

BrainDead

REELSENSE (PS:01)

Submitted By

Dhanashri A. Shah

An Explainable Hybrid Movie Recommendation System
Delivering personalized Top-K recommendations with
diversity, novelty, and natural language explanations.



PROBLEM STATEMENT

01 Rating Bias

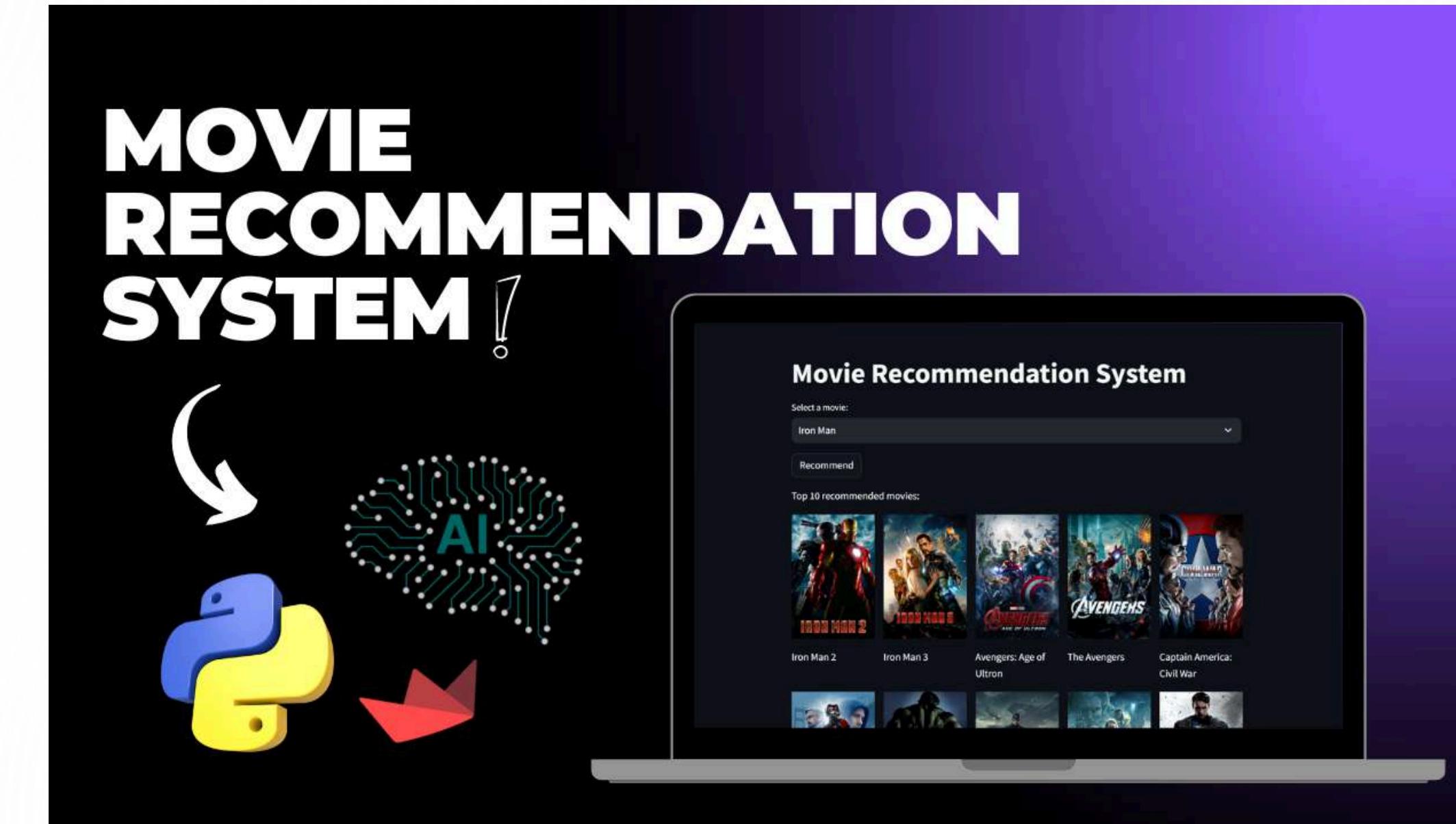
Traditional recommendation systems focus mostly on predicting user ratings, which can lead to repetitive or popular-only suggestions.

02 Explainable Recommendations

Users often don't know why a movie is suggested; providing clear, natural language explanations increases trust and engagement.

