SI507 Proposal

Tingjun Li, tingjunl, 13131899

Goals and objectives:

Analyze data from a movie database API, and recommend movies to users according to their choices.

Data sources:

I will retrieve data from a movie API:
https://developers.themoviedb.org/3/getting-started/introduction
Method to make a query with my API key:
https://api.themoviedb.org/3/search/movie?api_key=8e82d66faf368b0b37ec9ed221ab3ec7&query=Avengers
Detail about making queries:
https://developers.themoviedb.org/3/getting-started/search-and-query-for-details

Data accessing and storage:

1. JSON file of the movies' titles & their id's can be downloaded through the following request:
   http://files.tmdb.org/p/exports/movie_ids_10_10_2022.json.gz
   According to TMDB, the movie database API I am using, their server will run JSON export job once everyday. I will use this JSON package for my project.
2. Randomly pick 10000 IDs to generate a JSON file that store detail infos about the chosen movies and use it as my cached database.

Interaction with users:

I will give users the following skippable options:

1. Release date (a range)
2. Genres
3. Certification: (NR, G, PG, PG-13, R, NC-17)
4. Movie Rate
5. Movie length

Another idea in my mind:

The program will give the user three sets of movies. A user can choose their favorable movies inside those sets and the program will record them and find similar movies.

Data Processing:

The movies will be stored in several trees according to Genres, then each tree has a branches start from certification, movie rate, movie length then release date. (I am still confused about how to properly set up the data processing part and need some guidance during development).

Once the user finished their selections, I will search in the database and find a set of movies that matches their tastes.

Data Presentation:

Use Flask or Django to make a web application that is interactable with users. The resulting page will present the top 5 match solutions to the users. Ideally, user can click “explore more” to browse another set of movies. Each movie will have a picture, brief description, the cast crew