MOVIE RECOMMENDATION

GROUP MEMBERS

SHIFA KHAN- 13

SHRUTI PARAVE - 17

INTRODUCTION

In today's digital landscape, movie recommendation systems have become an essential tool for streaming platforms, helping users discover films tailored to their preferences. With an overwhelming number of choices available, a well-structured recommendation system enhances user experience by suggesting movies based on factors such as viewing history, genre preferences, ratings, and even social interactions.

PROJECT GOALS

- Evaluate Recommendation Models – Analyze the performance of collaborative filtering, content-based filtering, and hybrid models in predicting user preferences.

- Improve User Experience – Enhance the accuracy of movie suggestions to make personalized recommendations more relevant and engaging.

DATA COLLECTION

- The data was collected from publicly available sources like:
  - Movie Databases
  - User Interaction Data
  - Content Metadata

- The data was stored in CSV (Comma Separated Values) format, which was later loaded into Python Jupyter Notebook/Colab for analysis.

DATA CLEANING AND PREPROCESSING

To ensure accurate and efficient recommendations, data cleaning and preprocessing are essential steps.

1. Handling Missing Data
 • Movie Metadata: Fill in missing values for attributes like genre, director, and cast using external sources (IMDb, TMDb).

2. Removing Duplicates and Inconsistencies
 • Identify and eliminate duplicate entries for movies or users to prevent redundancy.

3. Data Normalization and Transformation
 • Ratings Scaling: Convert ratings into a uniform scale (e.g., 1-5 or 0-1) to standardize differen rating systems.

EXPLORATORY DATA ANALYSIS (EDA)

EDA is a crucial step in understanding the dataset, identifying patterns, and gaining insights before building a movie recommendation model.

1. Understanding the Dataset
  • Checking Data Shape:
  • Number of movies, users, and interactions.

2. Visualizing Ratings Distribution
  • Histogram of Movie Ratings:
  • Helps understand how users rate movies (e.g., are ratings skewed towards higher/lower values?).

3. Movie Popularity & Trends
  • Top Rated Movies:
  • List of movies with the highest average ratings.

4. User Behavior Analysis
  • Number of Ratings per User:
  • Helps segment users based on engagement levels (e.g., occasional vs. frequent raters).

THANK YOU

ANY QUESTIONS?