A3_DocumentNumber_1_Joel_Patteson

May 30, 2021

1 Content-based Movie Recommendation systems using Correlation

In the evergrowing world of movies it has become significantly harder to search through a large database of movies only to find a movie which you have wasted time watching because it was not to your own taste. Recommendation systems have now become ever-present after the large access to the internet from the masses, the idea is to make use of techniques such as filtering or matrixes to suggest similar or new items to a user in hope to keep them interest and coming back to the app, website, or program you have provided. For a media product that movies are companies like Netflix take very quick action in having the best recommendation system that is available even providing cash prizes for people who can get a more accurate system then what they have now. There are multiple systems used such as content-based, collaborative or hybrid systems however, it depends on what information has been presented before you decide on which system you would be working with. As myself i am an avid movie watcher decided to develop two simple content-based recommendation systems to evalutes the top 1000 movies on IMDb's website ranked by user rating descending, this was chosen as a base as I found this would be a good choice and easier to have user testing on as most people who actively watch movies would be able to judge if the movies recommended were a good choice or not. These two systems evaluate recommendations using a correlation type content-based system exploring a bag of words with Tf idf Vectorization and K-Nearest Neighbors checking for similar genres between movies. This aligns with the task provided and provides a solid base to expand upon in the future.