03 Diverse Suggestions

Recommendations should include novelty and variety, avoiding popularity bias while covering different genres and user interests



ReelSense is a machine learning project that bridges these gaps by delivering personalized, diverse, and explainable movie recommendations.

DATASET OVERVIEW

In this competition, we have provided some datasets as follows

01

Movies

Contains information about each movie, used for content-based feature modeling

Key Columns: movieId, title, genres

02

Ratings

Records of user–movie interactions, used for collaborative filtering and analyzing user preferences

Key Columns: userId, movieId, rating, timestamp

03

Tags

User-generated tags providing semantic context for movies, useful for explainable recommendations.

Key Columns: userId, movieId, tag, timestamp

04

Links

Connects movies to external databases like IMDb and TMDb, ensuring compatibility for external data enrichment.

Key Columns: movieId, imdbId, tmdbId

DATA CLEANING & PIPELINE

Most of the data was clean, and no need to do anything. There were just 8 null values. I dropped all, then I merged the two tables, as the rating table and the movies table

Why to join ?

Valid & Complete Data

Ensures every rating corresponds to a movie with complete metadata (title, genres).

Consistency & Cleanliness

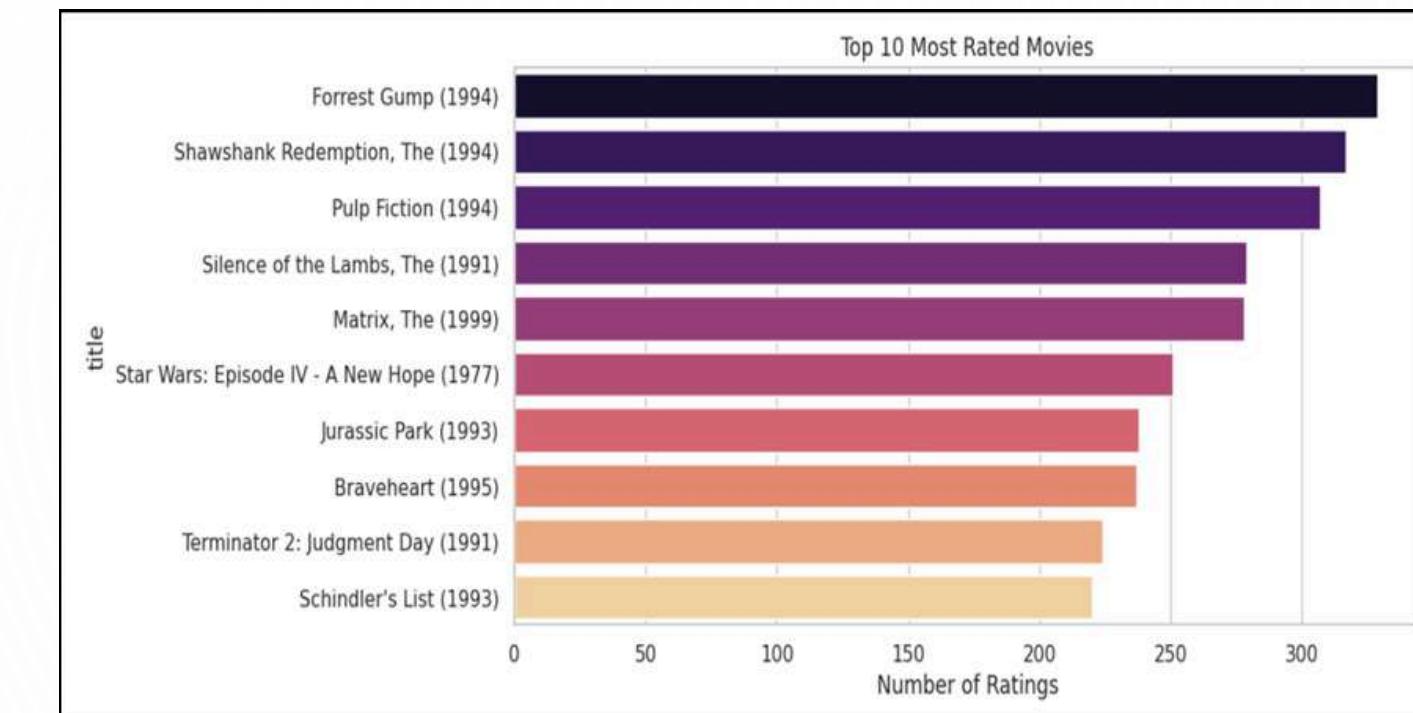
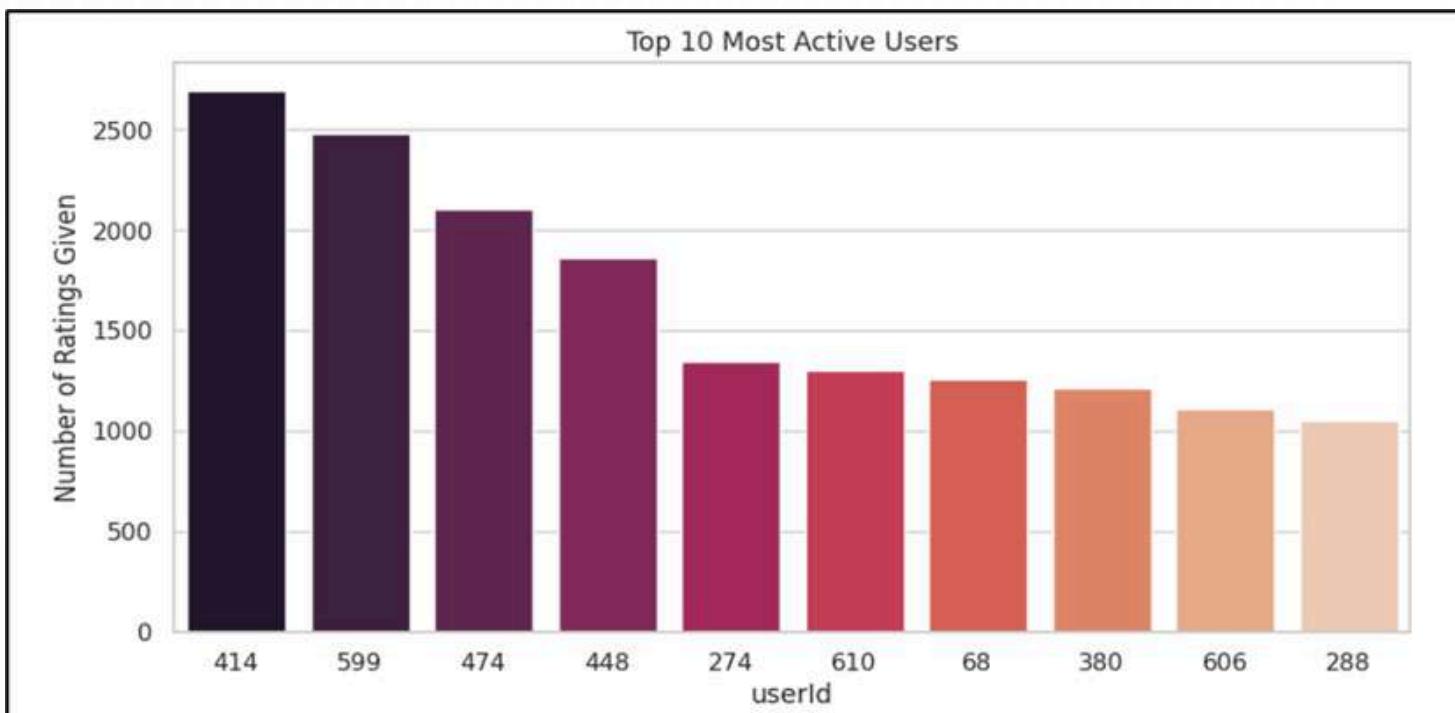
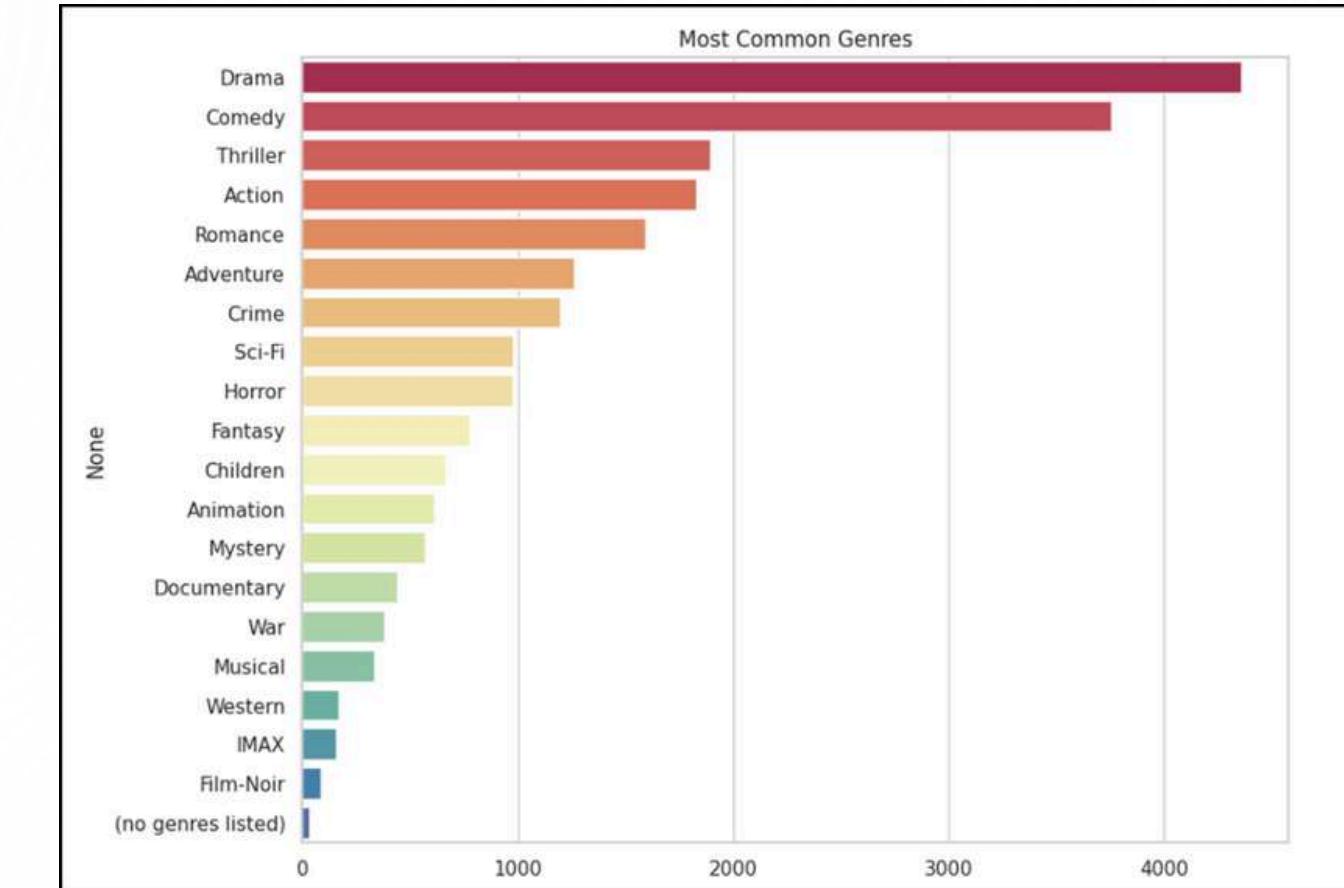
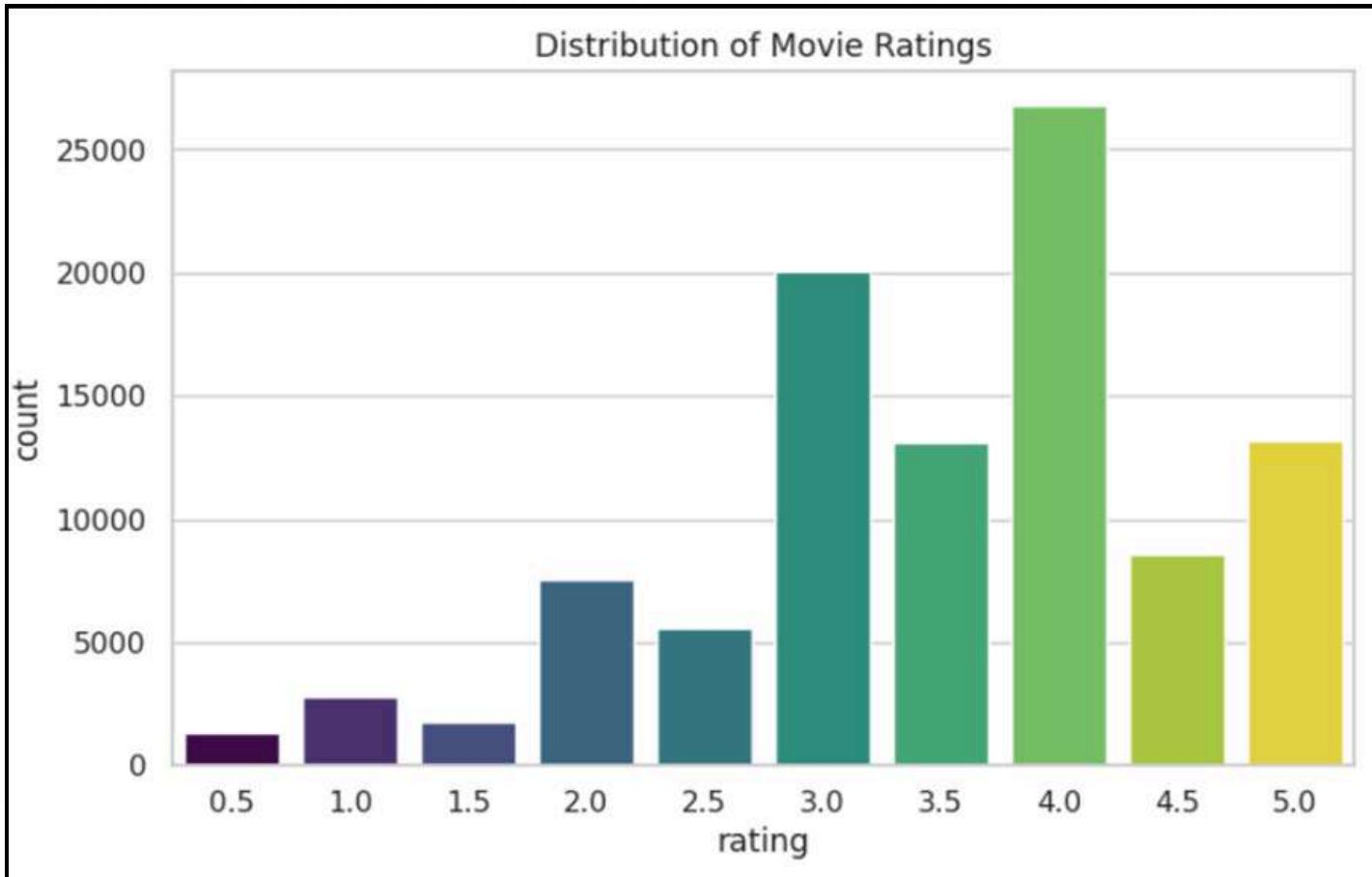
Removes incomplete or inconsistent records, maintaining reliable data for modeling.

