Movie Recommender memorandum

Introduction

The proposed project aims to build a movie recommender system using Python, leveraging the TMDB API to fetch movie metadata from 1980 until 2023. The TMDB (The Movie Database) is an online database that provides comprehensive information related to movies, TV shows, and other forms of visual media.

Context

Movie recommendation systems have gained popularity in recent years due to the overwhelming number of movies available for viewers to choose from. The rise of streaming services such as Netflix, Hulu, and Amazon Prime has made it difficult for viewers to decide which movie to watch. Thus, to address this issue, we have decided to develop a project that utilizes TMDB's API, which provides readily available movie metadata, as a solution for movie selection dilemmas.

Methodology

Movie metadata is fetched asynchronously from TMDB (The Movie DataBase) API, stored in a mySQL database, analyzed using multiple visualization techniques and finally is used to train a K-means based recommendation model that, given one movie, suggests five other movies that are similar to the input movie.

Results

We have successfully fetched metadata on more than 500,000 movies from 1980 until 2023 and stored it in a mySQL database. After, we have performed an in-depth exploratory analysis to get to know the fetched data and a pre-processing to prepare it to feed the recommendation model. Finally, a functional recommendation model has been obtained that provides the user with 5 movie suggestions.

Conclusions

Upon concluding this project, we can confidently state that we have gained an array of valuable skills. These skills include the ability to execute asynchronous API requests, implement advanced data visualization techniques, and utilize algorithms such as K-means for recommendation system development. Additionally, we have expanded our knowledge on the movie market's progression and status, spanning from 1980 until the present day. Ultimately, we have achieved our objective by deploying a fully operational model, capable of providing users with tailored movie recommendations based on their selected input.

Future work

To enhance this project in the future, a website could be developed that features a search bar enabling users to input a movie title. Upon searching, the website would display 5 clickable movie posters. When a user clicks on a poster, they would be redirected to a new page that displays the metadata of the selected movie. Also, tests should be done to ensure the correct functioning of the model.