

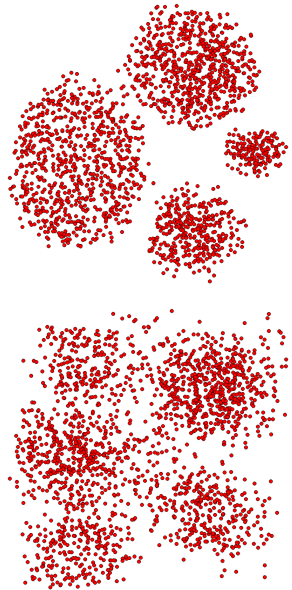
Aprendizaje no supervisado

Agrupamiento basado en densidad – Mean shift

Javier Sevilla

Tipos de algoritmos de agrupamiento

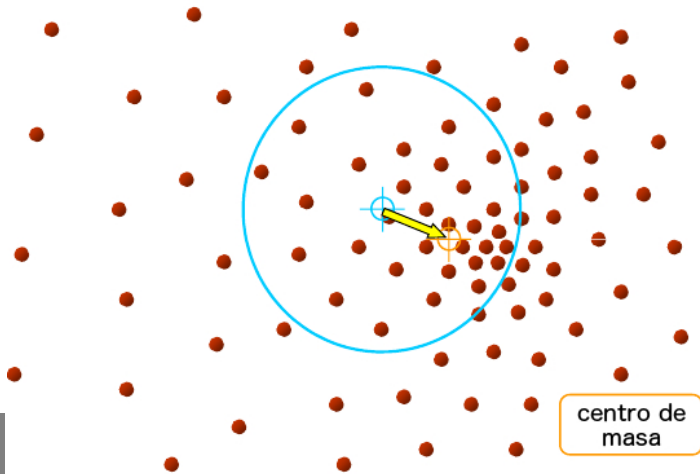
- ▶ Basados en particiones
- ▶ Jerárquicos
- ▶ Espectrales
- ▶ **Basados en densidad**
- ▶ Probabilísticos



Agrupamiento basado en densidad

Idea

Si desplazamos cada punto al centro de masa de su vecindario, los puntos se acaban agrupando de manera natural en grupos

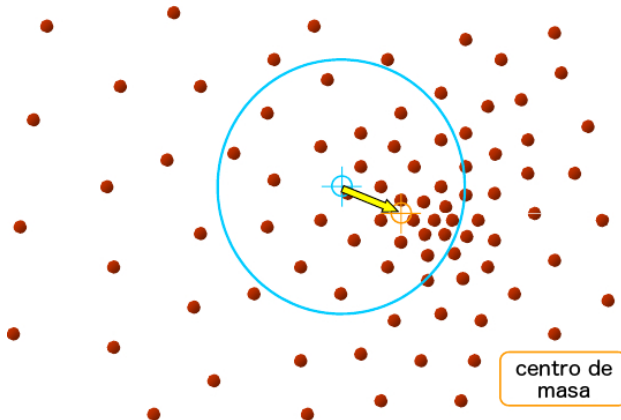


centro de
masa

Agrupamiento basado en densidad

Conceptos

- ▶ **Vecindario**, ¿qué casos se usan para calcular el centro?
- ▶ **Kernel**, ¿cómo se ponderan los casos usados para calcular el centro?



Agrupamiento basado en densidad

Media ponderada:

$$m(\mathbf{x}) = \frac{\sum_{i=1}^n k\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h^2}\right) \cdot \mathbf{x}_i}{\sum_{i=1}^n k\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h^2}\right)}$$

Agrupamiento basado en densidad

Media ponderada:

$$m(\mathbf{x}) = \frac{\sum_{i=1}^n k\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h^2}\right) \cdot \mathbf{x}_i}{\sum_{i=1}^n k\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h^2}\right)}$$

Kernel plano:

$$k(x) = \begin{cases} 1, & x \leq \lambda \\ 0, & x > \lambda \end{cases}$$

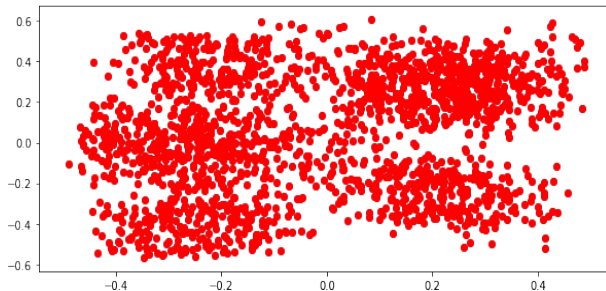
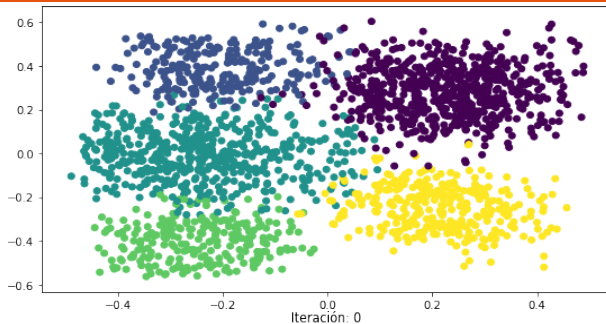
normalmente, $\lambda = 1$.

