Movie Recommendation System

Project Overview

This project implements a hybrid movie recommender system combining:

- SVD (Singular Value Decomposition) based collaborative filtering
- Item-based collaborative filtering (CF) using cosine similarity

The system provides:

- Personalized movie recommendations for users based on their past ratings
- Related movie suggestions based on item similarity

Data

- Dataset: MovieLens 100k
- Merged ratings with movie metadata (title, genres, IMDb URL, release date)
- Preprocessing steps:
  - Handled duplicate ratings by the same user for the same movie (took average rating)
  - Converted timestamps to datetime for analysis

Methodologies

Item-Based Collaborative Filtering

- Built user-item rating matrix
- Computed cosine similarity between movies
- For a selected movie, recommended top N similar movies
- Recommended movies were filtered to ensure they are reasonably well-rated (e.g. average rating ≥ 3)

SVD (Singular Value Decomposition)

- Used scikit-surprise SVD implementation
- Tuned hyperparameters using GridSearchCV
- Best RMSE achieved: 0.9269
- Best parameters: n_factors=150, reg_all=0.1, lr_all=0.01
- Model predicts unseen ratings for users

Hybrid Approach

- Recommend top N movies for a user using SVD
- For each selected movie, recommend top N similar movies using item-based CF

Results

- SVD RMSE after tuning: 0.9269
- Item CF worked well for similarity, easy to compute and interpret
- Hybrid model provides both personalized user suggestions (SVD) and related movie discovery (item CF)

Limitations

- scikit-surprise dependency makes cloud deployment (like Streamlit Cloud) challenging
- Cold-start problem: system struggles with new users or movies with no ratings
- CF similarity can be influenced by sparse rating data

Future Enhancements

- Add content-based filtering using genres, tags, or descriptions
- Explore deep learning recommenders like neural collaborative filtering or autoencoders
- Improve deployment (e.g. using Docker to support SVD models)