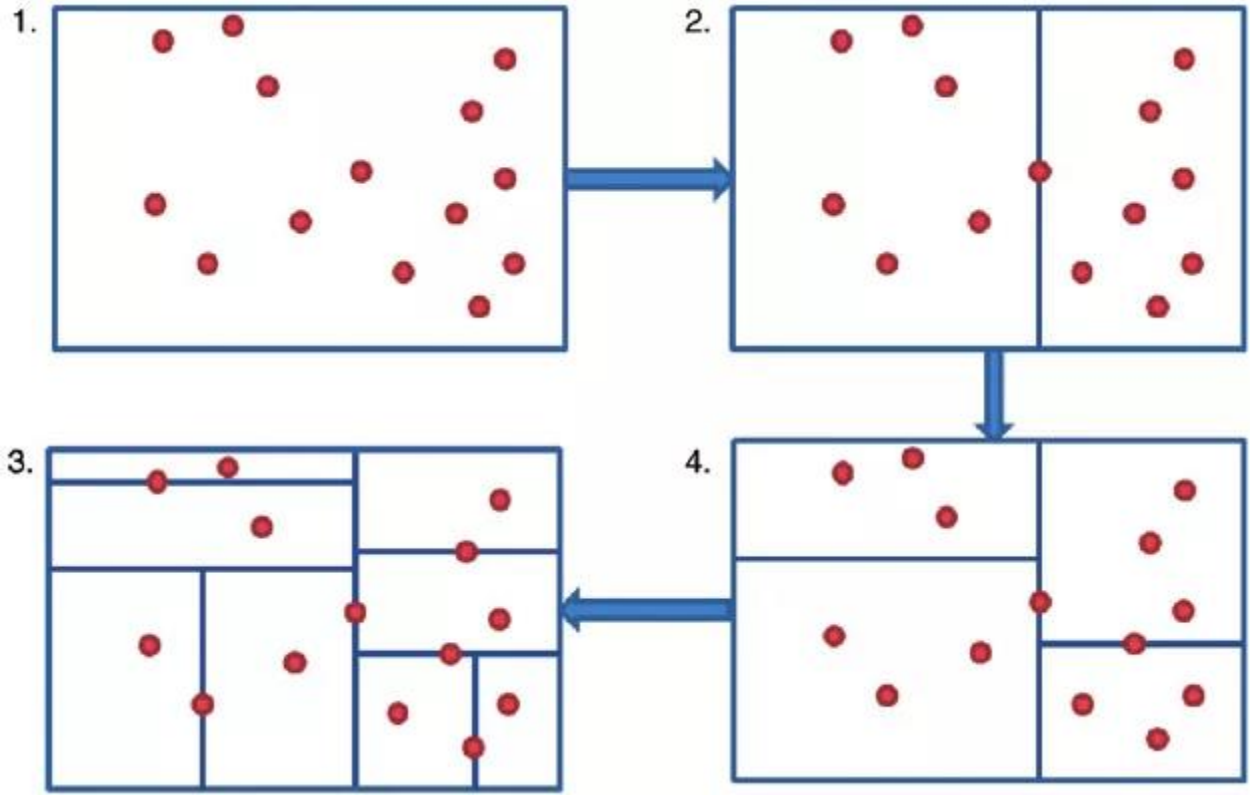
ESTRUCTURAS DE DATOS Y ALGORITMOS IIC2133

