🌟 Tamizhan Skills SE RISE Internship – Machine Learning & AI

📝 Task 5: Movie Recommendation System

👤 Name: Harshita
🏫 College: CRSSIET, Jhajjar
🎓 B.Tech CSE (AI & ML), 2023–2027

📄 Project Summary

🧩 Objective:
To build a movie recommendation system that suggests similar movies to users based on their preferences using item-based collaborative filtering and cosine similarity.

📁 Dataset Used:
used the open-source MovieLens + TMDB dataset from Kaggle:
🔗 https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset
Files Used:

- ratings_small.csv
- movies_metadata.csv

🔧 Tools & Technologies:

- Google Colab (Python)
- pandas, numpy, sklearn
- Cosine similarity (item-based filtering)

🔍 Methodology (Steps Followed):

1. Data Preprocessing:
Loaded ratings_small.csv and movies_metadata.csv
Cleaned and converted movieId and id to match format
Merged both datasets to connect movie ratings with titles

2. User-Movie Matrix Creation:

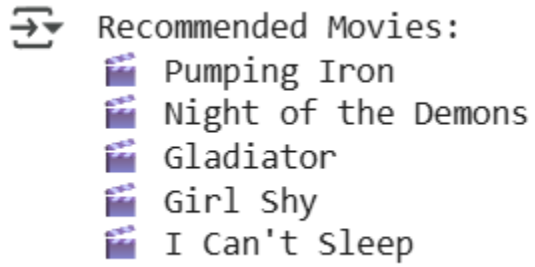
Created a pivot table where rows = users, columns = movie titles, values = ratings
Filled missing ratings with 0

3. Cosine Similarity Calculation:
Used cosine_similarity() from sklearn.metrics.pairwise to build a similarity matrix between all movie titles

4. Recommendation Function:
Created a function to fetch top 5 similar movies to a given input title
Implemented fuzzy matching (difflib.get_close_matches) to handle partial or incorrect input titles

5. Result Generation:
Successfully generated and displayed recommendations for example movie "Gator Bait" in Colab

🖼️ Sample Output Screenshot:



✅ Conclusion:
This project successfully demonstrates the use of collaborative filtering using cosine similarity to recommend movies. It simulates how platforms like Netflix or Prime Video suggest content based on previous preferences.