

Análisis de resultados Reto 3

Estudiante A: Pablo Alejandro Guatibonza Briceño (202014393)

Estudiante B: Juan Diego Calixto Nuñez (202020774)

Análisis de complejidad en notación O para cada uno de los requerimientos.

• Análisis de tiempo de ejecución y uso de memoria para cada uno de los requerimientos implementados.

Requerimiento 1:

1. Se obtienen una lista a partir de los valores de una tabla de hash con los eventos únicos que se encuentran en ambos csv (context, user_track) $O(1)$
2. Recorrido total de la lista filtrando aquellos eventos que se encuentren en el rango de la característica especificada por el usuario. $O(n)$
3. Los eventos filtrados se agregan a un mapa nuevo bajo el key artist_id y se lleva una cuenta de cuantos eventos hay. $O(1)$

Requerimiento 2:

1. Desde la carga de datos se crea un RBT donde las llaves son los valores de 'danceability' y sus valores son tablas de hash con los eventos únicos.
2. Se extraen los valores (tablas de hash con eventos únicos) de las llaves del RBT que se encuentran dentro de los rangos establecidos por el usuario. $O(2*\text{Log}(n))$
3. Se recorren todas las tablas de hash y por cada una se recorren todos sus eventos los cuales se filtran por un rango de 'energy' especificado por el usuario. $O(n*m)$, n= cantidad de tablas de hash encontradas en el RBT & m = cantidad de eventos únicos por tabla
4. Se agregan los eventos filtrados a un mapa y se retorna una lista a partir de sus valores. $O(1)$

Requerimiento 3:

1. Desde la carga de datos se crea un RBT, donde las llaves son los valores del tempo, y sus valores son listas con los eventos únicos.
2. Se extraen los valores (TAD listas con eventos únicos) de las llaves del RBT que se encuentran dentro de los rangos establecidos por el usuario. $O(2*\text{Log}(n))$
3. Se recorren todas las listas y por cada una se recorren todos sus eventos los cuales se filtran por un rango de 'instrumentalness' especificado por el usuario. $O(n*m)$, n= cantidad de TAD listas encontradas en el RBT & m = cantidad de eventos únicos por tabla
4. Se agregan los eventos filtrados a un mapa y se retorna una lista a partir de sus valores. $O(1)$

Requerimiento 4:

1. Desde la carga de datos se crea un RBT, donde las llaves son los valores del tempo, y sus valores son listas con los eventos únicos.

2. Se emplea una función para que el usuario pueda agregar o escoger géneros musicales para su análisis. Esto crea una lista con los géneros seleccionados. $O(1)$
3. Luego por cada género se crea una lista con los eventos únicos utilizando la misma función que en el requerimiento 3 a partir del RBT clasificado por tempo. $O(2 \cdot \log(n))$
4. Luego se crea una sublista con los primeros 10 eventos. $O(10)$
5. Se recorre la lista del género filtrado y se crea un mapa donde se guarda por `artist_id`, para retornar la cantidad de artistas únicos. $O(n)$

Requerimiento 5:

1. Se crea un RBT con llaves que son la fecha del evento (en segundos) y valores que son listas, con los eventos que están la misma fecha.
2. Se filtran las llaves del RBT y se seleccionan los valores los cuales su fecha está en el rango indicado por el usuario. $O(2 \cdot \log(n))$
3. Se recorren todas las listas obtenidas anteriormente y por cada una se recorren todos sus eventos los cuales se incluyen en : un diccionario que tiene como llave el nombre de los géneros y como valor, una tabla de hash con los eventos únicos que pertenecen a dicho género ; y en un diccionario de géneros que tiene como llave el nombre de los géneros y como valor, una tabla de hash con las canciones únicas de dicho género . Se retorna estos 2 diccionarios . $O(n \cdot m \cdot 8)$, n = cantidad de TAD listas encontradas en el RBT & m = cantidad de eventos por lista & 8 =cantidad de categorías musicales
4. Se ordenan descendientemente los diccionarios obtenidos anteriormente de acuerdo al tamaño de cada una de las tablas de hash de los eventos únicos. Se retorna el diccionario ordenado con llaves los géneros y valores la cantidad de eventos únicos, el nombre del género con mas reproducciones, y la cantidad de artistas únicos por el género con mas reproducciones . Se imprimen las llaves de los diccionarios junto con su valor(cantidad de reproducciones) $O(8 + 8 \cdot 8 + 8)$
5. Se escogen 10 canciones aleatorias del género con más reproducciones y por cada una de ellas se halla una lista con los hashtags que tiene esta canción. Se compara si este hashtag está dentro del csv de sentiment value y si lo está y es diferente de 0 , añadimos su `vader_avg` para hacer su promedio. Se imprime el `track_id` de las canciones aleatorias, junto con su hashtag y su promedio vader. Complejidad $O(1 \cdot 10 \cdot m)$ donde m es la cantidad de hashtags que tiene cada canción