ANTONIO LOPEZ – MANUEL BECKER

Ayudantía T1 KDTree

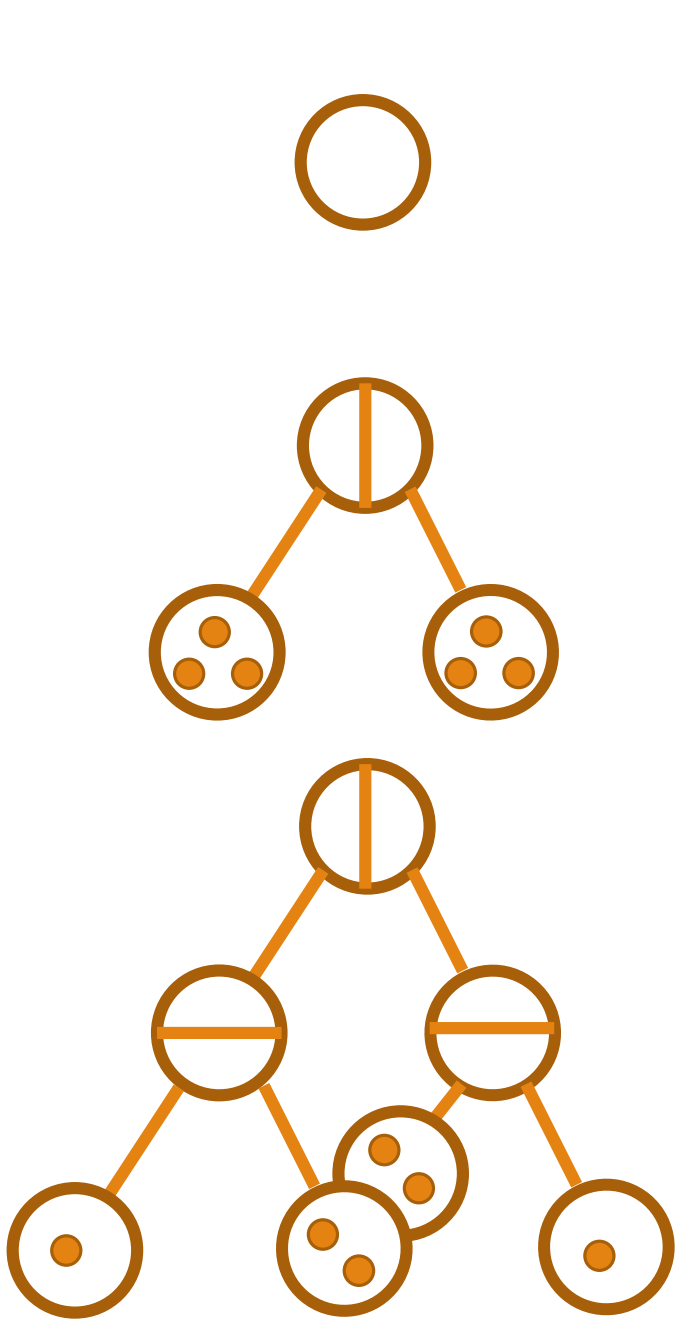
Enunciado T1

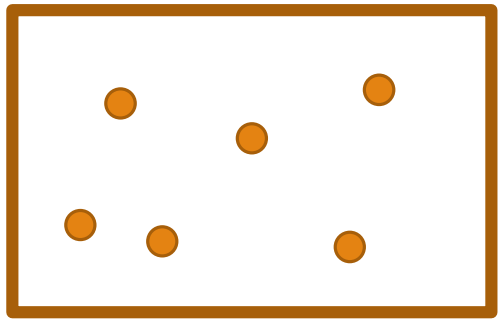
¿Qué es un KDTree?

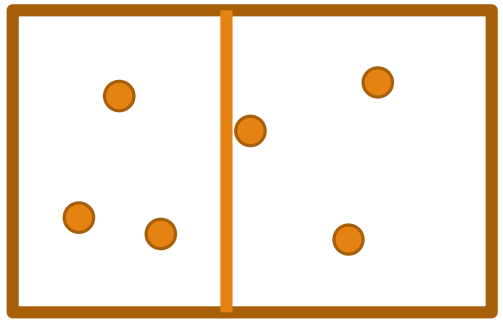


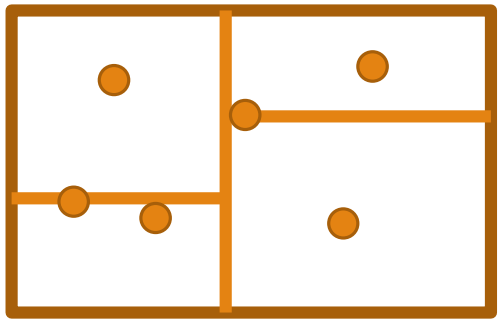
Vectores en hojas

- Dividimos el espacio según la mediana de los vectores
- Cada nodo desde la raíz corresponde a una división (horizontal o vertical) del espacio
- Subdividimos hasta alcanzar los casos bases:
  - Número de vectores en la caja
  - Altura del árbol
  - Otro
- Todos los vectores estarán en las hojas del árbol









