# Movie Recommendation System – Project Report

## 1. Project Title

Movie Recommendation System using Machine Learning

## 2. Objective

The objective of this project is to develop a Movie Recommendation System that suggests top 5 movies to users based on their preferences. The system leverages both content-based filtering (using genres) and collaborative filtering (using user ratings) to generate relevant and personalized movie suggestions.

## 3. Dataset

We used the MovieLens dataset available from Kaggle:  
https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset

Key files used:

- movies\_metadata.csv – contains movie details (title, genres, etc.)

- ratings\_small.csv – user ratings for movies

## 4. Data Preprocessing

1. Loading Datasets: Used pandas to load CSV files.  
2. Cleaning Issues:  
 - Fixed corrupted rows using on\_bad\_lines='skip'.  
 - Removed null or invalid entries.  
3. Merged Data: Combined metadata with ratings using movieId and id.  
4. Feature Engineering: Combined genres, tags, and keywords into a single string.

## 5. Model Development

A. Content-Based Filtering (Using sklearn)  
- Used CountVectorizer to extract features.  
- Calculated cosine similarity to find similar movies.  
  
B. Collaborative Filtering  
- Built user-item matrix using pivot tables.  
- Used cosine similarity between users for recommendations.

## 6. Streamlit Interface

Built an interactive web app using Streamlit to:  
- Accept user preferences  
- Choose recommendation method  
- Display top 5 recommended movies  
Run using: streamlit run app.py

## 7. Optional Add-ons

Planned future upgrades include:  
- Sentiment-based filtering using NLP  
- Hybrid model combining content and collaborative filtering

## 8. Tools & Libraries Used

Python, Pandas, Scikit-learn, Streamlit, NumPy, Matplotlib (optional)

## 9. Deliverables

- Notebook.py – Model training  
- app.py – Streamlit interface  
- CSV files – Cleaned datasets  
- report.docx – Project report

## 10. Conclusion

This project demonstrates a real-world application of machine learning in entertainment. Using user behavior and movie content features, we built a deployable recommender system. The solution is scalable and can be enhanced further for production use.

S