



# Tema 1

## Actividad 1

Autor

Juan José Méndez Torrero

# Generación de patrones frecuentes

## Introducción

Tras ejecutar por primera vez el algoritmo **Apriori** con un soporte mínimo de 0.1, se puede observar que el algoritmo ya nos indica que el soporte mínimo es de 0.35, con lo que para esta actividad se considerarán los soportes mínimos de 0.35, 0.5 y 0.7.

## Soporte mínimo: 0.35

### Resultados

Tras ejecutar el algoritmo con esta configuración, se obtienen los siguientes patrones frecuentes o *large itemsets*:

- L(1): Tiene un tamaño de 4 itemsets y los soportes absolutos de cada itemset son:
  - class=crew 855
  - age=adult 2092
  - sex=male 1731
  - survived=no 1490
- L(2): Tiene un tamaño de 5 itemsets y los soportes absolutos de cada itemset son:
  - class=crew age=adult 885
  - class=crew sex=male 862
  - age=adult sex=male 1667
  - age=adult survived=no 1438
  - sex=male survived=no 1364
- L(3): Tiene un tamaño de 2 itemsets y los soportes absolutos de cada itemset son:
  - class=crew age=adult sex=male 862
  - age=adult sex=male survived=no 1329

### Conclusión

Se puede observar que con un soporte mínimo de 0.35, el algoritmo Apriori es capaz de encontrar patrones frecuentes hasta en L(3). En este itemset podemos encontrar dos itemsets que nos indican el número absoluto de observaciones encontradas con esa transacción en todo el conjunto de datos.

## Soporte mínimo: 0.5

### Resultados

Tras ejecutar el algoritmo con esta configuración, se obtienen los siguientes patrones frecuentes o *large itemsets*:

- L(1): Tiene un tamaño de 3 itemsets y los soportes absolutos de cada itemset son:

- age=adult 2092
- sex=male 1731
- survived=no 1490
- L(2): Tiene un tamaño de 3 itemsets y los soportes absolutos de cada itemset son:
  - age=adult sex=male 1667
  - age=adult survived=no 1438
  - sex=male survived=no 1364
- L(3): Tiene un tamaño de 1 itemset y los soportes absolutos de itemset es:
  - age=adult sex=male survived=no 1329

### Conclusión

Tras la ejecución del algoritmo Apriori con un soporte mínimo del 0.5, se puede observar que se siguen consiguiendo encontrar patrones frecuentes incluidos en el itemset L(3). Aunque solamente se ha encontrado un patrón frecuente, siendo la transacción, adulto, masculino y no superviviente.

## Soporte mínimo: 0.7

### Resultados

Tras ejecutar el algoritmo con esta configuración, se obtienen los siguientes patrones frecuentes o *large itemsets*:

- L(1): Tiene un tamaño de 2 itemsets y los soportes absolutos de cada itemset son:
  - age=adult 2092
  - sex=male 1731
- L(2): Tiene un tamaño de 1 itemset y los soportes absolutos de itemset es:
  - age=adult sex=male 1667

### Conclusión

Con un soporte mínimo del 0.7, se puede observar que el algoritmo Apriori sólo es capaz de encontrar patrones frecuentes hasta L(2), siendo estos patrones bastante poco identificativo como para ser contado como patrón que discrimine cierta parte de la población.

## Conclusiones generales

Tras la ejecución del algoritmo Apriori con distintos valores para el soporte mínimo, hemos podido observar lo siguiente:

- El número más alto de patrones frecuentes encontrados ha sido con un soporte mínimo de 0.35, llegando hasta encontrar 2 patrones frecuentes en el conjunto de ítems L(3).
- Se puede observar que valores mas altos que 0.35 hace que el algoritmo Apriori no sea capaz de encontrar tantos patrones frecuentes, y este número se ve más reducido aún si el soporte mínimo supera el 0.6.