Aprendizaje no supervisado

Inicialización avanzada de K-means

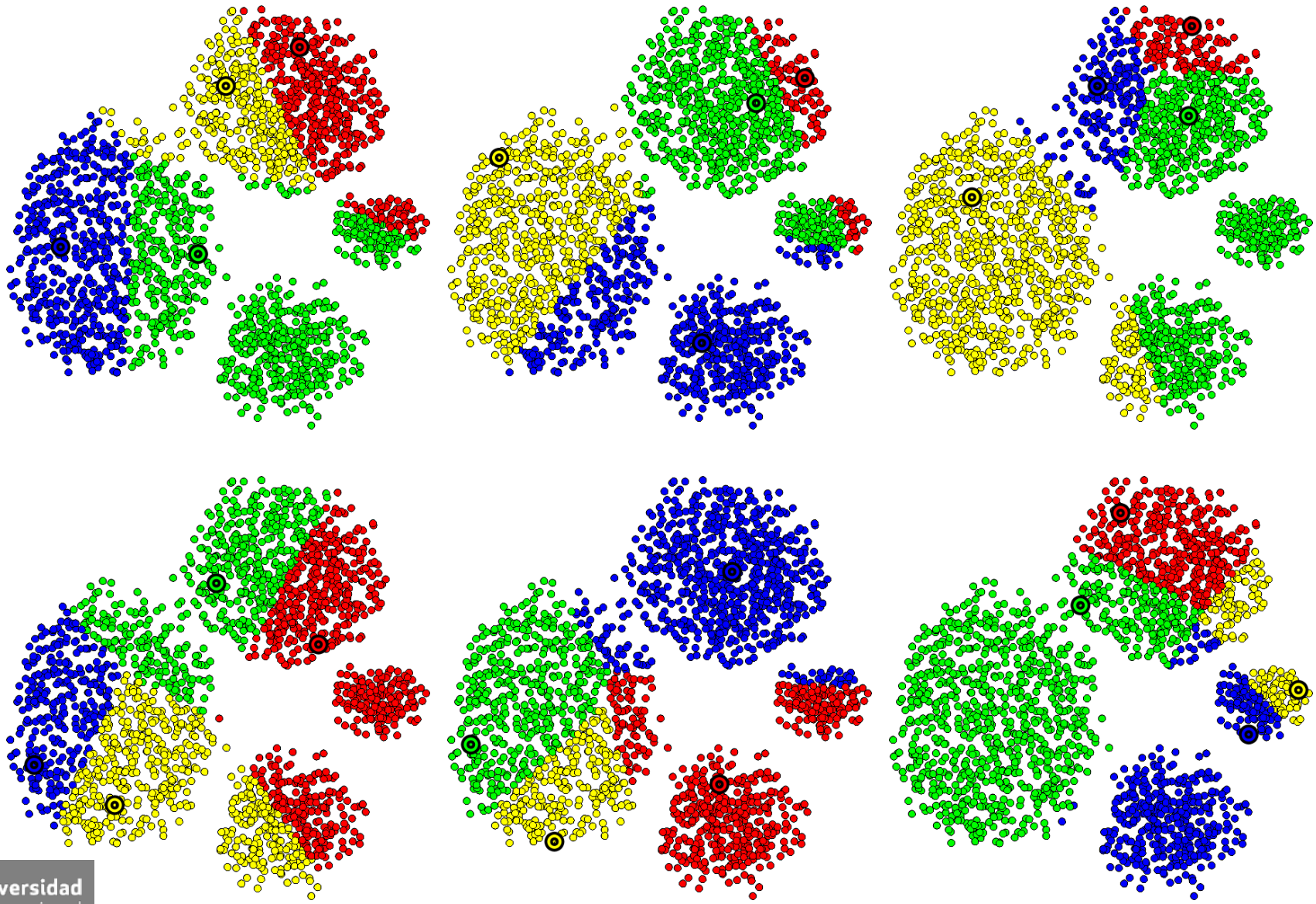
Javier Sevilla

viu Universidad Internacional de Valencia

25.02.2020

Agrupamiento basado en condiciones:

INICIALIZAR K-MEANS



Agrupamiento basado en condiciones:

INICIALIZAR K-MEANS

Método basado en repeticiones

- Inicializar aleatoriamente y ejecutar K-means R veces
- Medir la bondad de los R diferentes agrupamientos
- Devolver como resultado el mejor agrupamiento

K-means++

Inicialización avanzada del método K-means favoreciendo la separación de los centros iniciales

Intuición

Elección individual (y dependiente) de los centros

- **Centros**: Muestreo aleatorio no uniforme
- **Probabilidad**: Un ejemplo tiene mayor probabilidad de ser escogido como centro (inicial) cuanto mayor sea su distancia con los centros

Agrupamiento basado en condiciones:

