MOVIE RECOMMENDATION SYSTEM – FINAL PROJECT REPORT

INTRODUCTION

Recommendation systems are essential tools for modern content platforms, helping users discover content aligned with their preferences. This project focuses on developing a hybrid movie recommendation system that combines user behavior (collaborative filtering) and movie content (genre-based filtering). A user-friendly interface was built using Streamlit to allow real-time interaction.

ABSTRACT

Using the MovieLens dataset, this system delivers accurate movie recommendations based on two approaches:

- **Content-Based Filtering**: Uses genres and cosine similarity to suggest similar movies.
- **Collaborative Filtering**: Leverages user rating behavior and a K-Nearest Neighbors model built on a sparse matrix.

An optional sentiment analysis filter (TextBlob) refines results to suggest only positively perceived movie titles. The result is a fast, responsive web app where users input a movie they like and get back top recommendations.

TOOLS USED

- Python 3.10+
- Streamlit
- Pandas, NumPy
- Scikit-learn, SciPy
- TextBlob (optional sentiment analysis)

STEPS INVOLVED IN BUILDING THE PROJECT

1. **Dataset Preprocessing**:
   - Imported movies.csv and ratings.csv from MovieLens.
   - Filtered active users (rated >50 movies) and well-rated movies (>100 ratings).
2. **Content-Based Filtering**:
   - Vectorized genres with CountVectorizer.
   - Used cosine distance to identify similar movies.
3. **Collaborative Filtering**:
   - Created a sparse matrix of user-item ratings.
   - Used NearestNeighbors to recommend based on similar rating behavior.
4. **Sentiment Filtering (Optional)**:
   - Used TextBlob to analyze sentiment of movie titles.
   - Filtered out movies with negative polarity scores.
5. **Streamlit Interface**:
   - Allowed users to choose a movie and filtering method.
   - Displayed top 10 movie recommendations with or without sentiment filtering.

CONCLUSION

The Movie Recommendation System successfully demonstrates how machine learning can enhance entertainment discovery. The hybrid model improves accuracy by merging behavioral and content insights. With an intuitive UI and optional sentiment filtering, this system balances performance and usability—offering a real-world example of intelligent automation for user satisfaction.
