

Aprendizaje no supervisado

VC05: Agrupamiento basado en densidad - Propagación de Afinidad

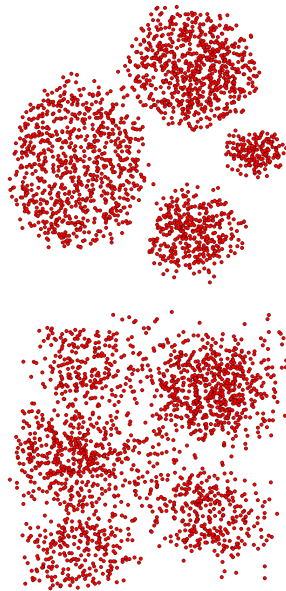
Félix José Fuentes Hurtado

felixjose.fuentes@campusviu.es

Universidad Internacional de Valencia

Tipos de algoritmos de agrupamiento

- ▶ Basados en particiones
- ▶ Jerárquicos
- ▶ Espectrales
- ▶ **Basados en densidad**
- ▶ Probabilísticos

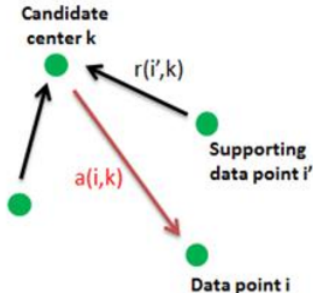
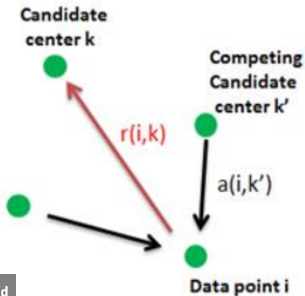


Agrupamiento basado en densidad

Idea

Pase de mensajes:

- Responsabilidad (R): evidencia de que x_k es el centroide más apropiado para x_i
- Disponibilidad (A): evidencia de lo apropiado que x_i escoja x_k como su centroide habiendo sido elegido como centroide por otros puntos



Inicialización

- ▶ Matriz de similitud (S): Similitud entre todos los pares de ejemplos del conjunto de datos
- ▶ Preferencias: creencia (autoconfianza) de que un ejemplo x_k puede ser centroide

Se modifica la **matriz de similitud** S para incorporar en su diagonal principal los valores de **preferencia** de cada ejemplo.

Responsabilidad

$$r(\mathbf{x}_i, \mathbf{x}_k) = s(\mathbf{x}_i, \mathbf{x}_k) - \max_{k' \neq k} (a(\mathbf{x}_i, \mathbf{x}_{k'}) + s(\mathbf{x}_i, \mathbf{x}_{k'}))$$

Disponibilidad

$$a(\mathbf{x}_i, \mathbf{x}_k) = \begin{cases} \min \left(0; r(\mathbf{x}_k, \mathbf{x}_k) + \sum_{i' \in \{1, \dots, n\}: i' \neq i \wedge i' \neq k} \max \left(0; r(\mathbf{x}_{i'}, \mathbf{x}_k) \right) \right), & i \neq k \\ \sum_{i' \in \{1, \dots, n\}: i' \neq k} \max \left(0; r(\mathbf{x}_{i'}, \mathbf{x}_k) \right), & i = k \end{cases}$$

