

Calificación de películas

Emmanuel Peto Gutiérrez
Rodrigo Fernando Velázquez Cruz

IIMAS
UNAM

7 de diciembre de 2022

Introducción

Manager-Workers

Entrada y salida de
datos

Descripción del problema

Calificación de
películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de
datos

Los usuarios de servicios de streaming tienen problemas para seleccionar películas que le agraden y requieren una recomendación basado en lo que les ha gustado y lo que no les ha gustado.

Objetivo

Calificación de
películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de
datos

Diseñar un sistema que filtre información sobre las películas y sobre los datos de los usuarios que han calificado las películas. También, mostrar estadísticas sobre los resultados filtrados.

El patrón Manager-Workers

Calificación de
películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

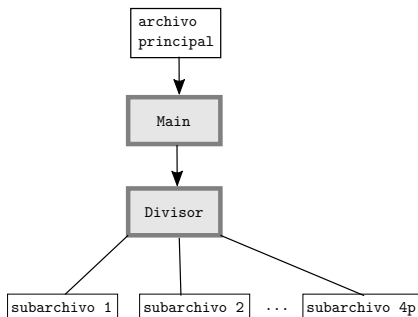
Manager-Workers

Entrada y salida de
datos

Se utilizó el patrón de diseño en paralelo *Manager-Workers*.
A continuación se explicarán los pasos realizados por el
programa.

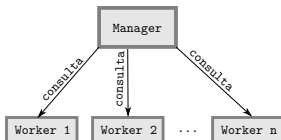
Paso 1

Se calcula el número de hilos y se divide el archivo grande en $4p$ subarchivos, donde p es el número de hilos disponibles en el CPU.



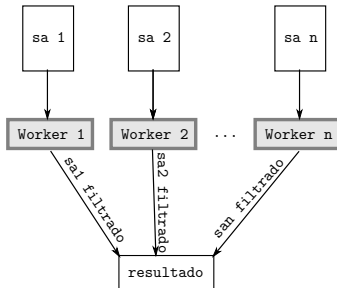
Paso 2

- Interpretar la información ingresada por el usuario, la cual es: selección de columnas y condición de filtrado para los registros.
- El hilo principal le pasa al manager la información interpretada y el número de hilos.
- El manager lanza $4p$ hilos y le pasa la consluta interpretada a cada uno.



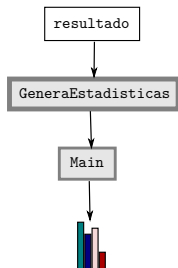
Paso 3

Una vez que el *worker* conoce la consulta, aplica el filtrado y la selección sobre su subarchivo. Después, escribe sobre un archivo compartido “resultado.csv”, de manera sincronizada.



Paso 4

Cuando ya se filtraron los subarchivos y se guardaron en el archivo “resultados”, se calculan estadísticas de forma secuencial como promedio, mediana, mínimo y máximo, y se muestran en gráficas de barras.



Selección

Primero, el usuario debe elegir los nombres de las columnas que desea seleccionar para el resultado. Estas se colocan en el campo de texto que está justo debajo de la etiqueta:

Indique las columnas a seleccionar.

Las columnas que seleccione deben estar separadas por comas y las opciones son:

idRating, userId, movieId, rating, timestamp, title, year, genres, name, lastname, age, imdb, themoviedb.

Calificación de películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de datos

Selección

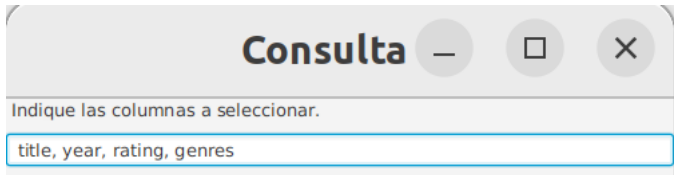
Primero, el usuario debe elegir los nombres de las columnas que desea seleccionar para el resultado. Estas se colocan en el campo de texto que está justo debajo de la etiqueta:

Indique las columnas a seleccionar.

Las columnas que seleccione deben estar separadas por comas y las opciones son:

idRating, userId, movieId, rating, timestamp, title, year, genres, name, lastname, age, imdb, themoviedb.

Ejemplo, se pueden seleccionar las columnas: title, year, rating, genres.



The screenshot shows a web browser window with a title bar that says "Consulta". Below the title bar is a text input field. The input field has a placeholder text "Indique las columnas a seleccionar." and the user has entered the text "title, year, rating, genres".

Calificación de películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de datos

Condición

Luego, se deben elegir las condiciones de filtrado y se deben escribir en forma normal disyuntiva.

Por ejemplo, se pueden filtrar los registros de ratings para todas las películas que hayan sido calificadas por personas de entre 20 y 30 años, para cualquier película que haya salido después del 2000, tal que su género sea Thriller u Horror con la siguiente expresión:

Calificación de
películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de
datos

Condición

Calificación de
películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Luego, se deben elegir las condiciones de filtrado y se deben escribir en forma normal disyuntiva.

Por ejemplo, se pueden filtrar los registros de ratings para todas las películas que hayan sido calificadas por personas de entre 20 y 30 años, para cualquier película que haya salido después del 2000, tal que su género sea Thriller u Horror con la siguiente expresión:

`(age >= 20 AND age <= 30 AND year > 2000 AND genres = Thriller) OR (age >= 20 AND age <= 30 AND year > 2000 AND genres = Horror)`

Introducción

Manager-Workers

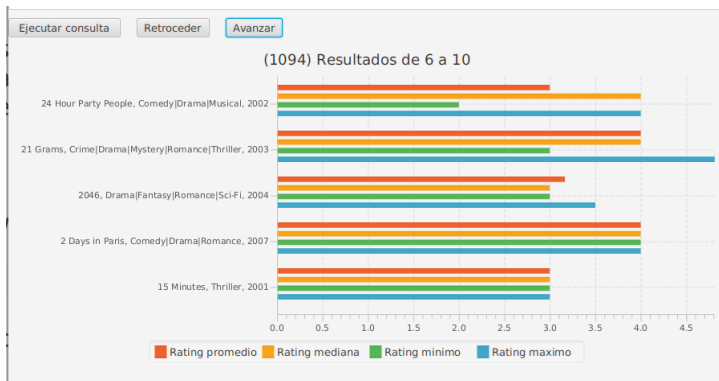
Entrada y salida de
datos

Escriba las condiciones de filtrado en forma normal disyuntiva.

`(age >= 20 AND age <= 30 AND year > 2000 AND genres = Thriller) OR (age >= 20 AND age <= 30 AND year > 2000 AND genres = Horror)`

Resultados

Los resultados se presentan en gráficas de barras y los datos que se observan son los de ratings: media, mediana, mínimo y máximo.



Calificación de películas

Emmanuel Peto
Gutiérrez
Rodrigo Fernando
Velázquez Cruz

Introducción

Manager-Workers

Entrada y salida de datos