Error Prevention

Avoids runtime errors during feature extraction, re-ranking, and evaluation by eliminating missing values.



EXPLORATORY DATA ANALYSIS (EDA)



BrainDead

LEARNING USER PREFERENCES

Focus: Learning user preferences and item representations for recommendations.

Tools: Python, NumPy, Pandas, Scikit-learn, Google Colab.

Methods:

Matrix Factorisation = Collaborative Filtering

Genre-based features = Content-Based Modelling

Procedure:

Construct a user–movie interaction matrix.

Train on training split, evaluate on test split.

Default hyperparameters; focus on hybrid approach reproducibility.





ASSESSING RECOMMENDATION QUALITY

Rating Prediction Metrics: RMSE, MAE (evaluate prediction accuracy).

Top-K Ranking Metrics: Precision@K, Recall@K, NDCG@K, MAP@K (evaluate recommendation ranking).

Comparison: Hybrid model vs. baseline collaborative filtering. Focus on personalisation, diversity, and explainability.

HYBRID MODEL

INSIGHTS

- Generates personalised and diverse recommendations.
- Trade-off: Accuracy vs. diversity; hybrid improves novelty without losing personalisation.
- Works effectively for individual users and multiple users.
- Content-based component enhances explainability.

Top-K Performance (K=10)

Average Precision@10: 10% → ~1 relevant movie in top 10.
 Average Recall@10: 9.32% → small portion of relevant movies retrieved.
 Average Genre Coverage@10: 15 → diverse genres per user.

Rating Prediction Accuracy

RMSE: 5.582 → deviation between predicted and actual ratings.
 MAE: 1.981 → average absolute difference.
 High RMSE expected due to dataset sparsity.

Diversity & Novelty Metrics

NDCG@10: 0.0 → relevant movies not highly ranked.
 MAP@10: 0.0 → top recommendations rarely match all relevant items.
 Intra-List Diversity: 0.749 → recommendations cover a wide range of genres.
 Novelty: 12.70 → less popular movies recommended, reducing popularity bias.

BrainDead

WORKING DEMO

When I Input clod user, it displayed recommendations for him as you can see in the table below.