Agrupamiento basado en densidad

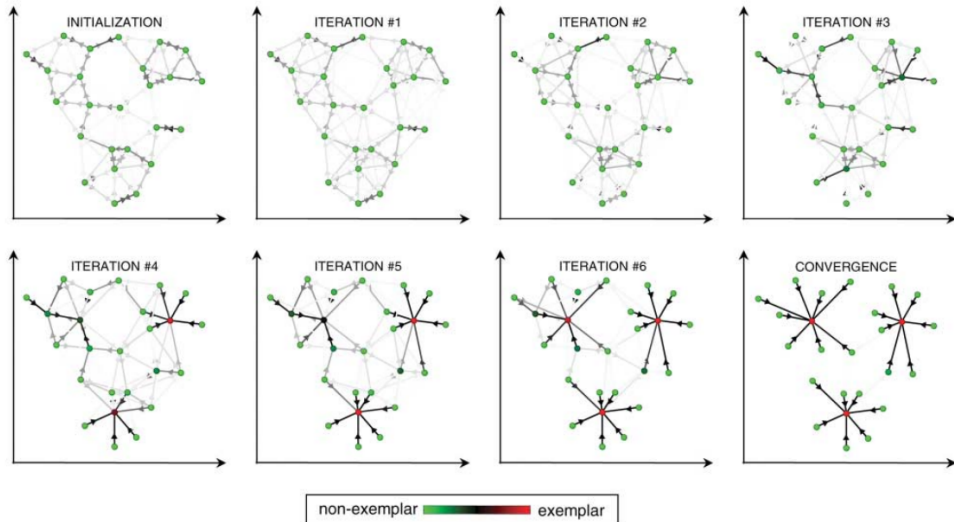
Affinity propagation

Algoritmo

1. Inicializar S con las preferencias
2. Inicializar A todo a cero
3. Repetir hasta convergencia:
 - 3.1 Actualizar R
 - 3.2 Actualizar S

Agrupamiento basado en densidad

Affinity propagation



Agrupamiento basado en densidad

Affinity propagation

Representantes o centroides

Aquellos puntos que tienen una responsabilidad y disponibilidad hacia ellos mismos positiva:

$$E \subset \{1, \dots, n\} : \forall k \in E, \left(r(\mathbf{x}_k, \mathbf{x}_k) + a(\mathbf{x}_k, \mathbf{x}_k) \right) > 0$$

Agrupamiento basado en densidad

Affinity propagation

Representantes o centroides

Aquellos puntos que tienen una responsabilidad y disponibilidad hacia ellos mismos positiva:

$$E \subset \{1, \dots, n\} : \forall k \in E, \left(r(\mathbf{x}_k, \mathbf{x}_k) + a(\mathbf{x}_k, \mathbf{x}_k) \right) > 0$$

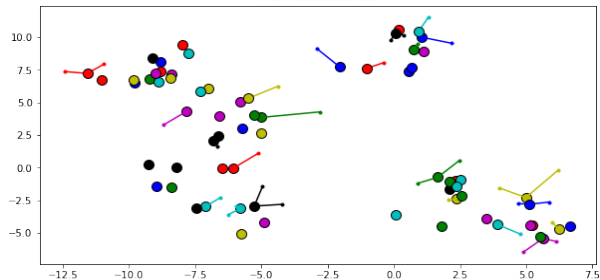
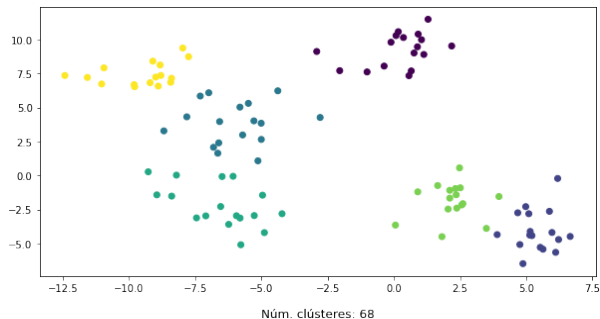
Asignación a clústeres (representantes)

Aquellos puntos que tienen una responsabilidad y disponibilidad hacia ellos mismos positiva:

$$C(\mathbf{x}_i) = \arg \max_k \left(r(\mathbf{x}_i, \mathbf{x}_k) + a(\mathbf{x}_i, \mathbf{x}_k) \right)$$