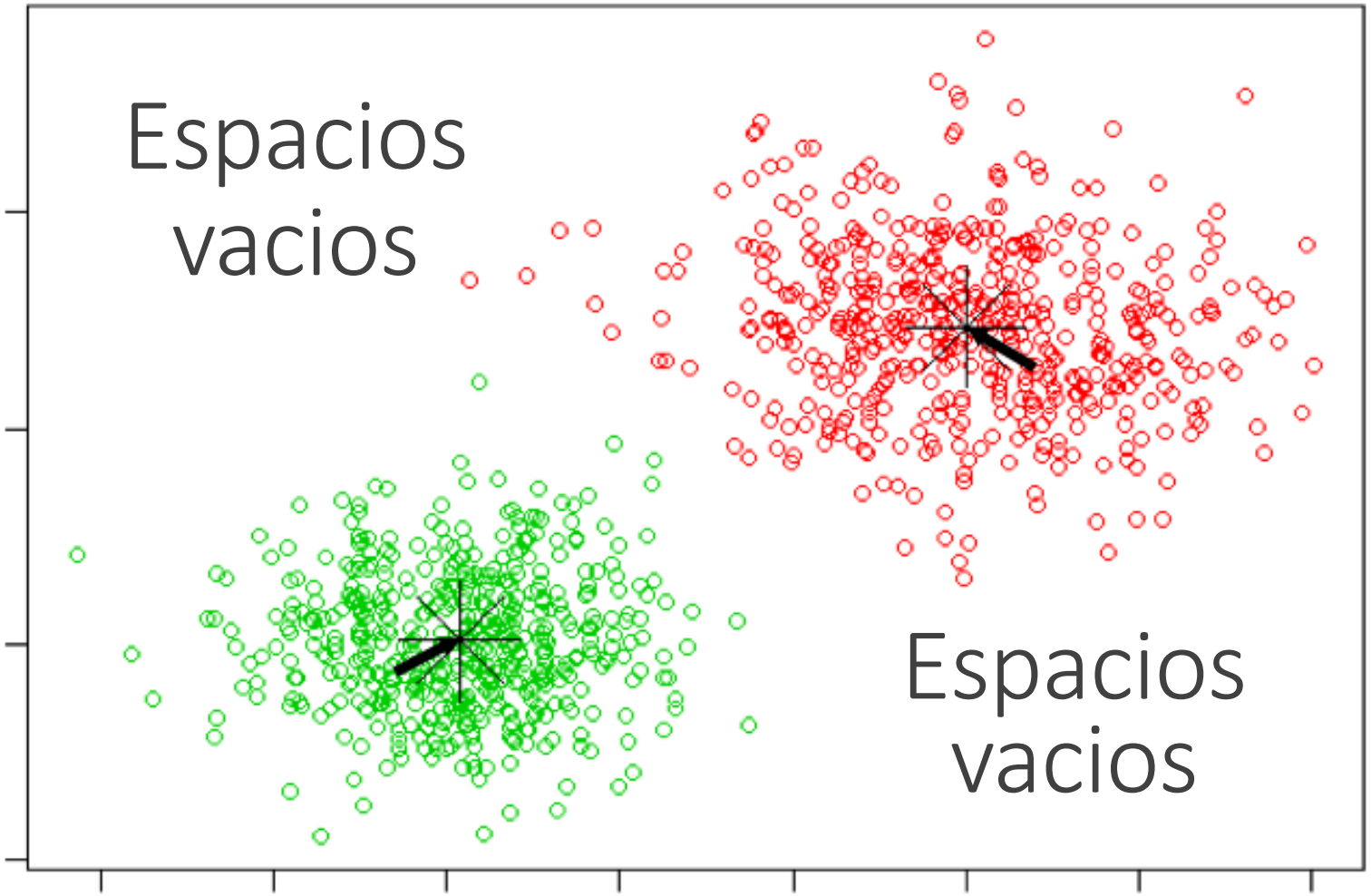
Paso a paso

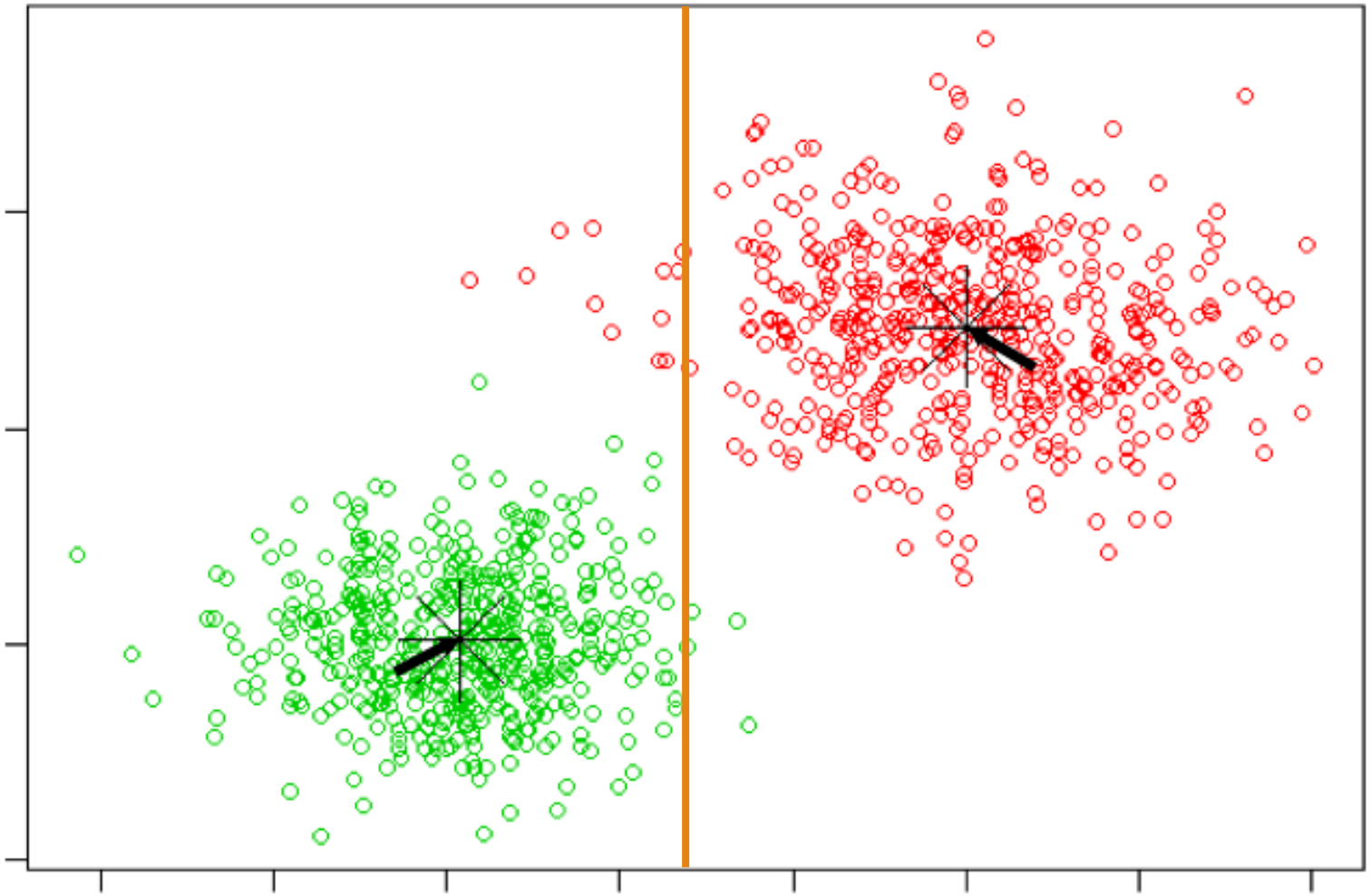
vectores x caja
<=
2

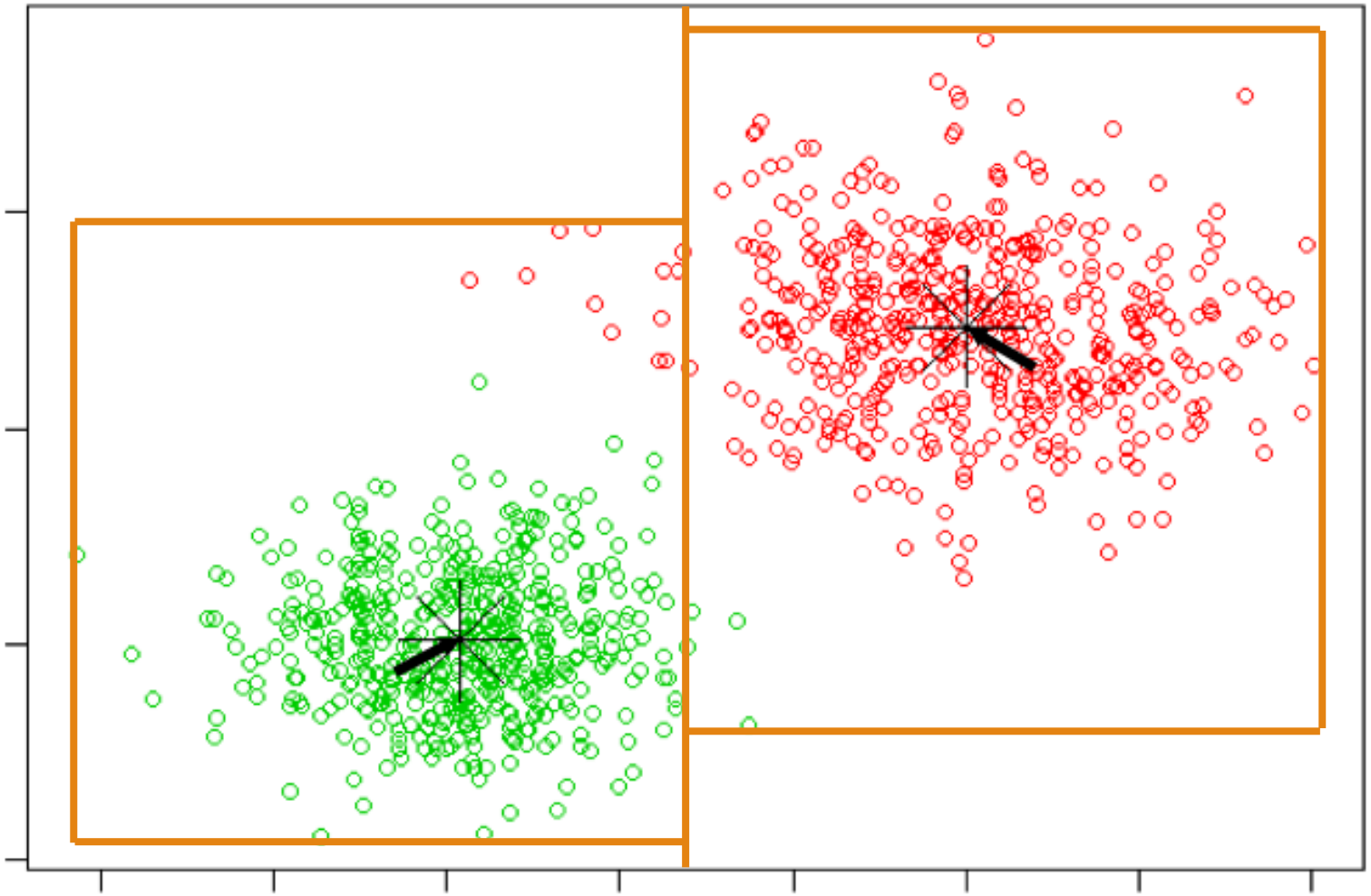
Si no se usa un KDTree

- Debemos iterar sobre cada vector del DATA SET de entrenamiento (N)
- Debemos comparar la distancia a cada uno de los vectores del test (R)
- Debemos revisar sobre cada uno de los K-vecinos ya revisados (K)

Complejidad O(N*R*K)


