Movie Recommender - Functional Specification

Group names: Priyanka Bijlani, Sharmeelee Bijlani, Laura Thriftwood, Lakshmi Venkatasubramanian

Background

When considering which movie to watch, users have access to an overwhelming number of options. Users want custom recommendations to ensure optimal use of their time watching content. Business models benefit from strong recommender systems by increasing user engagement and addiction to streaming platforms.

With this project, we can create our own movie recommendation system that takes user input of one movie and utilizes a rich dataset of movie titles, ratings and user information to output a recommended movie.

User Profile

Users

- Primary: Streamer or movie watcher with access to movie content and too many options to choose from.
- Secondary: Movie streaming platforms trying to increase user engagement by providing movie recommendations to users and accurately predicting movie ratings to curate the best content.

Background Knowledge

- Have a computing device (The recommender system is still a prototype and not yet on the web)
- Able to install Python and run python script (program does not yet have a front-end)
- Must have watched movies prior to using this system (input to the system)
- Secondary users should have understanding of recommendation system algorithms
- Secondary users need to provide historical data of movies, users, and ratings to train custom model with our program

Data Source

- MovieLens | GroupLens
    - Over 100k ratings
    - 1700+ movie titles
    - 1000+ users
- **Dataset Structure**
    - Ratings File: Each user has made at least 20 ratings.
        - Features:
            - UserID: integer ranging between 1 and 6040
            - MovieID: integer ranging between 1 and 3952
            - Rating: integer, 5-star scale (whole-star ratings only)

        - Timestamp: in seconds since the epoch as returned by time

- Movies File: The titles are sourced from IMDB and entered manually.
    - Features:
        - Title: string
        - Genres: nominal values
            - Action
            - Adventure
            - Animation
            - Children's
            - Comedy
            - Crime
            - Documentary
            - Drama
            - Fantasy
            - Film-Noir
            - Horror
            - Musical
            - Mystery
            - Romance
            - Sci-Fi
            - Thriller
            - War
            - Western

Use Cases

1. User will get a movie recommendation from the system based on their previous ratings
    a. Training input: users, movies, ratings
    b. User input: user name/id
    c. Outputs: movie(s)
    d. ML algorithm: Collaborative filtering (analyzes historical data)

2. User will be able to provide a movie name and get similar movies
    a. Training input: users, movies, ratings
    b. User input: movie
    c. Outputs: movie(s)
    d. ML algorithm: Collaborative filtering