Movie Recommendation System

Objective

The goal of this project was to build a Movie Recommendation System using the MovieLens dataset (100K). The system should suggest the Top-N movies for a given user, evaluated with metrics such as Precision@K, Recall@K, and NDCG@K.

I have implemented and compared two main approaches:

1. **User-Based Collaborative Filtering (CF)**
2. **Matrix Factorization using Singular Value Decomposition (SVD)**

Methodology

1. **User-Based Collaborative Filtering**
    - Built a user–user similarity matrix using cosine similarity.
    - The top 5 movies were recommended for the target user.
2. **Matrix Factorization (SVD)**
    - Predicted missing ratings by reconstructing the matrix.
    - The top 5 movies with the highest predicted ratings were recommended.
3. **Evaluation Metrics**
    - Precision@K
    - Recall@K
    - NDCG@K

Result

User-based:
Precision@5 = 0.234
Recall@5 = 0.101
NDCG@5 = 0.246

SVD (Matrix Factorization):
Precision@5 = 0.356
Recall@5 = 0.126
NDCG@5 = 0.385

This clearly shows that SVD worked better than the user-based method. Here's why:

1. **Captures latent factors:**
   Instead of relying only on direct user-user similarity (which is sparse in MovieLens 100k), SVD compresses the user-item rating matrix into latent features that represent hidden preferences.

2. **Better generalization:**
   User-based CF struggles when two users don't share enough rated movies. SVD generalizes better by projecting all users and movies into a shared embedding space.

3. **Handles sparsity:**
   The dataset is sparse. SVD reduces noise and sparsity by approximating the rating matrix.

Deployment

The model is deployed on Hugging Face Spaces using Gradio. It allows users to input a User ID and select between User-based Collaborative Filtering or SVD (Matrix Factorization) to get top-5 movie recommendations.