Espacios
vacios
Espacios
vacios





K-Nearest Neighbour

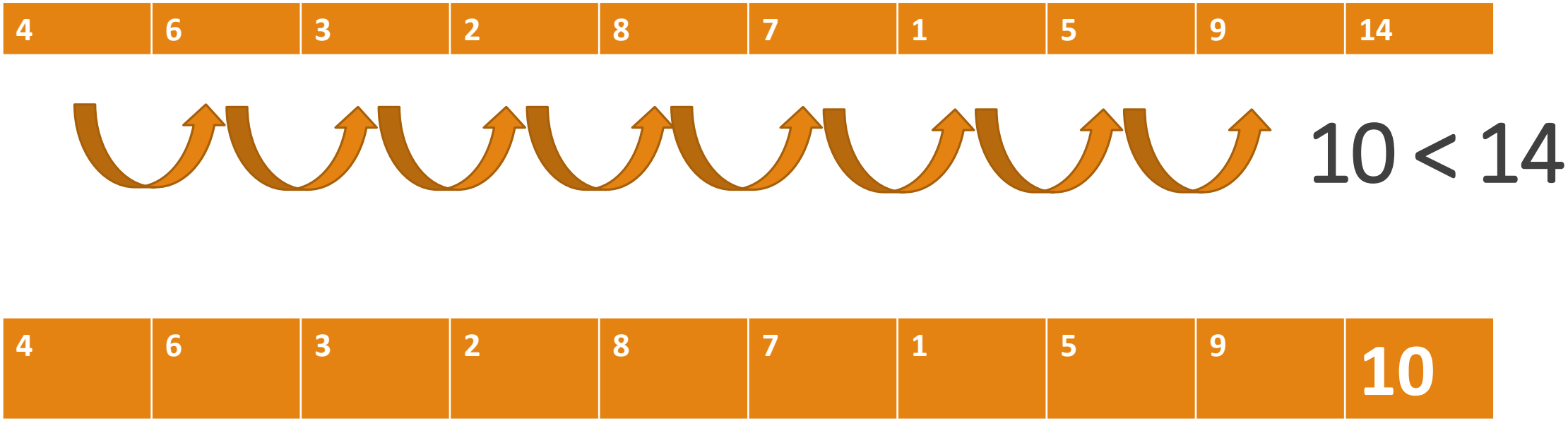
- Necesitamos asignar un Label a cada vector del test
- Debemos encontrar los K-elementos más cercanos
- **¿Cómo?**

