

COMSC-341 Final Project Code Submission Notes

Jenny and Suravi

Datasets that need to be downloaded

1. Netflix Movie Datasets:
<https://www.kaggle.com/code/laowingkin/netflix-movie-recommendation/input>
 - a. combined_data_1.txt
 - b. combined_data_2.txt
 - c. combined_data_3.txt
 - d. combined_data_4.txt
 - e. movie_titles.csv
2. Movielens Dataset: <https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset>
 - a. movies_metadata.csv
3. GloVe Model for vector representation: <https://nlp.stanford.edu/projects/glove/>
 - a. glove.6B.300d.txt

NOTE: For easier access, all the datasets needed are in the [google drive folder here](#), please remember to change the file paths in the code accordingly.

Uncommon libraries/packages such as `surprise`, `openai`, `nltk` are used and the code already includes the installation of these libraries.

How to run the code?

After downloading all the above mentioned files, please take note of the following things:

1. Change the file paths in the code according to where the files are saved in your local machine. Here is a list of places you want to change the file paths:
 - a. 4 dataframes' file path in cell 4
 - b. Movie lens and movie titles file path in cell 10
 - c. GloVe file path in cell 17
2. Insert your own OpenAI API key in cell 15
3. Remove/uncomment cell 3 you are not using Google Colab drives for data

Run the whole program and the recommender systems will start up!