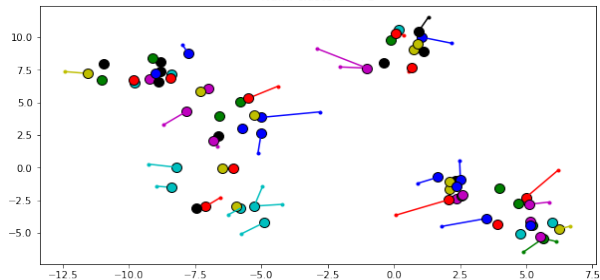
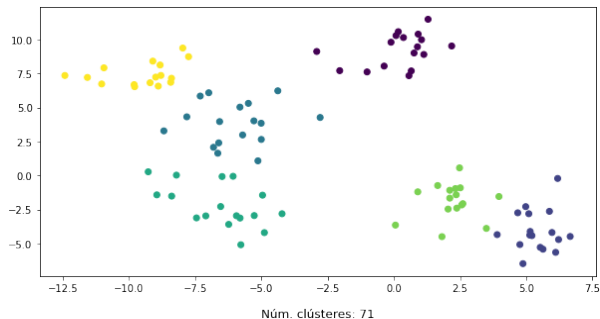
Agrupamiento basado en densidad

Mean-shift



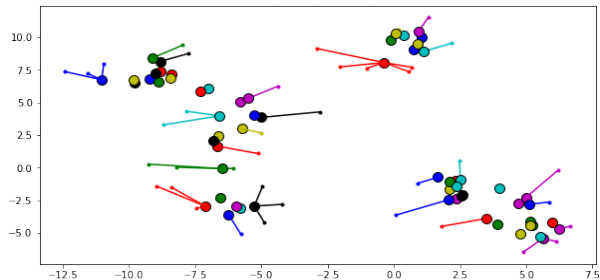
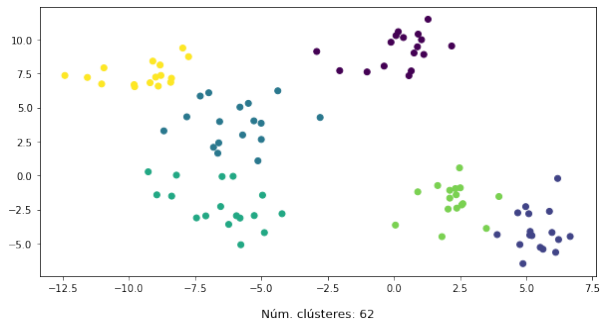
Agrupamiento basado en densidad

Mean-shift



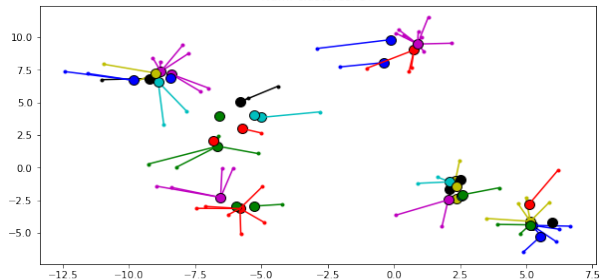
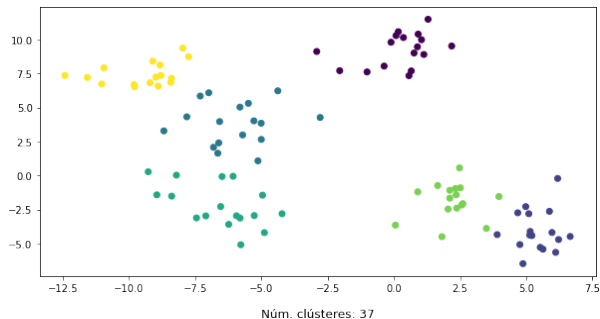
Agrupamiento basado en densidad