K = 10 | | Valor representa distancia

Insertar hasta que esté lleno no es problema

Y ahora? Vector a distancia 10…

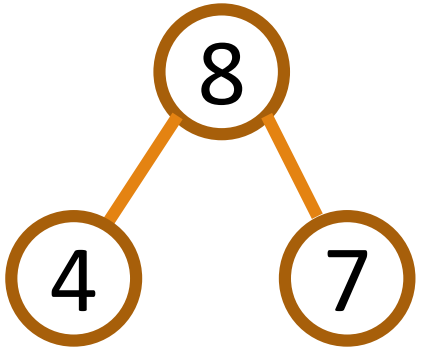
Debo revisar cada uno…

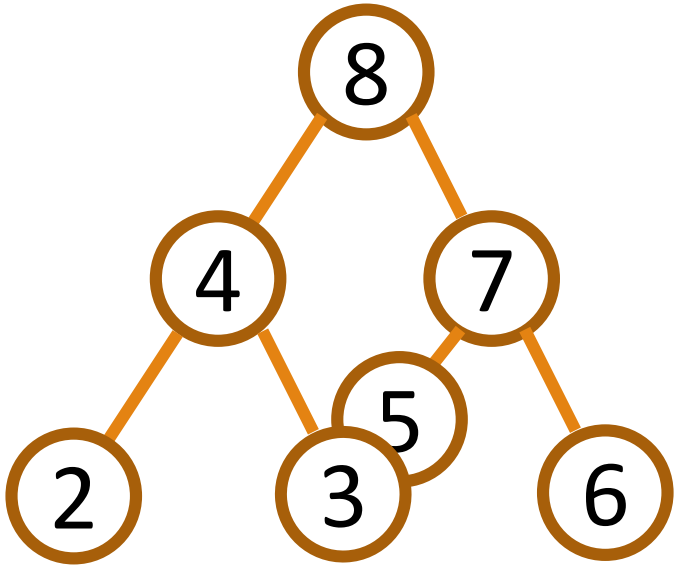


Complejidad O(K)…

MAX-HEAP

K = 7

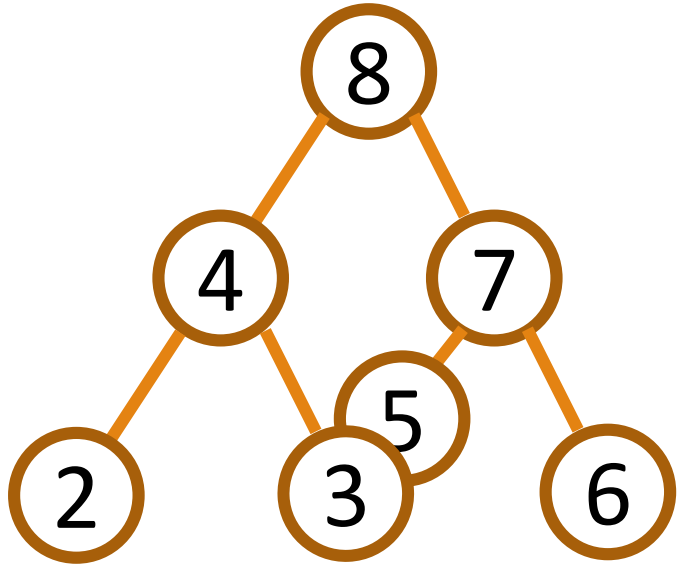




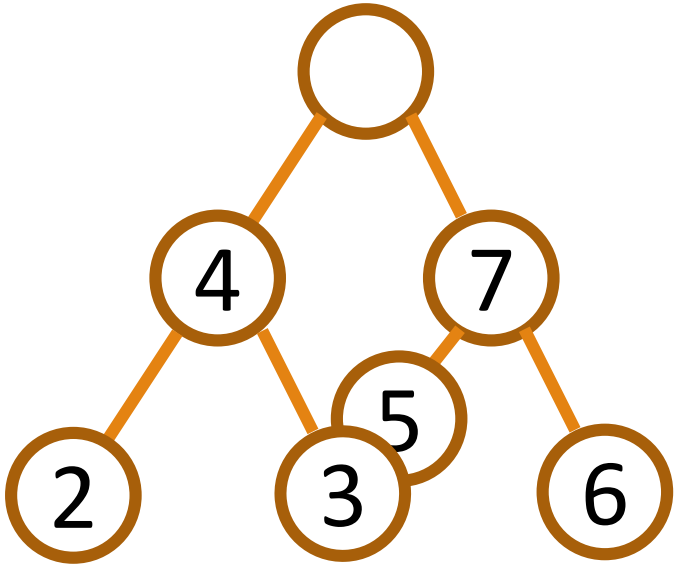
Revisar vector a distancia 1

- Sabemos cuál es el mayor
  - Eliminar raíz
  - Insertar 1 en raíz
  - HEAPIFY

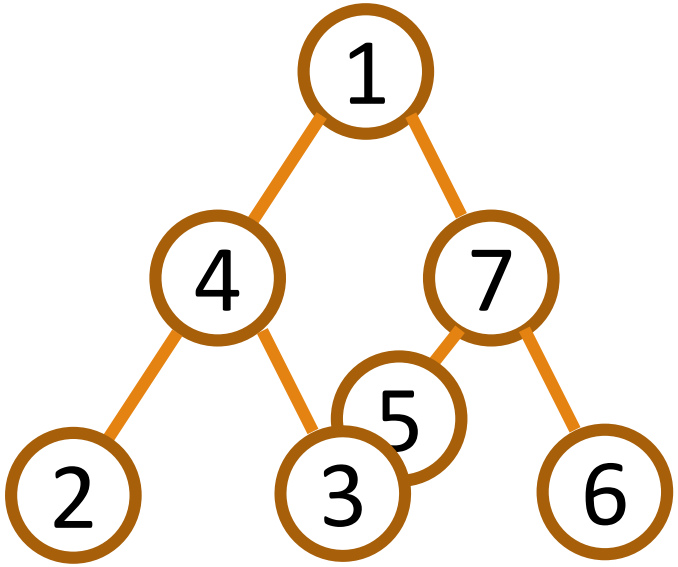
$8 > 1$



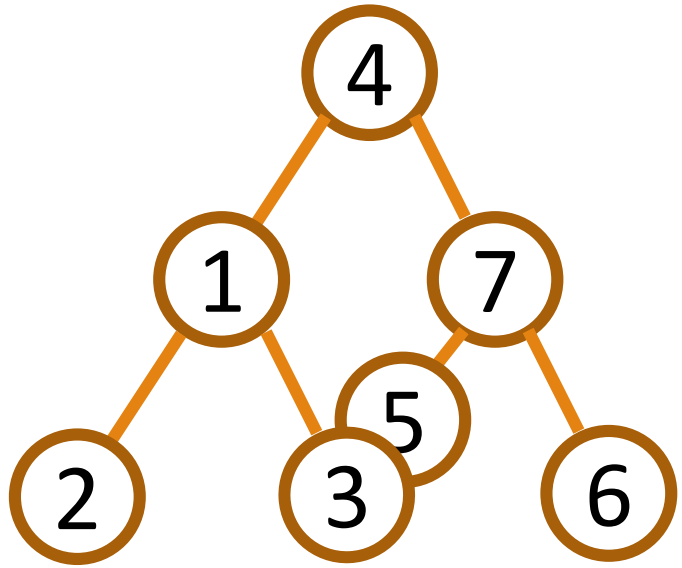
- Sabemos cuál es el mayor
- **Eliminar raíz**
- Insertar 1 en raíz
- HEAPIFY

