

Aprendizaje No supervisado

Reglas de asociación caso Movielens MovieLens

Grupo:

- Martín Hunziker
- Ramiro Caro
- Claudio Sarate

Introducción

El objetivo del trabajo práctico fue utilizar el algoritmo Apriori para realizar recomendaciones de películas.

Desarrollo

Para la realización del trabajo práctico se utilizaron dos implementaciones del algoritmo A priori, la implementación del paquete `efficient_apriori` (presentado en clase) y la del paquete `mlxtend`.

Luego de realizar diversos test se verificó que ambas implementaciones tienen performances similares, sin embargo el paquete `mlxtend` presenta los resultados dentro de un dataframe de Python que le da mas flexibilidad a la hora de analizar los resultados.

Se trabajó sobre el dataset reducido porque ninguno de los dos algoritmos soportó trabajar sobre el dataset completo aduciendo falta de memoria, lo que nos permitió experimentar lo exigente en recursos que es el método.

De la misma forma, cuando el valor de la variable `min_support` disminuye la cantidad de casos analizados aumenta en forma exponencial hasta el `min_support` de 0.01 donde se presentan todas las combinatorias posibles. Como era de esperar al igual que los tiempos de procesamiento como se puede observar en la tabla a continuación

min_support	Casos	Tiempo
0.001	1 411 132	2.61 seg
0.01	1 411 132	2.55 seg
0.1	1 316	0.02 seg
0.2	10	0.01 seg

El tipo de recomendación de este algoritmo nos permite investigar una relación subyacente dentro de las películas.

Sería interesante ver otros algoritmos que permitan trabajar con dataset mayores.