Mean-shift



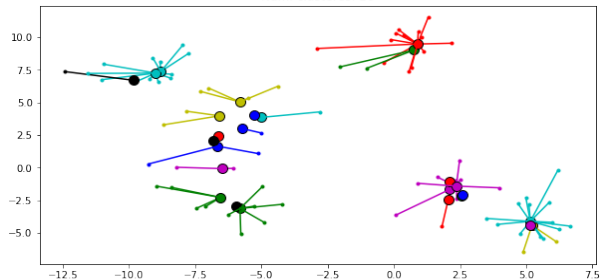
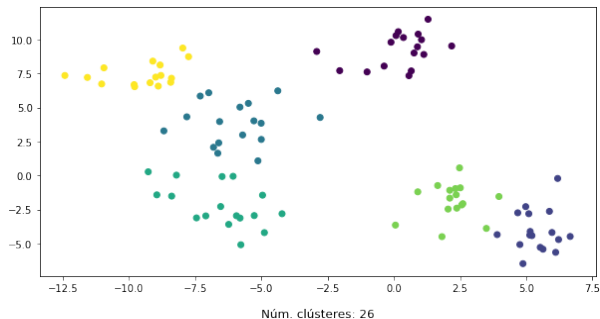
Agrupamiento basado en densidad

Mean-shift



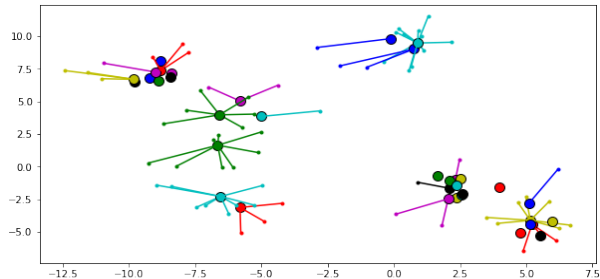
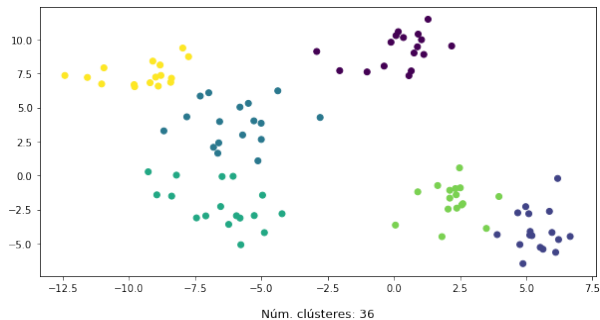
Agrupamiento basado en densidad

Mean-shift



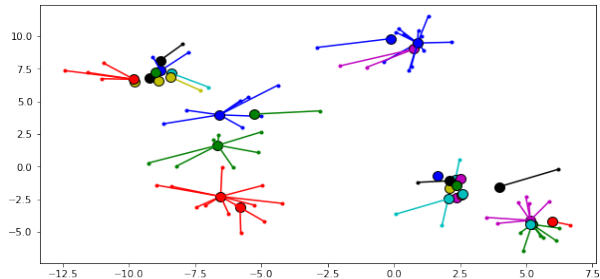
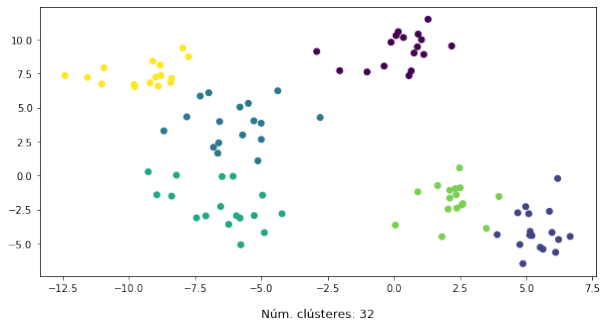
Agrupamiento basado en densidad