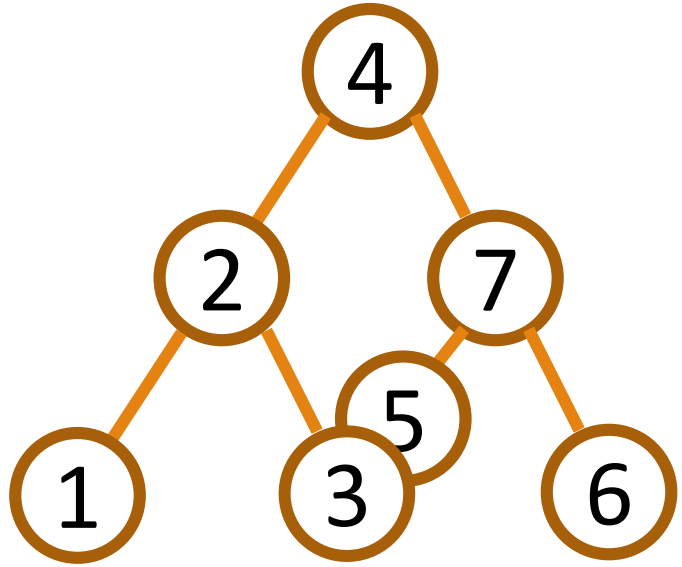


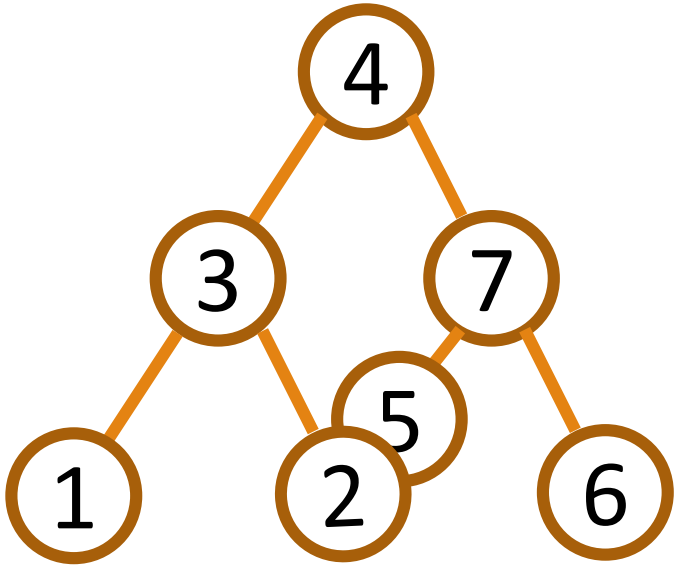
- Sabemos cuál es el mayor
- Eliminar raíz
- Insertar 1 en raíz
- HEAPIFY