Kernel Gaussiano:

$$k(x) = e^{-x}$$

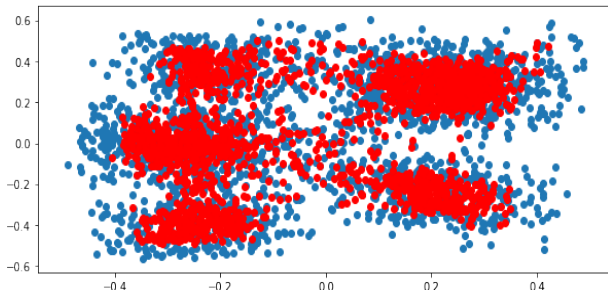
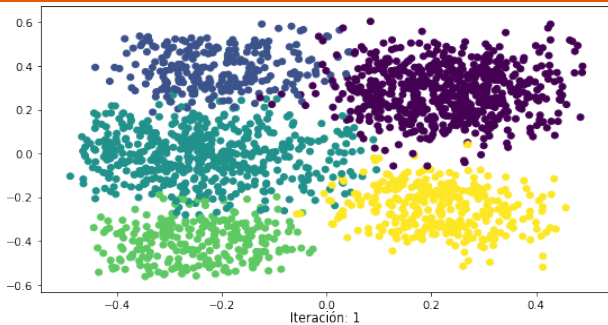
Agrupamiento basado en densidad

Mean-shift



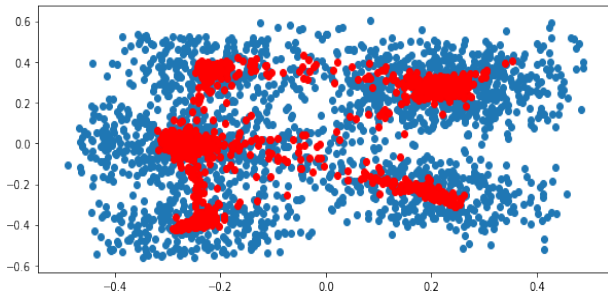
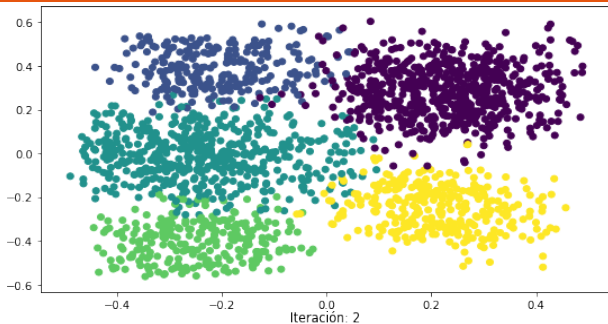
Agrupamiento basado en densidad

Mean-shift



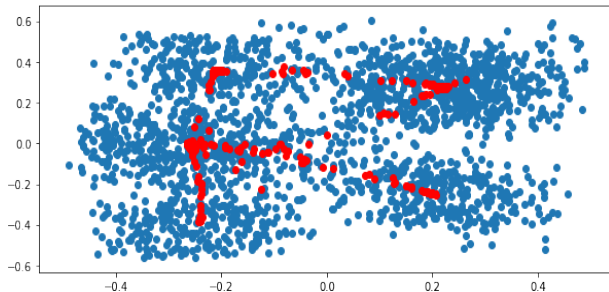
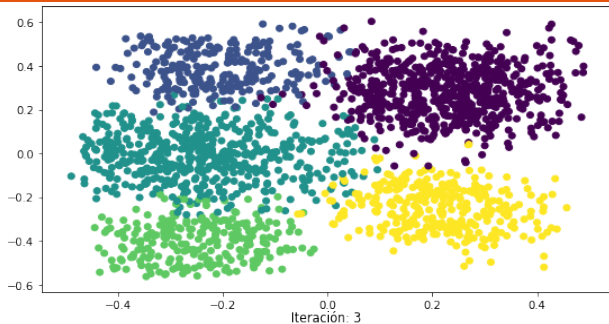
Agrupamiento basado en densidad