Mean-shift



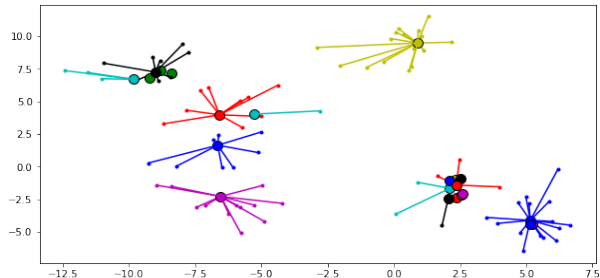
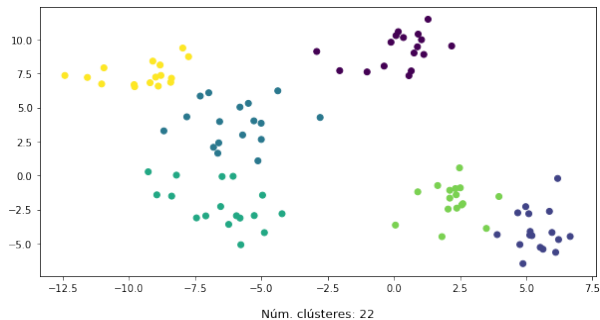
Agrupamiento basado en densidad

Mean-shift



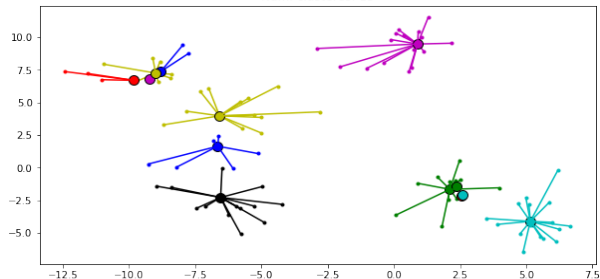
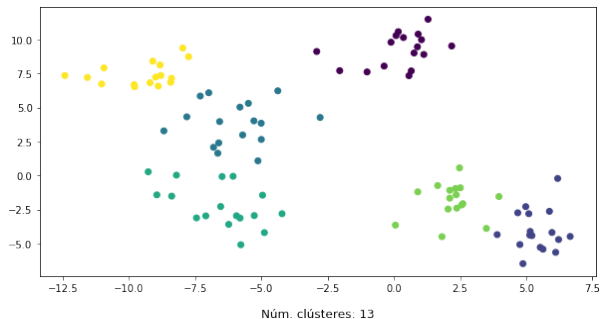
Agrupamiento basado en densidad

Mean-shift



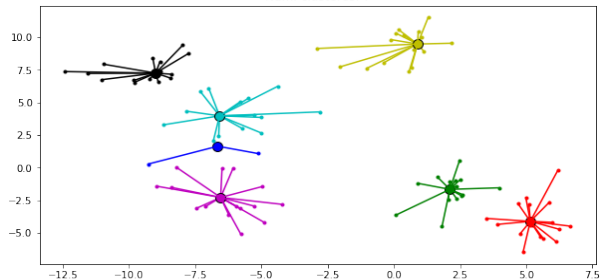
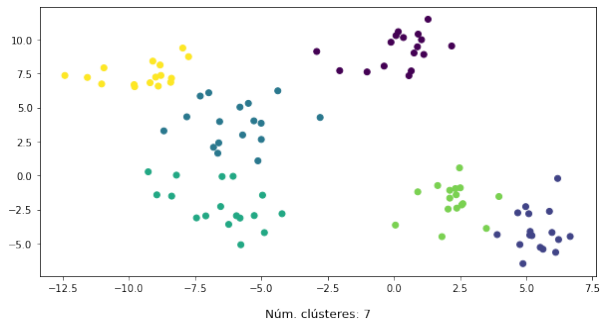
Agrupamiento basado en densidad

Mean-shift



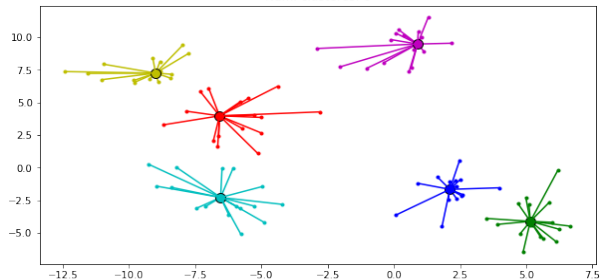
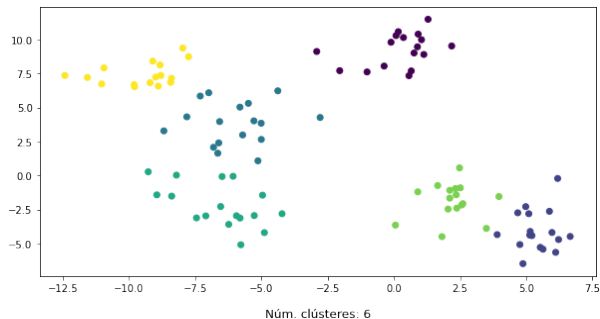
Agrupamiento basado en densidad