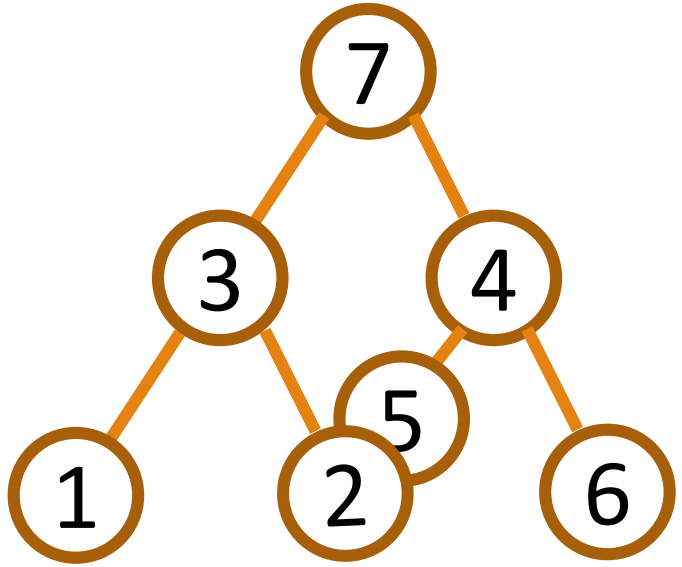


- Sabemos cuál es el mayor
- Eliminar raíz
- Insertar 1 en raíz
- HEAPIFY

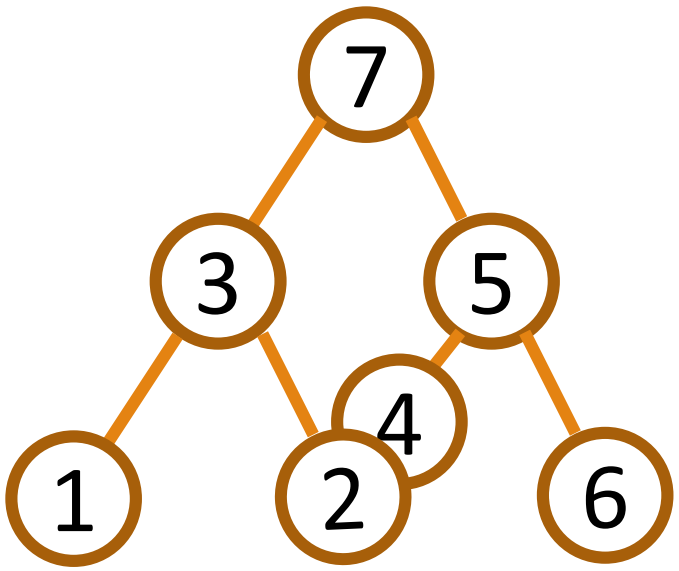


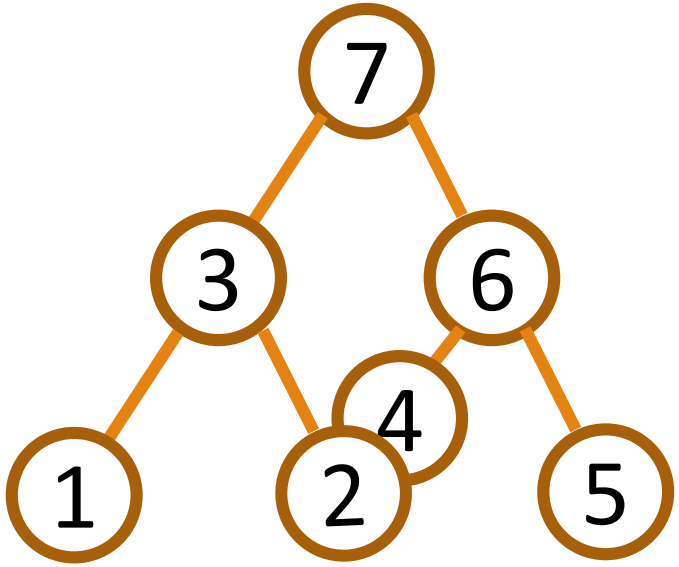






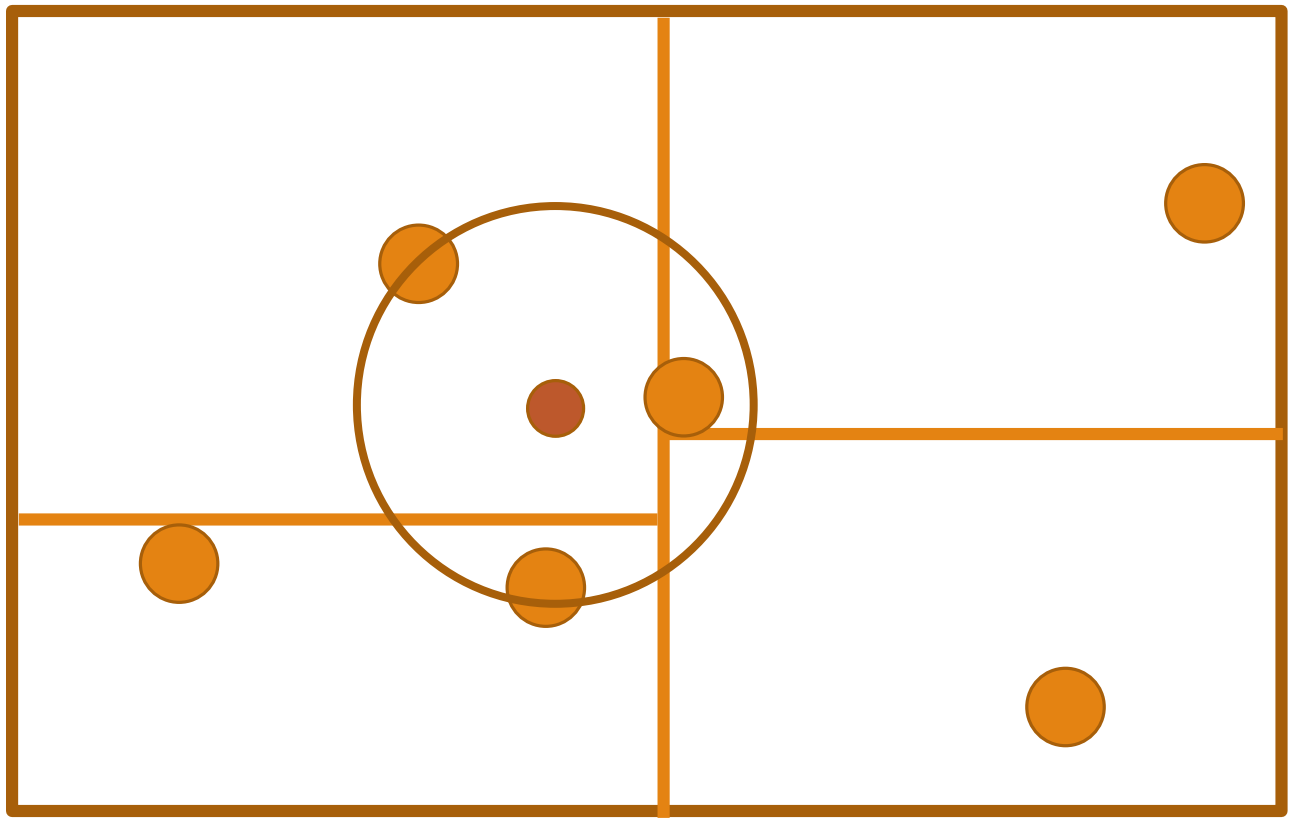
HEAPIFY hasta que se cumpla la condición del Heap en cada nodo…





Búsqueda en el KDTree

```
search(kd, vector, heap):
        if (!collision(kd, heap)):
                return;
        if (leaf):
                iterar por los vectores
                search(hijo que corresponde)
                search(el otro hijo)
```



Existe collision con otra caja… revisaré esa caja

ESTRUCTURAS DE DATOS Y ALGORITMOS IIC2133

ANTONIO LOPEZ – MANUEL BECKER

Ayudantía T1 KDTree