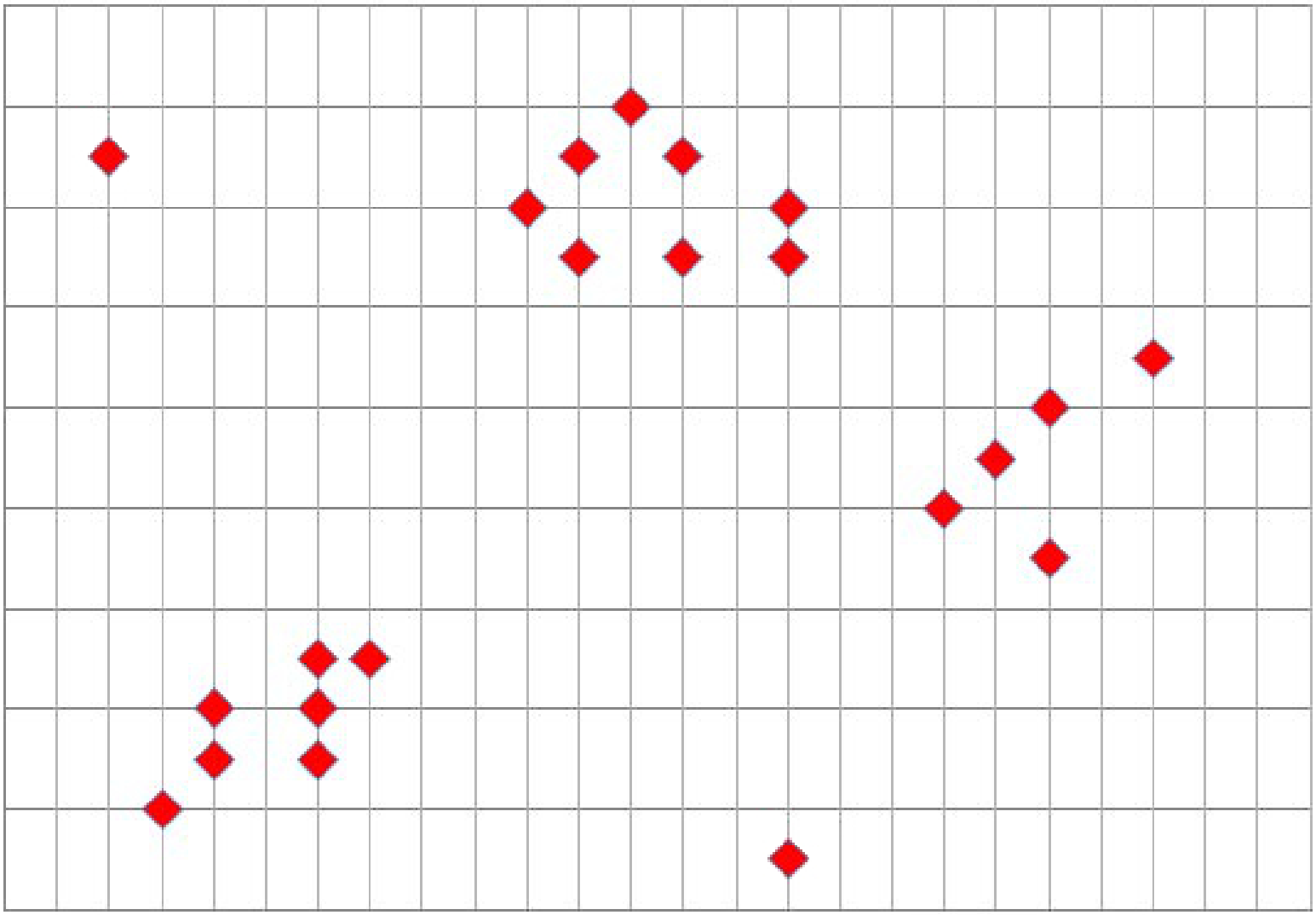
K-MEANS ++

La probabilidad de un ejemplo es proporcional al cuadrado de su distancia mínima a un centro ya incluido, $D(\mathbf{x}, S)$

$$D(\mathbf{x}, S) = \min_{k \in \{1,\ldots,|S|\}} ||\mathbf{x} - \bar{\mathbf{x}}_k||^2$$

Agrupamiento basado en condiciones:

K-MEANS ++



Agrupamiento basado en condiciones:

K-MEANS ++

K-means++ (inicialización)

Recibe: Conjunto de entrenamiento, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$; número de clústeres, K

1. Elección (aleatoria) de 1 punto del conjunto de entrenamiento como primer centro, $S = \{\bar{\boldsymbol{x}}_1\}$.

2. Mientras $|S| < K$, repetir

2.1. Para todos los ejemplos de entrenamiento, calcular $D(\boldsymbol{x}_i, S)$, la distancia al centro más cercano: $D(\boldsymbol{x}_i, S) = \min_{k \in \{1,\ldots,|S|\}} ||\boldsymbol{x}_i - \bar{\boldsymbol{x}}_k||^2$

2.2. Muestrear un nuevo caso $\boldsymbol{x}'$ del conjunto de entrenamiento, donde el caso $\boldsymbol{x}$ tiene probabilidad $D(\boldsymbol{x}, S)^2 / \sum_{i=1}^{n} D(\boldsymbol{x}_i, S)^2$ y añadir a S: $S = S \cup \{\boldsymbol{x}'\}$

Devuelve: Conjunto de centros, $\{\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_K\}$

Gracias

viu Universidad Internacional de Valencia