# HLD Movie Recommendation System

## Introduction

**Recommender Systems-** We live in an era where every product has thousands of options to choose from. We can take the example of online shopping, gaming, social media and list goes on. Recommender systems help to personalize a platform and help the user find something they like.

**Movie recommendation system-** A movie recommender system, is an approach to filtering or predicting the users' movie preferences based on their past choices and behaviour. It's an advanced filtration mechanism that predicts the possible movie choices of the concerned user and their preferences towards a domain-specific item, aka movie.

## Types of Recommender Systems

Machine learning algorithms in recommender systems typically fit into two categories: content-based systems and collaborative filtering systems. Modern recommender systems combine both approaches.

**A) Content-Based Movie Recommendation Systems**
Content-based methods are based on the similarity of movie attributes. Using this type of recommender system, if a user watches one movie, similar movies are recommended. For example, if a user watches a comedy movie starring Adam Sandler, the system will recommend them movies in the same genre or starring the same actor, or both. With this in mind, the input for building a content-based recommender system is movie attributes.

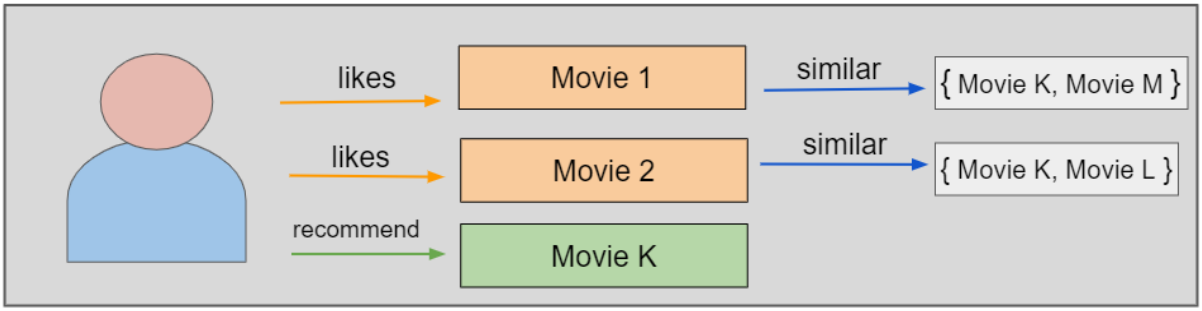



Figure 1: Overview of content-based recommendation system (Image created by author)

**B) Collaborative Filtering Movie Recommendation Systems**
With collaborative filtering, the system is based on past interactions between users and movies. With this in mind, the input for a collaborative filtering system is made up of past data of user interactions with the movies they watch.
For example, if user A watches M1, M2, and M3, and user B watches M1, M3, M4, we recommend M1 and M3 to a similar user C. You can see how this looks in the figure below for clearer reference.

