# Hybrid Movie Recommendation System Using User-User Collaborative Filtering and Content-Based Filtering

1. Introduction

This project focuses on developing a Hybrid Movie Recommender System combining two recommendation techniques: User-User Collaborative Filtering and Content-Based Filtering. The main objective of our hybrid approach is to effectively blend personalization, achieved through collaborative filtering, with the novelty provided by content-based methods. Utilizing datasets such as user ratings, movie metadata, and genre information, our recommender system aims to deliver accurate and relevant movie recommendations tailored specifically to user preferences, enhancing the overall user experience.

2. Project Flowchart

The flowchart provided visually outlines the project's pipeline, capturing each essential phase, from initial data collection and processing to final recommendation evaluation. Below is the visualization of the flowchart:
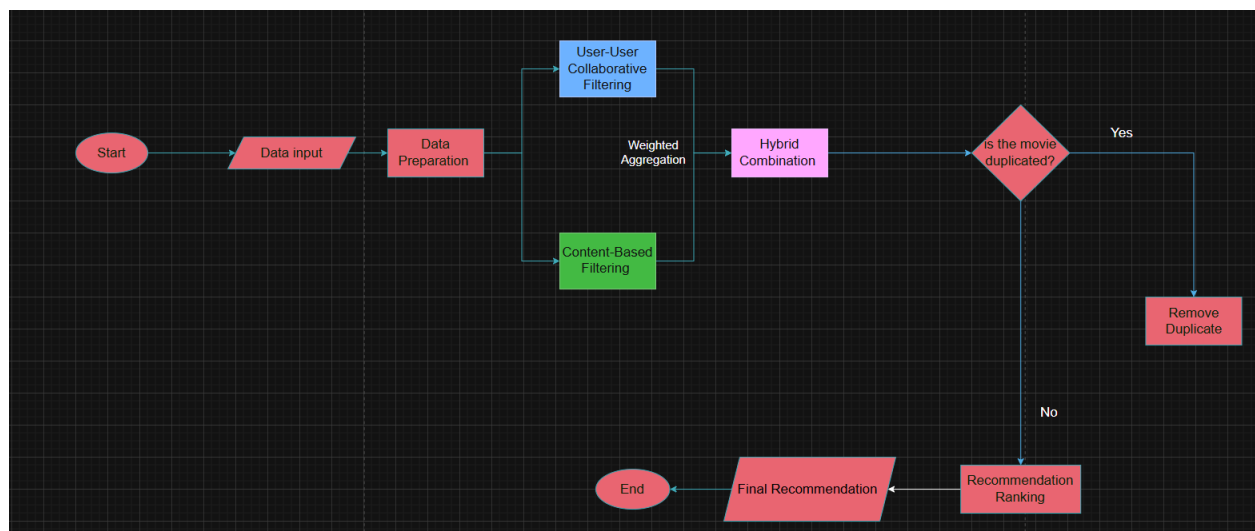
2.1 Data Input

Initially, multiple datasets (ratings.csv, tags.csv, Genres_encoded.csv, Films_metadata.csv, Movies.csv, links.csv) are loaded. These datasets contain critical user-item interactions, movie metadata and movie genre information which will be used for the recommendation process.

2.2 Data Preparation

This stage involves preprocessing the input datasets by merging them using a common identifier, **movie_id**. A user-item matrix – aka Utility matrix – is constructed from the ratings data, and a movie feature matrix is created from genres of these movies. This structured preprocessing ensures accurate and efficient recommendations.

2.3 User-User Collaborative Filtering

User-user collaborative filtering computes similarity scores among users based on their historical ratings. It leverages cosine similarity technique between user vectors to predict ratings for movies that users haven't rated yet, focusing heavily on personalized user experiences.

2.3 Content-Based Filtering

In parallel, content-based filtering computes similarities between movies using genres as features. Recommendations here are based on movie characteristics rather than user interactions alone, ensuring novel and diverse suggestions.

2.4 Hybrid Combination

The results from both collaborative and content-based methods are combined through taking n requested recommendations from both systems and sorting the output based on global average rating then taking only the number of requested recommendations. This hybridization balances personalization from collaborative filtering with the novelty of content-based filtering, optimizing recommendation quality.

2.5 Recommendation Ranking & Duplication Removal

Combined scores are ranked, ensuring that the most relevant recommendations surface to the top. A check for duplicate movie recommendations is performed, ensuring that each movie is unique within the final recommendation list.

2.6 Final Recommendations

Finally, the refined and evaluated recommendations are presented to the user, offering an optimized list of unique movies that cater precisely to their preferences. This comprehensive yet streamlined approach ensures the recommender system effectively meets user expectations by delivering high-quality, personalized movie recommendations.

3. Summary

In summary, the hybrid recommender system leverages the strengths of collaborative and content-based filtering methods to deliver personalized and novel movie recommendations. Each phase, from data preprocessing to final evaluation, is carefully structured to maximize accuracy and relevance, ensuring high user satisfaction and an improved recommendation experience.