TubeHarvest

Session: 2022 – 2026

Submitted by:

Muhammad Wali Ahmad	2022-CS-65
Leena Zaheer	2022-CS-72

Supervised by:

NAZEEF UL HAQ

Department of Computer Science

University of Engineering and Technology

Lahore Pakistan

Table of Contents

1. <u>Problem Statement:</u>

We are making an app that shows YouTube videos stats in a user-friendly way. Each video has different info. Our app makes this data easy to understand with tabular representation. User can also compare two videos based on likes, comments, and length. This helps them decide if a video is worth watching.

Users don't need to sign up, it's hassle-free. Users can search and sort videos themselves. The app's design is user-friendly. While collecting data, we added features like pausing, starting, resuming, and stopping. A progress bar shows how it's going.

Sorting:

User can check which videos are the most popular and which ones are the least popular based on various factors like number of likes, comments, and views. He can use different methods to sort the videos, and it is up to him to decide how he wants to sort them e.g., ascending or descending. This sorting can be done using any characteristic of the videos.

Searching:

User can find videos using different criteria like the video's title, video link, or the name of the channel. If there are multiple results, user can see a preview of all of them. This makes it possible to use search methods effectively, so user can easily find the videos he is looking for among many records.

2. <u>Benefits:</u>

Unlike YouTube Studio, where you can only see stats about your own videos, this platform will allow you to compare and analyze a lot of entries using various filters. It is like websites such as GSMArena, IMDb, or Wikipedia, but specifically for analyzing YouTube videos. The main goal is to provide user with insights and information about videos so they can decide what to watch or what to avoid.

3. <u>Technical Details:</u>

Following table show the entities that we scrape from videos of YouTube. We prefer news channels videos because they have huge number of videos, but we did not scrape 5 or 10 channels we scrape almost 100 plus channels to meet the count of one million. Firstly, we scrape video links from each channel then we scrape each link individually we get 8 attributes of a single video.

Name	Data type	Description
Title	String	Name of the video
Channel Name	String	Name of the channel which published the video.
Channel Subscribers	Integer	Total subscribers the channel has.
Likes	Integer	Total likes on the video.
Comments	Integer	Total comments on the video.
Duration	Integer (Seconds)	The actual length of the video.
URL	String	URL of the video.
Views	Integer	Total views on the video.

Sample of Scrapping Source:

4. <u>Requirements Details:</u>

4.1   Sorting Algorithms:

We apply following sorting algorithms in our project and the time consumed is shown in label:

Comparison based sorting algorithms:

- Selection Sort
- Bubble Sort
- Insertion Sort
- Merge Sort
- Hybrid-Merge Sort
- Quick Sort
- Odd Even Sort
- Heap Sort

Linear time sorting algorithm

- Counting Sort
- Radix Sort
- Bucket Sort
- Pigeonhole Sort
- Bead Sort

4.2   Searching Algorithms:

We apply following searching algorithms in our project and the time consumed is shown in label:

- Linear Search
- Binary Search
- Jump Search
- Exponential Search
- Fibonacci Search

Also, we implement Filters for string type searching:

Contains: Search for the keyword provided by the user in the title/URL of the video and the channel's name.
Ends with: Search for the keyword provided by the user at the end of the title/URL of the video and the channel's name.
Starts with: Search for the keyword provided by the user at the end of the title/URL of the video and the channel's name.

Equals: Search for the keyword provided by the user in the title/URL of the video and the channel's name if it matches exactly.

4.3   Multi-Level Sorting:

When there are videos with the same number of comments, we use a two-level sorting approach. First, we sort them based on the number of comments, and if they have the same number, we then sort them based on the number of likes. If, by chance, we still have videos with the same values for both attributes, we can go to a third level of sorting, like using the video title or something unique, such as the video's URL, to arrange them.

5. <u>Other Feature:</u>

The additional feature of our app is that the user can compare two videos' stats in a graphical way. App needs two videos link and check if these videos are present in our data set or scrape their data on spot. After getting stats of each videos the app shows the stats of both video in graphical way
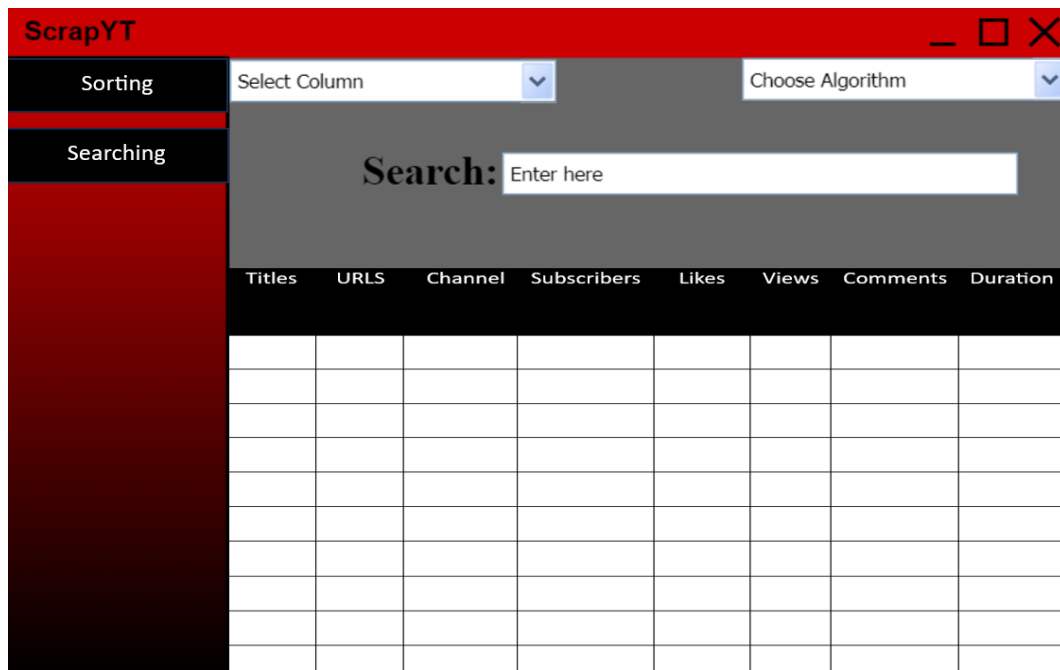
6. <u>Project UI:</u>

