

# Análisis de Algoritmos II

Un algoritmo de barrido de línea para agrupación espacial.

**Profesores:**

Jorge Urrutia Galicia

Adriana Ramírez Viguera

Diego Jesús Favela Nava.

---

Aguilera Moreno Adrian.

Facultad de Ciencias, UNAM



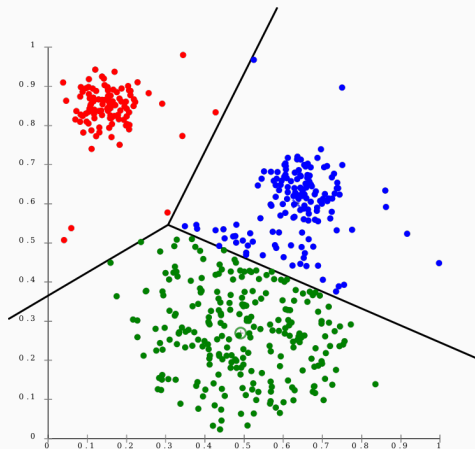
1

## Introducción.

---

- Categorías.
- Alternativas.
- Agrupación Espacial.

Un conjunto de datos agrupados en  $k$  grupos “similares”.



Agrupamiento por K-means.

¿Cómo agrupar? ...

# Categorías.

Existen dos categorías principales para realizar agrupamientos:

## 1. Algoritmos de agrupamiento *jerárquico*.

### 1.1 Aglomerantes.

Inicialmente cada objeto es un grupo, conforme transcurren las iteraciones los objetos se deben ir fusionando.

### 1.2 Divisivos.

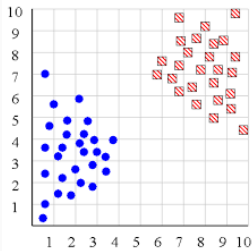
Inicialmente todos los objetos son un solo grupo, conforme transcurren las iteraciones se van subdividiendo estos grupos.

## 2. Algoritmos de agrupamiento *particional*.

Cada objeto se asocia con el centro de agrupamiento del más cercano.



Agrupamiento Particional.



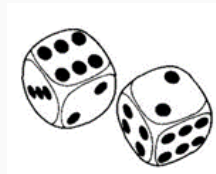
Agrupamiento Jerárquico.

# Alternativas.

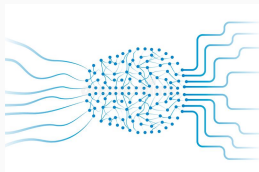
Algunas alternativas para agrupar son:

1. **Redes neuronales.**
2. **k-means + Algoritmos genéticos.**
3. **Muestreos aleatorios.**

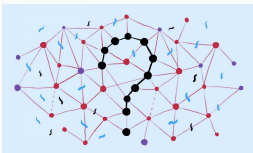
Algunas alternativas para agrupar basadas en el entrenamiento inteligente, búsquedas aleatorias (como las heurísticas), y uso de algoritmos genéticos (como las colonias de hormigas) son recurridas cuando no podemos garantizar un “buen” agrupamiento.



Aleatorios.



Redes Neuronales.



K-means + Genéticos.

# Agrupación Espacial: Propuestas I.

La agrupación espacial es un subconjunto espacial de agrupación. Este tipo de agrupamiento es relacionado, con frecuencia, a métodos gráficos.



1era ley de la geografía.

## Propuestas:

- Zahn sugiere trabajar con un gráfico completo (con vértices cada elemento en el espacio), construir el árbol de expansión mínima y eliminar los “bordes” más largos comparando las longitudes de los arcos con la longitud promedio, eliminando aquellos con longitud mayor al doble de la longitud promedio.



## Propuestas:

- Narendra sugiere el uso de diagramas de Voronoi para agrupar en tiempo  $\mathcal{O}(n \log n)$ . El problema de esta solución es que los algoritmos son difíciles de implementar.
- Kang usó triangulaciones de Delaunay y un diagrama dual de Voronoi. Después de construir la triangulación en  $\mathcal{O}(n \log n)$ , eliminamos las aristas con longitudes mayores a  $d$ .
- Yujian presentó un algoritmo de agrupamiento en subárboles máximos en distancia.

## ¿Problemas? ...