál es el **número óptimo de clusters** que necesitamos
  - Buscar la altura o las ramas en las cuales los clusters se fusionan
  - Buscar patrones dentro de los clusters para ver si nos cuadran, buscar sus características distintivas
  - Visualizar los clusters
  - **Evaluar los clusters:** con nuevas instancias determinar en qué cluster caerían y si tiene sentido.

# Métricas de evaluación en clustering

- **Índice de silueta:** mide como de compactos son los clusters y su separación

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

*a*: Distancia promedio entre un punto y los puntos de su propio clúster.

*b*: Distancia promedio entre un punto y los puntos del clúster más cercano al que no pertenece.

*S*: valores entre -1 y 1.

**Valores cercanos a 1 → división correcta; cercanos a -1 → división incorrecta**

# Métricas de evaluación en clustering

- **Índice de Davies-Bouldin:** mide como de compactos son los clusters y su separación

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \frac{\sigma_i + \sigma_j}{d(c_i, c_j)}$$

$\sigma_i$ : dispersión dentro del cluster i

$d(c_i, c_j)$ : distancia entre los centroides de los clusters i y j

**Cuanto menor sea el valor de DB, mejor separación de los clusters**