6.1   Sorting Page UI
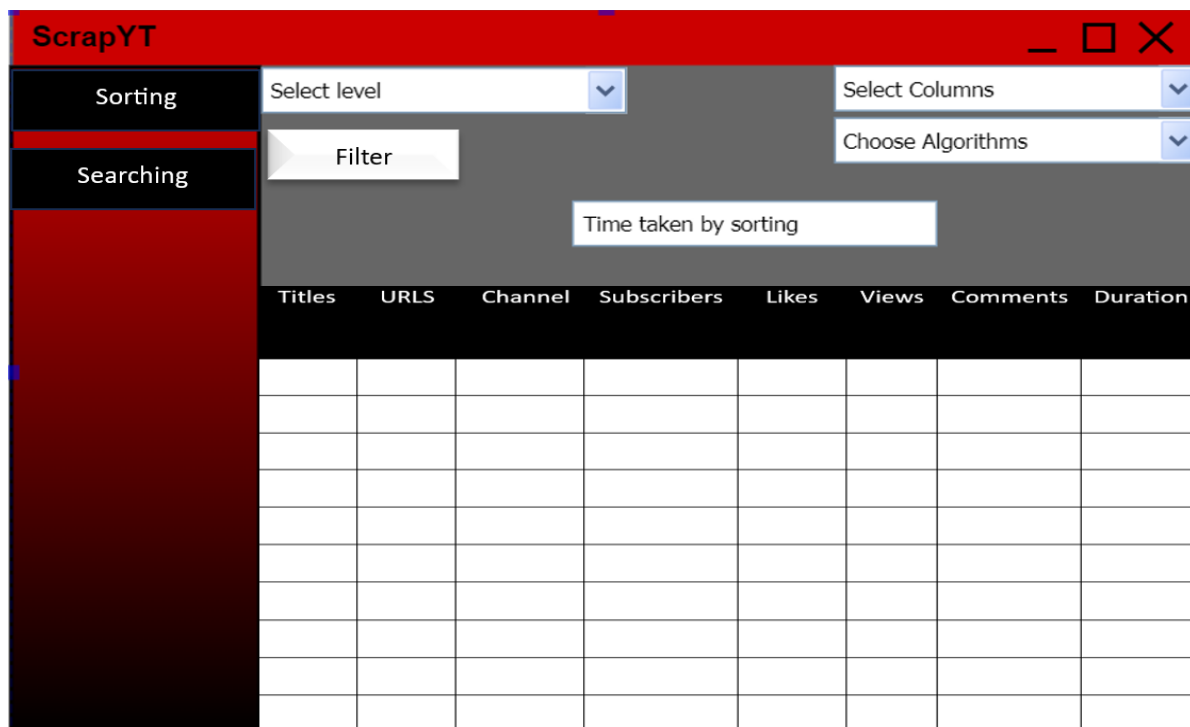
UI components	Component Types	Use of Components
Sorting, Searching button	Side Menu	Click buttons to do sorting or searching
Select Columns	Drop Down Menu	Select no. of columns to sort
Choose Algorithm	Drop Down Menu	Select algorithm on basis of which sorting will be done.
Select level	Button	It opens a new UI to select levels for sorting.
View Time taken	Label	It show time taken by algo in sorting
Table	Grid view	It shows information on all attributes of the video

6.2   Searching Page UI

UI components	Component Types	Use of Components
Sorting, Searching button	Side Menu	Click buttons to do sorting or searching
Select Columns	Drop Down Menu	Select no. of columns to sort
Search Bar	Text box	User will enter entity to search

Choose Algorithm	Drop Down Menu	Select algorithm on basis of which sorting will be done.
Table	Grid View	It shows information on all attributes of the video

6.3 Searched Result UI

UI components	Component Types	Use of Components
Title	Text box	Displays title of searched video
URL	Text Box	Displays url of searched video
Channel	Text box	Displays channel name of searched video
Subscribers	Text Box	Displays subscribers of channel of video
Likes	Text Box	Displays likes of searched video
Views	Text Box	Displays no. of views of searched video
Duration	Text Box	Displays time duration of video
Comments	Text Box	Displays no. of comments of video
Time taken	Text box	Displays time taken in searching