Group Members:
Everett Hayes, Himanshu Bainwala, Rayaan Attari

Title:
Film Recommendations

Project:
Being confined indoors due to required quarantines, movies have become even more important to our group members. With an ever increasing amount of streaming services and content available, how do we know what to choose? Our project will attempt to streamline this process for users. The user will enter some of their favorite movies, and our algorithm will handpick movies that we think will be to their liking. We plan on using an IMDB dataset on over 3000 movies to facilitate this application. Our algorithm will use IMDB score, popularity, production company, and keywords to derive similar movies. The keywords element will heavily rely on NLP techniques. Source (1) features an example of the algorithm we will be implementing.

Timeline:
Week 1: Find applicable dataset. Convert to JSON object and clean dataset for use.
Week 2: Develop weighted algorithm for determining user's interests given choices.
Week 3: Develop GUI/Android App for representing user choices etc.
Week 4: Wrap Up/Finishing touches

Dataset:
I've pasted a single JSON object from the dataset below. (from the latest James Bond Film)
{"genres":"[{\"id\": 28, \"name\": \"Action\"}, {\"id\": 12, \"name\": \"Adventure\"}, {\"id\": 80, \"name\": \"Crime\"}]","id":"206647","keywords":"[{\"id\": 470, \"name\": \"spy\"}, {\"id\": 818, \"name\": \"based on novel\"}, {\"id\": 4289, \"name\": \"secret agent\"}, {\"id\": 9663, \"name\": \"sequel\"}, {\"id\": 14555, \"name\": \"mi6\"}, {\"id\": 156095, \"name\": \"british secret service\"}, {\"id\": 158431, \"name\": \"united kingdom\"}]","original_title":"Spectre","popularity":"107.376788","production_companies":"[{\"name\": \"Columbia Pictures\", \"id\": 5}, {\"name\": \"Danjaq\", \"id\": 10761}, {\"name\": \"B24\", \"id\": 69434}]","runtime":"148","tagline":"A Plan No One Escapes","vote_average":"6.3"}

Sources:
(1) https://towardsdatascience.com/how-to-build-from-scratch-a-content-based-movie-recommender-with-natural-language-processing-25ad400eb243
(2) https://towardsdatascience.com/how-to-build-from-scratch-a-content-based-movie-recommender-with-natural-language-processing-25ad400eb243

(3) https://www.geeksforgeeks.org/parse-json-java/

Conclusion:

We plan on developing an application to recommend movies to users. Hopefully over algorithm can be generalized to almost anything we collect data on, such as music, TV, video games, etc.