



Movie Recommendation System

By: Dalya Lami, Ali Alam, Jahn Ferdinandus, Dayana Imanova

Project Objective:

We are focused on developing a movie recommendation model that incorporates user ratings to provide personalized recommendations. The project involves exploring and analyzing a dataset of user ratings and movie information, which will allow us to better understand individual preferences and provide tailored recommendations that align with each user's unique taste and interests. As a learning process we utilized a content and collaborative based filtering to recommend new movies based on the movie key word.

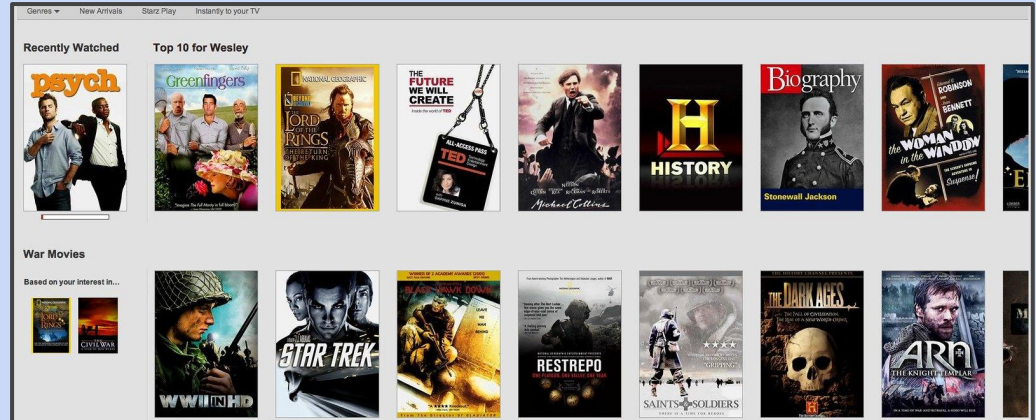


Table of Content:

Theory

1. Underlying Theory
2. Content-based Filtering
3. Collaborative Filtering

Overview of the data

Examine the dataset using Tableau's graphics and Python libraries.

movielens

Models

1. Content-based & Collaborative Models used to get similar movie recommendations
2. SVD Model used to predict ratings and get the top 5 movies with highest predicted ratings

Theory

Underlying Theory

The underlying theory is that movie with more user ratings and greater popularity tend to be enjoyed by a wider audience.

Content-based

Content-based model uses metadata to suggest similar movies based on filtering specific criteria, in our case “Movie Title”. The model will suggest movies with titles that are similar to the input movie. (fig 1)

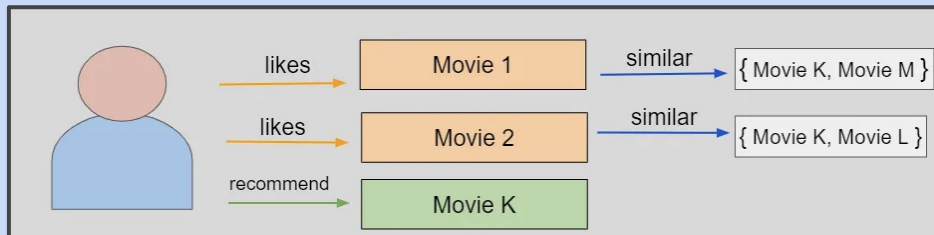


Fig 1

Collaborative Filtering

The system's collaborative filtering is derived from movie ratings by computing an overall score by “similar users” and “all users” for each movie. (fig 2)

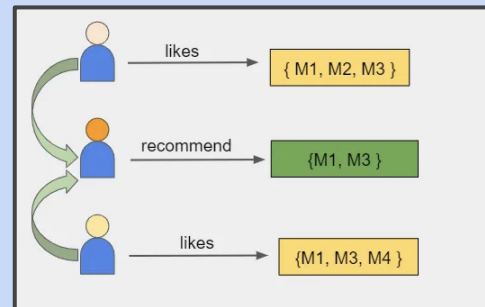


Fig 2

Learning Process

Content-based

Movie Title:

	movieid	title		genres	clean_title
20827	107630	High School (2010)		Comedy	High School 2010
61563	203282	The School (2018)	Horror Mystery Thriller		The School 2018
48367	175305	School Life (2016)		Documentary	School Life 2016
56460	192393	Night School (2018)		Comedy	Night School 2018
46276	170953	Night School (2016)		Documentary	Night School 2016

By entering a movie title, the system utilizes content-based filtering to recommend movies based on similar titles.