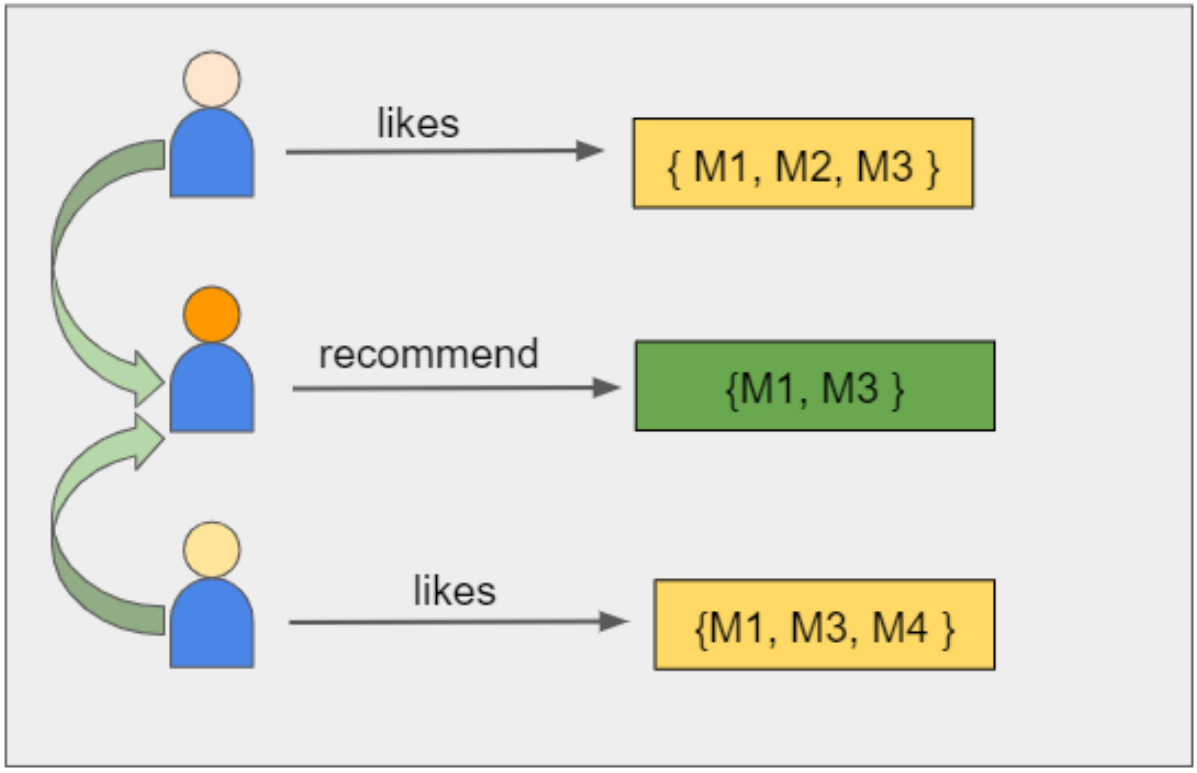


Figure 2: An example of the collaborative filtering movie recommendation system (Image created by author)

This data is stored in a matrix called the user-movie interactions matrix, where the rows are the users and the columns are the movies.

## Problem Statement

Given a user id, recommend 5 movies for that user. You should recommend a movie which the user has the highest probability of liking. You must not recommend a movie the user has already seen. Current data set is for 100K ratings. Final data set would be for 1M ratings. So make sure your solution can scale

## Why is this feature important/ what are the benefits?

The easiest and simplest way to do this is to recommend the most popular items. However, to really enhance the user experience through personalized recommendations, we need dedicated recommender systems. From a business standpoint, the more relevant products a user finds on the platform, the higher their engagement.

## Functional Requirement

Building a movie recommender system, which will recommend top 5 movies for a particular input userId.

## Scope

HLD of various design proposal for movie recommendation system. Preparation of dataset is not covered in the scope of this design.

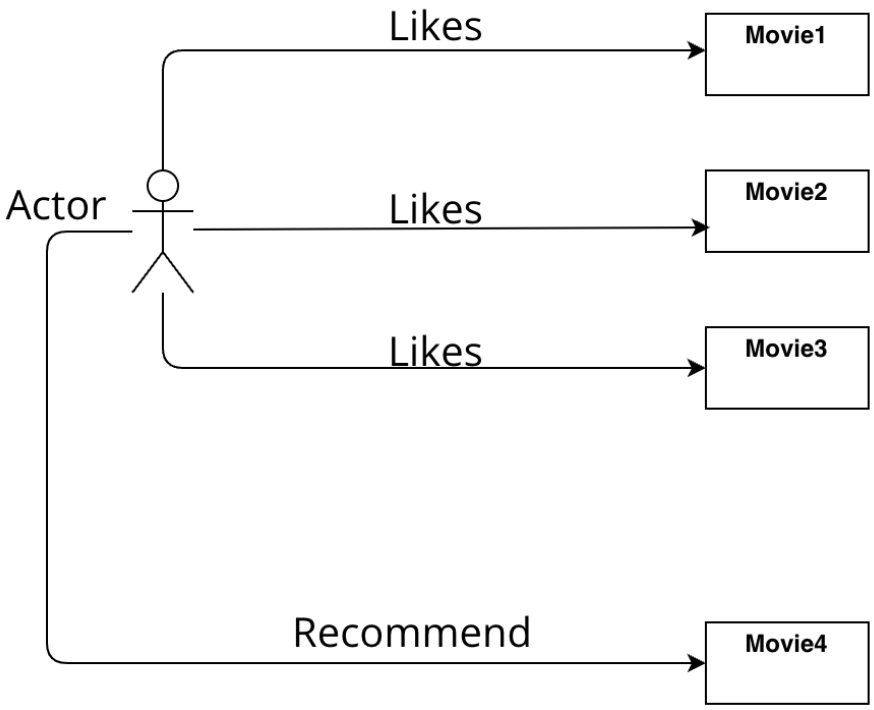
## Assumption (if any)

Provided dataset is suitable for this problem statement.
We are assuming that user already exist in dataset, we are not recommending any movie to new user.

