YouTube Best Collection Project

Team 083

1. **Summary**

As the world's leading video-sharing platform, YouTube generates vast amounts of data every day. These trend data, such as trending date, likes, and view count, are valuable resources for understanding user preferences and developing website functionalities. However, with millions of videos uploaded to YouTube every day, it becomes harder and harder for users to select videos that they would like to watch. To solve this problem, YouTube created a functionality called “Trending” that collects some of the most popular videos across different categories with a special formula. However, this functionality has its own drawbacks. For example, users cannot find trending videos for a specific category or region. To enhance the experience of watching YouTube videos, our group decided to create a project called “YouTube Best Collection” that makes further categorization based on the selected trending videos.

In our project, we will implement functionalities like sorting, filtering, and basic searching on our webpage. Our primary goal is to improve users’ watching experience and convenience, and spread high-quality videos. Also, through these implements, we will be able to create weekly recommendations based on the preferences of users.

2. **Description**
   a. **Usefulness**

      Based on our research, applications, and websites with similar functionality are rare. Websites like YouTube have a “trending” functionality but it cannot be filtered or sorted based on user preferences. Most video websites have the functionality of searching, but few of them search by tag instead of title name and provide the option of sorting based on attributes. Therefore, our web application would be much more targeted for users who would like to find anything specific that they would like to watch.

   b. **Functionality**

      Our front-end functionalities include:

      a. Users are able to see the top 10 videos/channels/tags with the highest views and likes for each day, listed by category and region.
      b. Users are able to see a list of must-watch videos per week summarized globally.
      c. Users are able to see all trending videos, sort and filter based on their preferences.
      d. Users are able to search for relevant videos based on their input and are able to sort and filter by attributes.
      e. Users can register for an account, login into account, and change profile information for the account

f. Users can add and delete videos from their favorites according to their preferences.
g. For every video listed above, users are able to access the original video on Youtube by clicking the information panel

Our back-end functionalities include:

a. Perform database maintenance(like add, delete trending videos, update trending video information) automatically based on changes in data provided on Kaggle everyday.
b. Retrieving information from the database based on various requirements provided by the front-end.
c. According to the users' commands, insert or delete videos from their favorites in the dataset.
d. According to the users' commands, add, change, and delete user information.

Front-end & back-end interaction

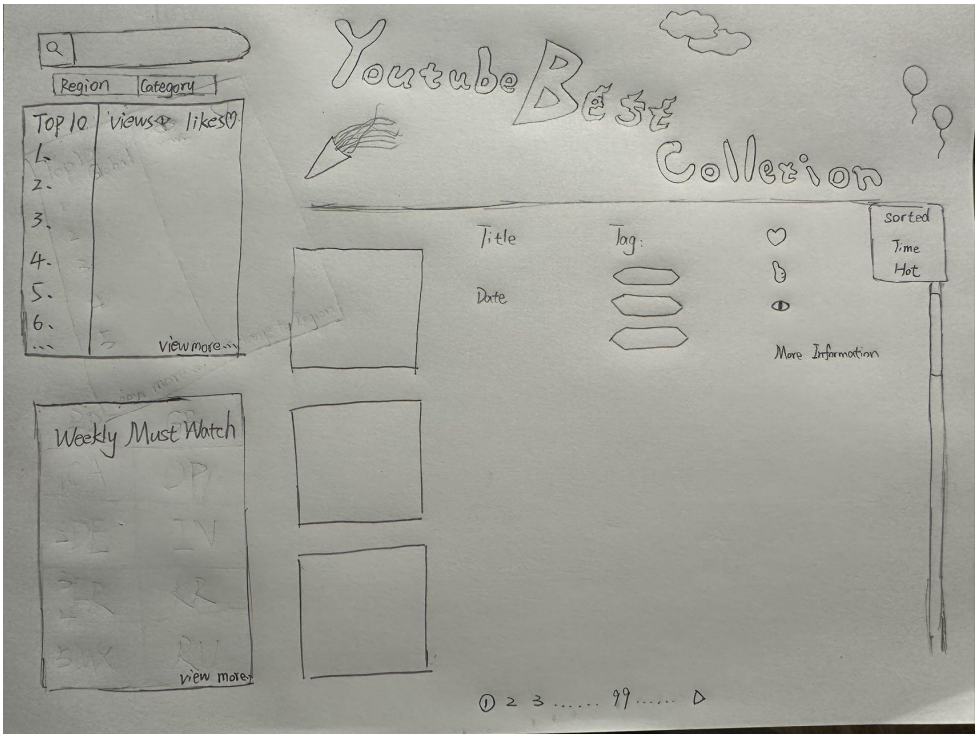
a. Browsing and sorting will be implemented with read commands
b. Register for account will be implemented with create command
c. Logging into account will be implemented with read command
d. Changing account profil will be implemented with update command
e. Adding and deleting videos from favorites will be implemented with create and delete commands

c. Creativity

A creative functionality we are planning to embed into our project is searching. This functionality would be a significant improvement for our website because it offers users the freedom to search for content they are interested in. We are planning to implement the search bar by listing all the videos such that at least one of the tags of the video is contained in the prompt. This would make searching more precise than merely searching by the titles of videos, as it is more likely that the title includes phrases that are literally the same as the prompt but different in meaning. Another functionality that we will implement is a list of best-watch videos. These videos are collected from all trending videos across the eight countries the database includes and select approximately 10 best videos based on a particular formula. These creative functionalities would definitely help users better select video types of their interests and benefit their watching experiences.

d. UI design

Our design of the front-end UI is shown below:



3. **Source**

   The data stored in the real-world dataset about YouTube trending videos was acquired from Kaggle. The original dataset consists of trending video data from eleven different countries, including US, UK, and Japan, each with 16 columns describing the attributes of the video. In our dataset, we are going to columns of store video ID, title, published date, channel ID, channel title, category ID, trending date, tags, view count, likes, dislikes, comment count, thumbnail link, and description. We also need to obtain tables about the meaning of each category ID from the database so that we can group each video into corresponding categories. We might be generating some random data for testing purposes but would not use them in production environments. The dataset link is at https://www.kaggle.com/datasets/rsrishav/youtube-trending-video-dataset?select=US_youtube_trending_data.csv.

4. **Work Distribution**

   The following is an outline of the distribution of work:

   a. Front-end development

      - Zihuang Li: Select top 10 videos/channels/tags with highest views and likes for each day, listed by category and region
      - Tianyi Liu: Summarize the must watch videos per week globally
      - Jiayi Zhang: Display all trending videos where users are able to sort and filter based on their preferences

        - David Fu: Search for relevant videos based on user input and able to sort and filter

b. Back-end development
    - Jiayi Zhang: Link construction of front-end and back-end servers
    - Tianyi Liu: Cloud server deployment
    - David Fu: database operation and maintenance,
    - Zihuang Li: database construction