Project deliverable 3

Movie Recommendation System

1. **Final training results:**

- In the previous work, the model was simply the cosine similarity matrix of movies based on their ratings. In this final work, I integrate both content-based method and collaborative method into my model.
- Content-based method:
    - I first created a word soup that includes the contents, casts, directors, and keywords of every movies. Following, a vector count matrix is created to count number of each words of the soup appear in each movie's meta data.
    - Then, given a movie title, the matrix will return most similar movies based on those meta data.
- Collaborative method:
    - After I have a list of suggested movies, I went on to create a model to pick 10 movies from these list that will each specific user will like.
    - The algorithm is simple; it is clustering and k-nn algorithm.
    - The model tries to group users based-on their ratings. Then, if we want to predict the rating a user to a movie, we simply compute the mean ratings of n users that are closest to this user.
    - For this model, I use the scikit-surprise library, which is a scikit-learn library for recommendation system.
    - Using small ratings dataset, I split it into 5 folds to train and compute the RMSE of final model. The mean RMSE is 0.9 (see the code for more details)
- All the models are saved to disk using python pickle library.
- Results:
    - The content-based model did give a list of movies that are similar to the original movie in term of content.

- The collaborative model did pick different movies from the list for different users (I did some tests on the code)

2. Final demo proposal:

- For the web app, I intend to integrate both content-based model and collaborative model to it.
- User can simply insert a movie name and the app will return list of suggested movies.
- The app can also list some movies and ask user to rate them, then it will these new ratings to suggest movies.
- The back-end will be made using Flask framework, while the front-end is just usual HTML, CSS, and JS.