Mean-shift



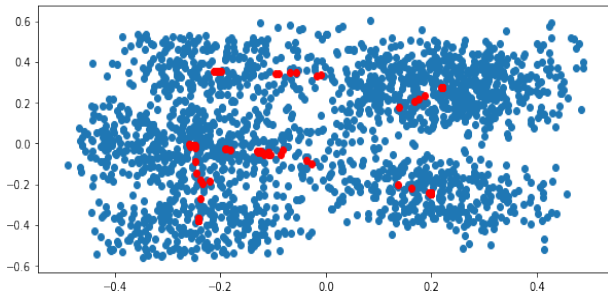
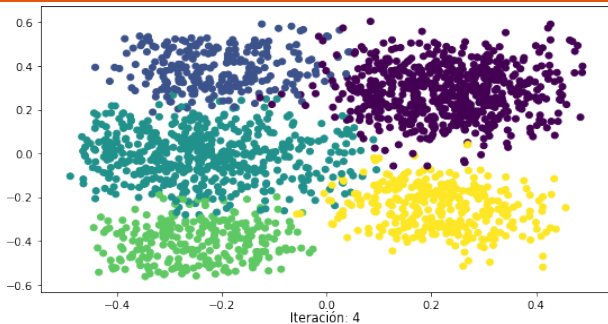
Agrupamiento basado en densidad

Mean-shift



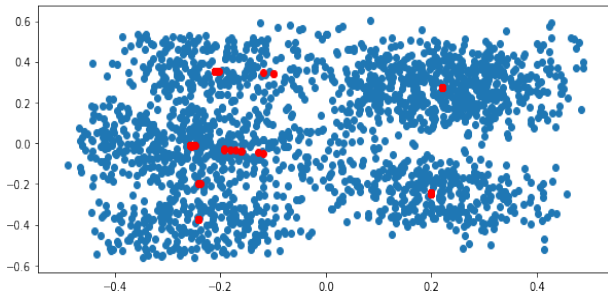
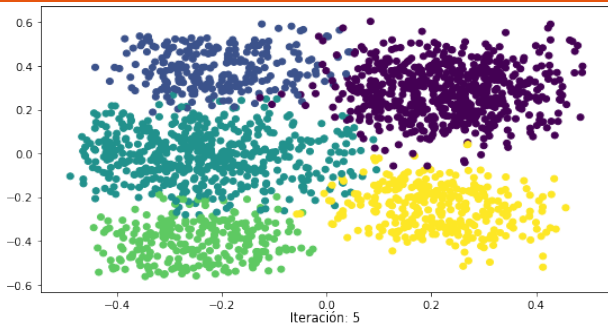
Agrupamiento basado en densidad

Mean-shift



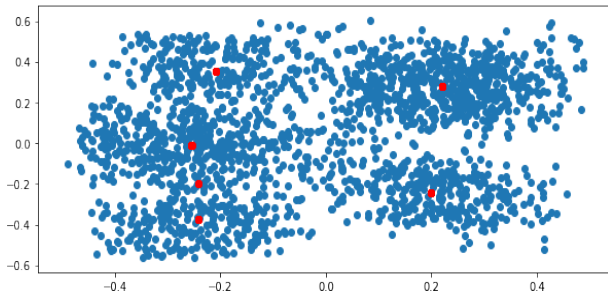
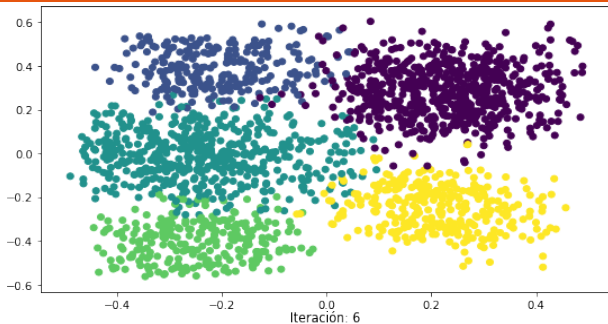
Agrupamiento basado en densidad

Mean-shift



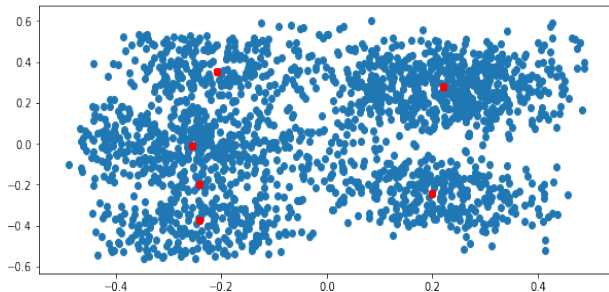
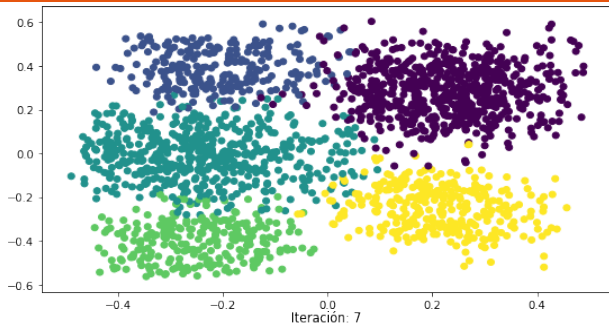
Agrupamiento basado en densidad