## RACI Matrix

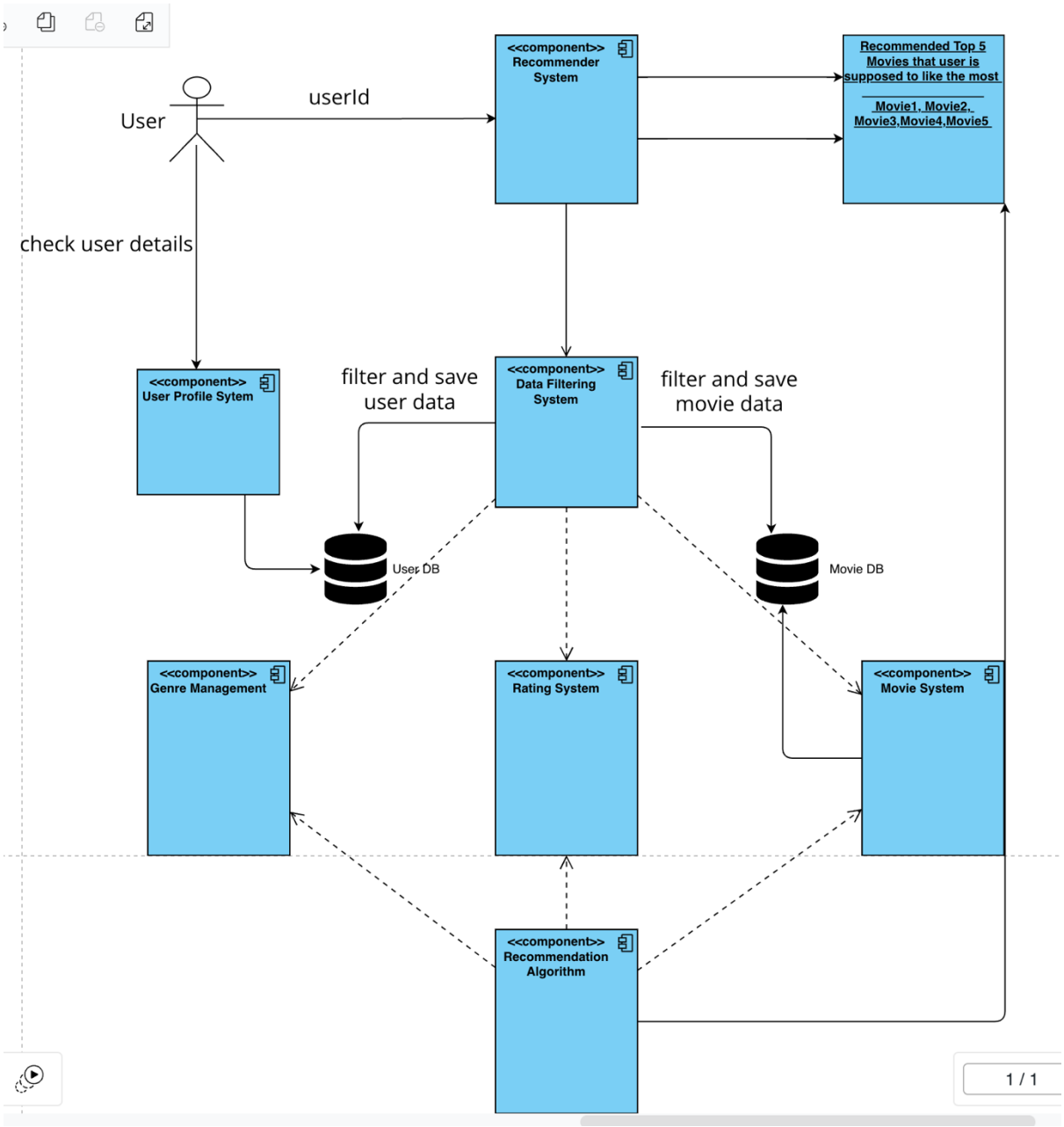
| | |
|---|---|
| Responsible | Himanshu Attri |
| Accountable | Avinash Garg |
| Consulted | Avinash Garg, Shivendra Kumar |

## High Level Design




**Movie Recommendation System**

On the basis of existing movies rated by user, we have to recommend movies to user.




It can be divided into following major component

1. **Data Filtering System**- This system will process raw data and will convert into meaningful information.
2. **Movie System**- It will be responsible for managing movies like storing details of movie, finding relation between different movies on basis of genre.
3. **User Profile System**- It will store data related to user like user attributes and his/her ratings.
4. **Rating System**- This system will be responsible for processing user data, finding most likes movie/ genre.
5. **Genre Management**- It will store available genre types and related movies for all the genre.
6. **Recommendation Algorithm**- This algorithm will be designed with the help of historical data present inside movie, rating system. It will be used to find related movies.

## Appendix

**Open Questions**
How are we going to validate the results of recommender system?
How to measure accuracy of the system?

**Reference**

https://towardsdatascience.com/how-to-build-a-movie-recommendation-system-67e321339109
https://www.geeksforgeeks.org/user-based-collaborative-filtering/

**Meeting Notes**

| Date | Attendees | Notes |
|---|---|---|
| 15 March 2023 | Himanshu Attri, Shivendra Kumar | |