Collaborative-based

Movie Title:

	score	title		genres
20827	48477.6	High School (2010)		Comedy
60889	48477.6	Unless (2016)		Drama
41364	48477.6	Shocking Blue (2010)		Drama
60950	48477.6	Love Is Thicker Than Water (2017)	Comedy Drama Romance	
39141	24238.8	The Daughter (2016)		Drama
49288	24238.8	Pink Floyd: The Story of Wish You Were Here (2...		Documentary
9908	24238.8	Between Your Legs (Entre las piernas) (1999)	Drama Mystery Romance Thriller	
56115	24238.8	Shot! The Psycho-Spiritual Mantra of Rock (2016)		Documentary

By entering a movie title and leveraging collaborative filtering, the system utilizes ratings to assign an overall score and provides quick suggestions for additional movies we might like.

SVD Modeling

Singular Value Decomposition

Strengths of SVD in Recommendations:

Effective Prediction & Matrix Factorization:

- Proficient in accurately predicting user preferences and decomposing the user-item matrix using factorization.

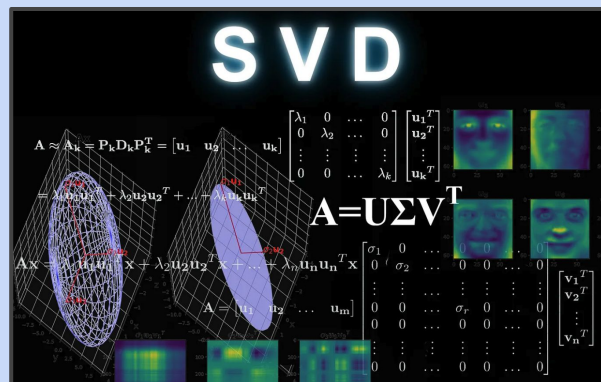
Collaborative Filtering Basis:

- Forms the foundation for collaborative filtering algorithms, enhancing collaborative recommendations.

Limitations of SVD:

Scalability Issues:

- Challenges with large datasets may impact computational efficiency.



Breakdown of what the code does

User Input Retrieving & Filtering:

- Retrieves the user ID for whom recommendations are to be generated
- Filters out movies that the user has already rated, preparing a list of potential recommendations.

Generate Movie Recommendations:

- It employs a collaborative filtering algorithm, utilizing the Surprise library, to predict ratings for unrated movies for the specified user.
- Predicts the user's ratings for each unrated movie, and a list of tuples containing movie IDs and their estimated ratings is created.
- It sorts and displays the top 5 recommended movies, including their titles and estimated ratings.

User ID:

Recommend

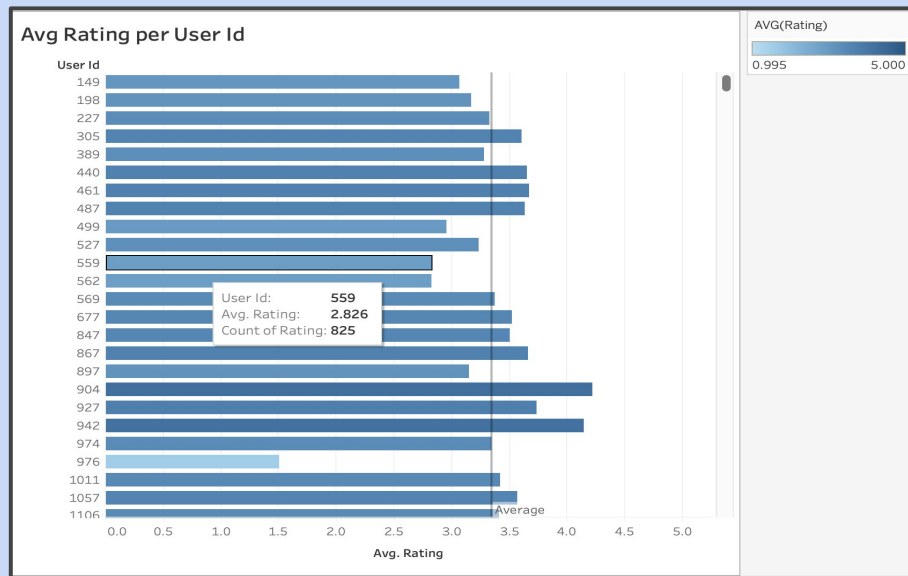
Top 5 Most Recently Watched Movies by User:

Requiem for a Dream (2000), Rating: 5.0
Willy Wonka & the Chocolate Factory (1971), Rating: 4.5
Mrs. Doubtfire (1993), Rating: 4.0
Dumb & Dumber (Dumb and Dumber) (1994), Rating: 0.5
Godfather, The (1972), Rating: 4.5

Top 5 Recommended Movies:

1. Monty Python and the Holy Grail (1975), Estimated Rating: 4.35426718845973
2. Grand Day Out with Wallace and Gromit, A (1989), Estimated Rating: 4.35267104554605
3. Old Boy (2003), Estimated Rating: 4.291601802345809
4. Godfather: Part II, The (1974), Estimated Rating: 4.2788868972254495
5. City of God (Cidade de Deus) (2002), Estimated Rating: 4.230013326009706

Visual Analysis

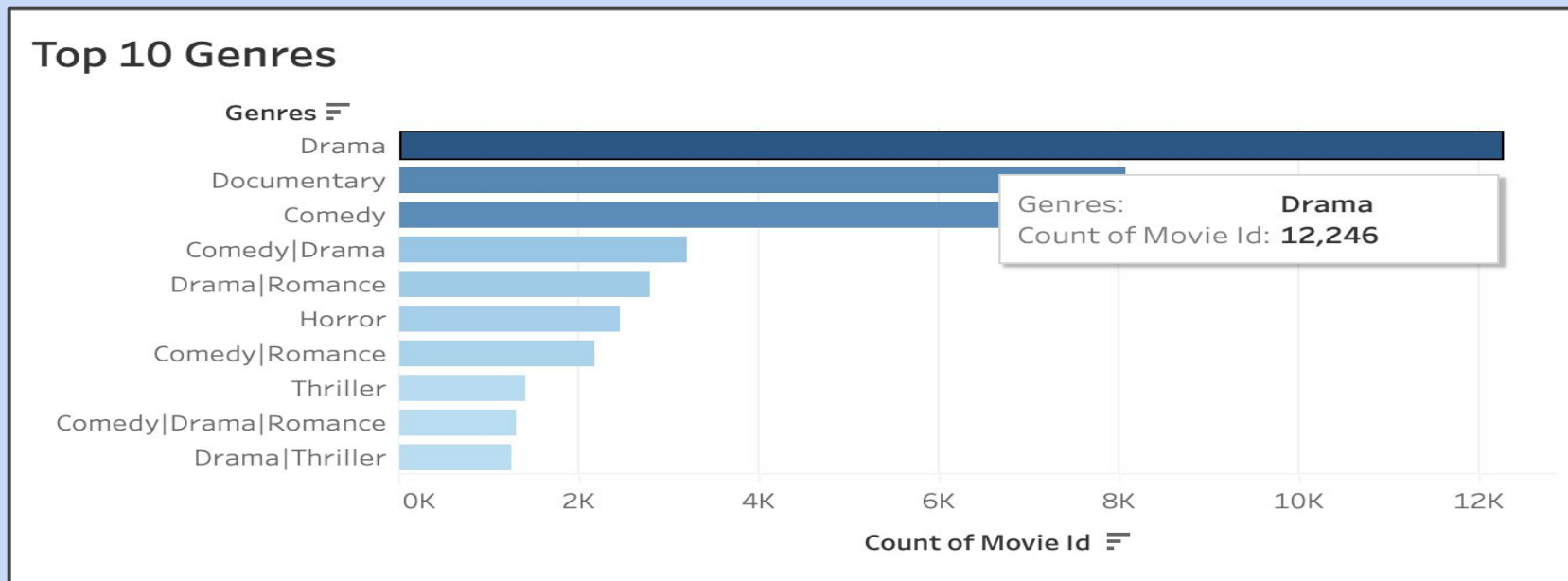


This image shows each user with their avg rating total for all the movies, while also showing the total number of movies they've rated. The vertical line shows the average total for all the user ratings



This image shows each movie with it's avg rating total, while also showing the total number of users that have rated the movie.

Visual Analysis

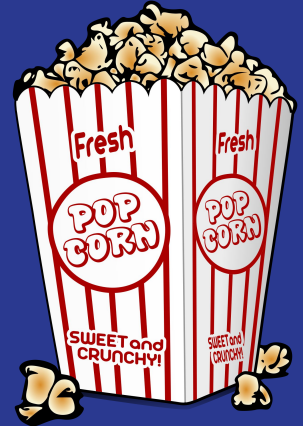


Top ten genres with the genres title and the total number of movies in that genres.

Demonstration

Movie Recommendation Systems

Now we are going to go over our code in Jupyter Notebook and we will demonstrate the type of results we can get.





Q & A