Mean-shift



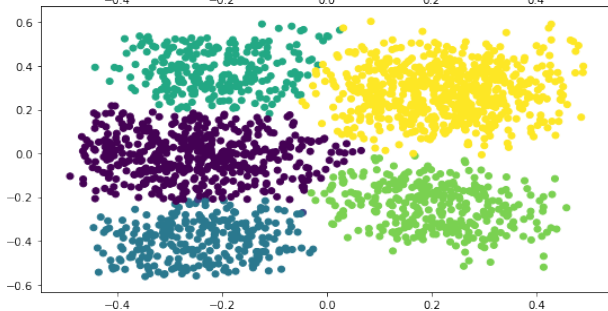
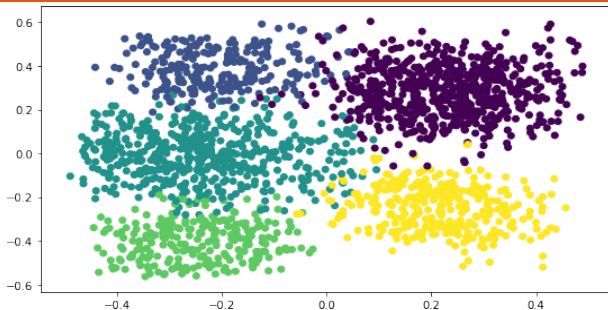
Agrupamiento basado en densidad

Mean-shift



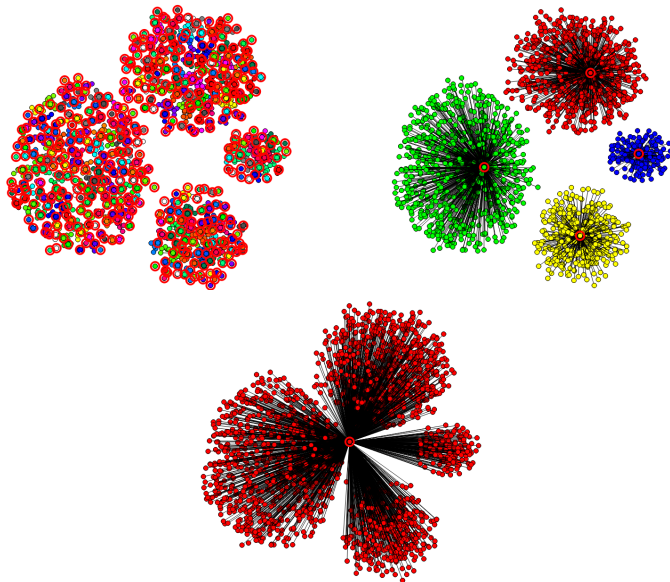
Agrupamiento basado en densidad

Mean-shift



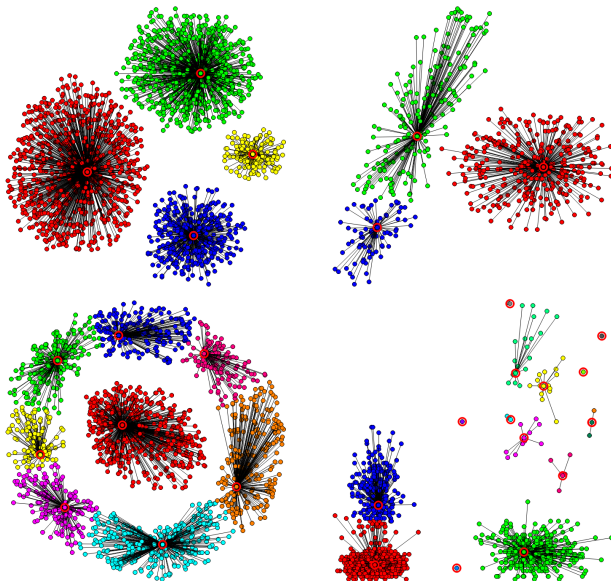
Agrupamiento basado en densidad

Mean-shift: efecto de h



Agrupamiento basado en densidad

Mean-shift: efecto de h



Agrupamiento basado en densidad

Mean-shift



Agrupamiento basado en densidad

Mean-shift

Ventajas

- ▶ Conceptualmente sencilla
- ▶ No es necesario especificar K
- ▶ Definición basada en densidad
- ▶ Funciona con clústeres de diferente tamaño y formas
- ▶ Diferentes kernels

Desventajas

- ▶ Clústeres no máximos
- ▶ Problemas al lidiar con clústeres de diferente densidad
- ▶ Sin demostración de convergencia en entornos reales

Gracias