Project link : [dhanashrishah1306-svg/Hackathon_-Brain_Dead-](#)

title	genres	Reason
Enchanted (2007)	Adventure Animation Children Comedy Fantasy Mu...	Matches your interest in Adventure, Animation,...
Turbo (2013)	Adventure Animation Children Comedy Fantasy	Matches your interest in Adventure, Animation,...
Space Jam (1996)	Adventure Animation Children Comedy Fantasy Sci...	Matches your interest in Adventure, Animation,...
Yellow Submarine (1968)	Adventure Animation Comedy Fantasy Musical	Matches your interest in Adventure, Animation,...
Futurama: Bender's Game (2008)	Action Adventure Animation Comedy Fantasy Sci-Fi	Matches your interest in Adventure, Animation,...
Robots (2005)	Adventure Animation Children Comedy Fantasy Sci...	Matches your interest in Adventure, Animation,...
Home (2015)	Adventure Animation Children Comedy Fantasy Sci...	Matches your interest in Adventure, Animation,...
Tale of Despereaux, The (2008)	Adventure Animation Children Comedy Fantasy	Matches your interest in Adventure, Animation,...
Monsters, Inc. (2001)	Adventure Animation Children Comedy Fantasy	Matches your interest in Adventure, Animation,...
Shrek the Halls (2007)	Adventure Animation Comedy Fantasy	Matches your interest in Adventure, Animation,...

LIMITATIONS

The proposed hybrid recommendation system, while effective, has several limitations. The collaborative filtering component struggles with the cold-start problem, making recommendations for new users or new movies less reliable. The system also relies heavily on explicit rating data, which may not fully capture true user preferences, and genre-based features can oversimplify complex tastes by ignoring context, mood, or temporal factors. Additionally, all evaluations were conducted offline, so real-world user interactions and feedback may differ from the reported results.

**01****Cold-start problem**

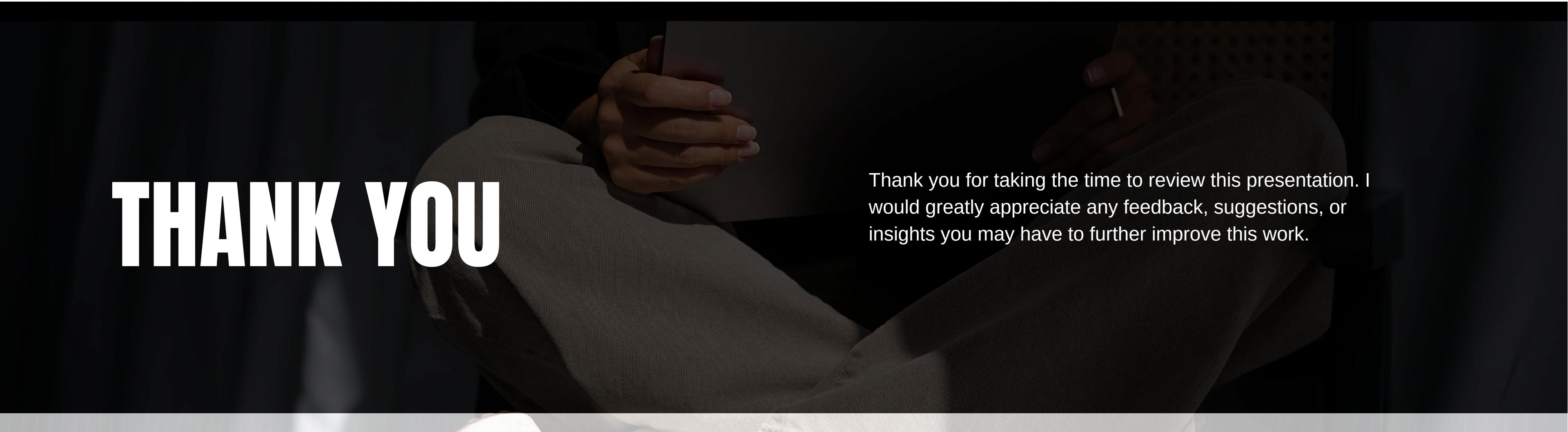
new users/movies may have unreliable recommendations.

02**Heavy reliance on explicit ratings**

genre-based features may oversimplify user preferences.

03**Offline evaluation**

may not fully reflect real-world interactions



THANK YOU

Thank you for taking the time to review this presentation. I would greatly appreciate any feedback, suggestions, or insights you may have to further improve this work.



dhanashrishah1306@gmail.com