



Recommender Systems Project:
Movie recommendation
System

Group Members:

- i) Bilal Ahmed Khan (20k-0183)
- ii) Muaaz Alam (20k-0212)
- iii) Muhammad Shayan (20k-0494)
- iv) Syed Talha Ali (20k-1621)

Introduction:

In this project, we have developed a movie recommender system that employs four different algorithms to provide personalized movie recommendations to users. The algorithms we have implemented include User-based Collaborative Filtering (CF), Item-based CF, Content-based filtering, and Naïve CF.

User-based CF relies on the similarity between users' preferences to recommend movies. By analyzing the viewing history and ratings of similar users, the system can suggest movies that one user might enjoy based on the preferences of users with similar tastes.

Content-based filtering takes a different approach, considering the features and attributes of movies to make recommendations. By analyzing factors(features) such as genre, actors, directors, and plot keywords, the system suggests movies that share similar characteristics with the ones a user has liked previously.

Lastly, the Naïve CF algorithm provides recommendations based on simple rules, such as popularity or average ratings. While it may not be as sophisticated as the other algorithms, it serves as a baseline for comparison and can sometimes produce surprisingly accurate recommendations.

Techniques implemented:

- i) User based CF
- ii) Content-based Filtering
- iii) Naïve based CF

Workflow:

- **Data Collection:** The first step in building a movie recommender system is to gather the necessary data. This can include movie metadata such as titles, genres, actors, directors, and user ratings. This data can be obtained from various sources, including online movie databases, user reviews, and previous user interactions within the system.
- **Data Pre-processing:** Once the data is collected, it needs to be processed and prepared for further analysis. This involves cleaning the data, handling missing values, and transforming it into a suitable format for the algorithms. For example, NaN values need to be removed, and user ratings may need to be normalized.
- **Algorithm Implementation:** The next step is to implement the four algorithms: User-based CF, Item-based CF, Content-based CF, and Naïve CF. Each algorithm has its own specific

logic and calculations. User-based CF involves finding similar users based on their rating patterns and recommending movies liked by similar users. Item-based CF focuses on finding similar movies based on their ratings and suggesting movies like the ones a user has enjoyed. Content-based CF uses movie features and attributes to recommend movies with similar characteristics. Naïve CF employs simple rules such as popularity or average ratings to make recommendations.

- **User Interface Development:** To make the movie recommender system user-friendly, a user interface needs to be developed. Tkinter library is used to make a user-friendly interface.

GUI:

Movie Recommender System

Select Dataset

Selected file: C:/Users/Bilal/Desktop/RS Project/ml-latest-small/movies.csv Select Dataset

Sample Data

```
movieId, title, genres
1, Toy Story (1995), Adventure|Animation|Children|Comedy|Fantasy
2, Jumanji (1995), Adventure|Children|Fantasy
3, Grumpier Old Men (1995), Comedy|Romance
4, Waiting to Exhale (1995), Comedy|Drama|Romance
5, Father of the Bride Part II (1995), Comedy
6, Heat (1995), Action|Crime|Thriller
7, Sabrina (1995), Comedy|Romance
8, Tom and Huck (1995), Adventure|Children
9, Sudden Death (1995), Action
```

Collaborative filtering

Tom and Huck (1995) Submit

```
Once Were Warriors (1994)
Wild Bunch, The (1969)
Othello (1995)
Dangerous Minds (1995)
Hellraiser: Bloodline (1996)
```

Updated changes:

User based Collaborative filtering:

Select Dataset

Selected file: C:/Users/Bilal/Desktop/RS Project/ml-latest-small/ratings.csv

Select Dataset

Sample Data

```
userId, movieId, rating, timestamp
1, 1, 4.0, 964982703
1, 3, 4.0, 964981247
1, 6, 4.0, 964982224
1, 47, 5.0, 964983815
1, 50, 5.0, 964982931
1, 70, 3.0, 964982400
1, 101, 5.0, 964980868
1, 110, 4.0, 964982176
1, 151, 5.0, 964984041
```

Collaborative filtering

95

Submit

```
Star Wars: Episode IV - A New Hope (1977)
Matrix, The (1999)
Star Wars: Episode V - The Empire Strikes Back (1980)
Pulp Fiction (1994)
Raiders of the Lost Ark (Indiana Jones and the Raiders of the Lost Ark) (1981)
```

User based Naïve bayes filtering:

Naive based filtering

64

Submit

```
Andrew Dice Clay: Dice Rules (1991)
Robot Carnival (Roboto kânibauru) (1987)
Doctor, The (1991)
Black Narcissus (1947)
Yards, The (2000)
```

Content Based Filtering:

Content based filtering

Toy Story (1995)

Submit

```
Turbo (2013)
Monsters, Inc. (2001)
Moana (2016)
Emperor's New Groove, The (2000)
Adventures of Rocky and Bullwinkle, The (2000)
```