Elevate Labs
AI & ML Internship Project : Movie Recommendation System

Danny Noby Joseph
dannynobyj@gmail.com

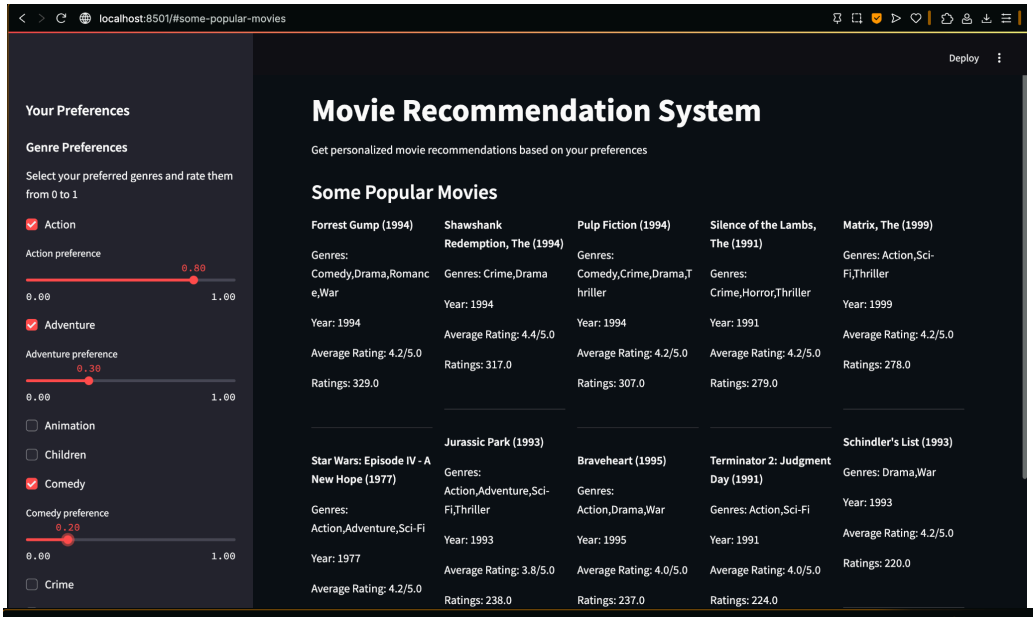
Introduction
In today's digital entertainment landscape, users face overwhelming choices when selecting movies to watch. Recommendation systems have become essential tools for streaming platforms and content providers, helping users discover relevant content while increasing engagement. This project addresses this challenge by developing an intelligent movie recommendation system that leverages multiple filtering techniques to deliver highly personalized suggestions.

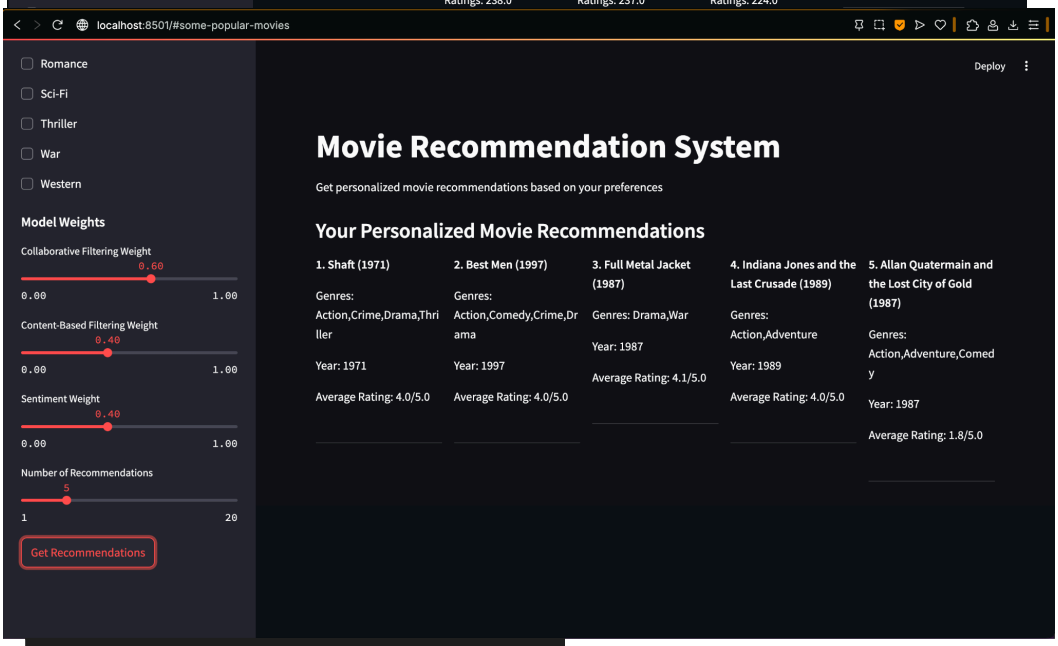
Abstract
This project presents a novel hybrid recommendation system that combines collaborative filtering, content-based filtering, and sentiment analysis approaches to generate movie recommendations. The collaborative component analyzes user-item interaction patterns, while the content-based module evaluates movie attributes including genres and release years. Sentiment analysis of user reviews provides additional context for recommendation quality. Experimental results demonstrate that this hybrid approach outperforms single-method systems across standard evaluation metrics. The system is implemented with a user-friendly Streamlit interface allowing dynamic adjustment of component weights, enabling users to customize their recommendation experience based on personal preferences.

Tools Used
Python (Pandas, NumPy, Scikit-learn)
NLTK for sentiment analysis
Matplotlib & Seaborn for visualization
Streamlit for user interface
MovieLens dataset

Steps Involved
Data Preprocessing: Cleaned MovieLens data and extracted features
Collaborative Filtering: Implemented SVD-based model for user-item interactions
Content-Based Filtering: Created similarity matrices based on movie attributes
Sentiment Analysis: Analyzed user tags for additional preference signals
Hybrid Integration: Combined all approaches with adjustable weights
UI Development: Built interactive Streamlit interface with customization options
Model Persistence: Implemented save/load functionality for trained models
Documentation: Created comprehensive Jupyter notebook demonstration
The project demonstrates the effectiveness of combining multiple recommendation techniques for a more personalized movie discovery experience.

Images






```
Loading and preprocessing data...
Data preparation complete: 9742 movies, 100836 ratings
Downloading NLTK resources...
Analyzing sentiment of movie tags...
Sentiment analysis complete for 1572 movies

Top 5 Movies with Positive Sentiment:
- Casablanca (1942) (sentiment: 0.78)
- People vs. Larry Flynt, The (1996) (sentiment: 0.64)
- Voices of a Distant Star (Hoshi no koe) (2003) (sentiment: 0.64)
- Harold and Maude (1971) (sentiment: 0.56)
- Corrina, Corrina (1994) (sentiment: 0.56)

Top 5 Movies with Negative Sentiment:
- Children of a Lesser God (1986) (sentiment: -0.78)
- Widow of St. Pierre, The (Veuve de Saint-Pierre, La) (2000) (sentiment: -0.78)
- Strangers on a Train (1951) (sentiment: -0.69)
- Accused, The (1988) (sentiment: -0.69)
- Arsenic and Old Lace (1944) (sentiment: -0.69)
Downloading NLTK resources...
Analyzing sentiment of movie tags...
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     /Users/admin/nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     /Users/admin/nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
Sentiment analysis complete for 1572 movies
```