Movie Recommender System

A Mahesh Reddy

[Unified Mentor]

1. Introduction

Recommendation systems have become an essential tool for personalized content delivery in various fields such as e-commerce, social media, and entertainment. This project aims to develop a movie recommendation system using collaborative filtering techniques. By leveraging user preferences and ratings, the system can suggest movies that align with users' tastes, improving user engagement and satisfaction.

The goal of this project is to provide personalized movie recommendations based on user behavior and preferences.

2. Dataset Description

The dataset used for this project includes the following features:

- **Movie ID**: A unique identifier for each movie.
- **Title**: The title of the movie.
- **Genres**: Categories associated with the movie.
- **User ID**: Unique identifier for each user.
- **Rating**: The rating given by a user to a movie (typically on a scale of 1 to 5).
- **Timestamp**: Time at which the rating was provided.

The dataset is divided into training and testing sets to evaluate the recommendation algorithm's performance.

3. Problem Statement

The task is to build a movie recommendation system that suggests movies based on past user interactions (ratings). The system should address the following challenges:

- **Cold Start Problem**: How to recommend movies to users with few or no ratings.
- **Scalability**: Efficiently handling large datasets with millions of users and movies.
- **Accuracy**: Making recommendations that align closely with user preferences.

4. Methodology

The recommendation system is built using **Collaborative Filtering**, which focuses on user-item interaction data. The two main techniques used are:

- **User-Based Collaborative Filtering**: Recommends movies based on the preferences of similar users. Users who have similar movie preferences are identified, and recommendations are made based on the movies that similar users have liked.
- **Item-Based Collaborative Filtering**: Recommends movies based on the similarity between items (movies). Movies that are similar to those a user has rated highly are recommended.

Outputs:

```
pop= df2.sort_values('popularity', ascending=False)
import matplotlib.pyplot as plt
plt.figure(figsize=(12,4))

plt.barh(pop['title'].head(6),pop['popularity'].head(6), align='center',
        color='skyblue')
plt.gca().invert_yaxis()
plt.xlabel("Popularity")
plt.title("Popular Movies")
```

```
Text(0.5, 1.0, 'Popular Movies')
```