Mean-shift



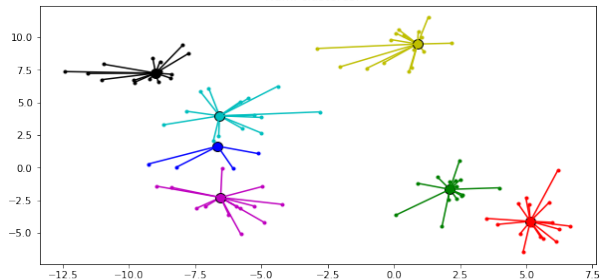
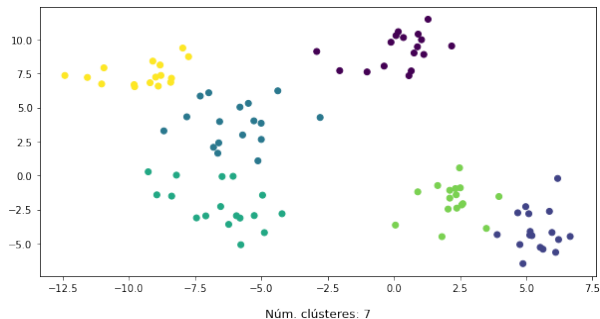
Agrupamiento basado en densidad

Mean-shift



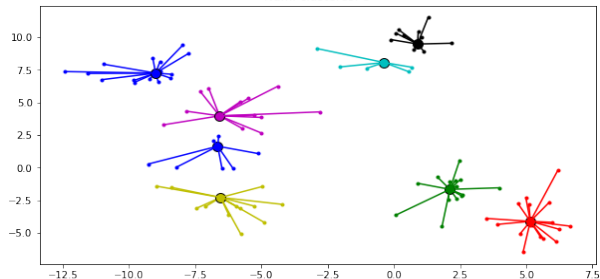
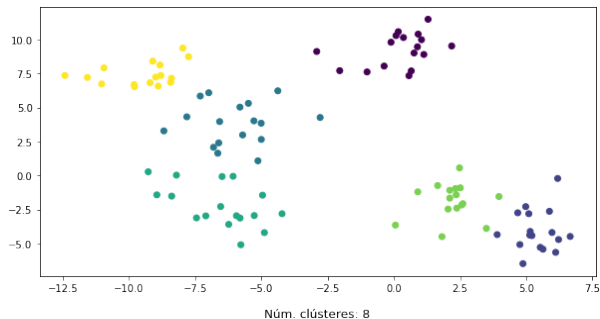
Agrupamiento basado en densidad

Mean-shift



Agrupamiento basado en densidad

Mean-shift



Agrupamiento basado en densidad

Affinity propagation

Ventajas

- ▶ No es necesario especificar K
- ▶ Definición basada en similitud
- ▶ Podría incorporar diferentes medidas de similitud
- ▶ Funciona con clústeres de diferente tamaño

Agrupamiento basado en densidad

Affinity propagation

Desventajas

- ▶ Conceptualmente difícil de explicar
- ▶ Problemas para lidiar con clústeres de formas diversas
- ▶ Selección de las preferencias

Aprendizaje no supervisado

VC05: Agrupamiento basado en densidad - Propagación de Afinidad

Félix José Fuentes Hurtado

felixjose.fuentes@campusviu.es

Universidad Internacional de Valencia