Corey Myers

Professor Bo Liu

COMP 5600-001

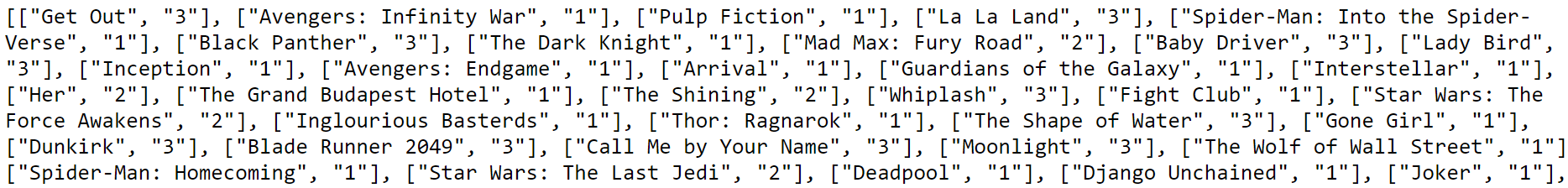
9 December 2019

Movie Recommendation Engine

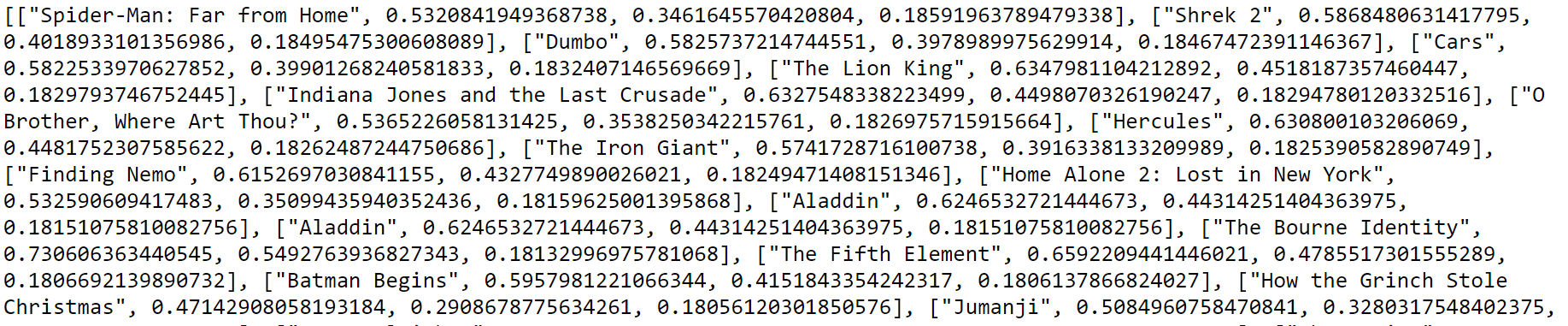
A common problem in today’s society is the uncertainty of not knowing if one enjoys a movie. This could be an impactful problem given that many people buy movie tickets with the chance of not liking the movie they are paying to watch. Thus, there must be a need for statistical evidence to see if a movie is liked or disliked. The foundation of this engine is to show how reviews of popular movies can be used to create percentages of likelihood if the user will like and dislike a certain movie.

Initially, the user had to input three to five movies that they liked and disliked. Then, the reviews for the movies will be parsed, creating a vocabulary of important words found. The system will then locate any movies that have similar or contrary word vocabularies to the user-inputted liked and disliked movies. The system will then output three to five different movies that the user may like and dislike based on the similarities and dislikes of words between the input movies and output movies.

After consideration, the most optimal strategy for the final iteration of the system was to have the user input a list of movies that they have liked, disliked, or have not seen. Then, the system will parse through the top 100 reviews of each of the top 1000 most popular movies. The system will then output the list of movies with a percentage score depending on if the system thinks the user will like, dislike, or has not seen a certain movie, depending on the word vocabulary in each review.



The above picture shows a sample of movies that the user may input into the system: putting a “1” if they liked the movie, a “2” if they disliked the movie, or “3” if they have not seen the movie. The system will then create dictionaries of words from the reviews of movies that the user input. The words will then be compared to the words in reviews of the top 1000 movies and create a percentage of how likely the user would like, dislike, or not have seen the movie, as seen in the picture below.



One negative consequence of implementing the initial structure of the system was slow processing. In the initial iteration of the engine, the system was using dynamic loading to load movies, reviews and words into the dictionaries. This method proved to be very slow and impractical based on the assumption that a movie recommendation system should be relatively quick to produce an output.

The article “A movie recommendation algorithm based on genre correlations” uses genres and how they combine to recommend movies using an algorithm to attempt to solve different issues with current movie recommendation systems. Genres and movie ratings are input by the user to figure out which movies the user might like. One issue the article’s authors found as the sparsity problem. However, the system described in this paper includes a large preset database with large vocabularies and dictionaries. Another potential problem is the cold state problem. However, the system described in this paper retrieves dictionaries from IMDB movie reviews, which are sometimes available before the movie is released to the general public.

“A hybrid approach for movie recommendation” proposes creating a movie recommendation engine using a plethora of content-based and collaborative filtering techniques. The former concept describes the extraction of some features from a source and comparing the features to that of other sources in order to create recommendations. The latter concept predicts similarities between the current user and others. The predictions are made by comparing the closest group of similar user preferences. This reference was vital if the system described in this paper potentially showed that content-based filtering did not prove efficient and effective.

Works Cited

Choi, Sang-Min, et al. “A Movie Recommendation Algorithm Based on Genre Correlations.” *Expert Systems with Applications*, vol. 39, no. 9, 2012, pp. 8079–8085., doi:10.1016/j.eswa.2012.01.132.

Lekakos, George, and Petros Caravelas. “A Hybrid Approach for Movie Recommendation.” *Multimedia Tools and Applications*, vol. 36, no. 1-2, 2006, pp. 55–70., doi:10.1007/s11042-006-0082-7.