Software Requirement Specification

Document

MOVIE RECOMMENDATION SYSTEM

Submitted To:

Santosh Singh Rathore Sir

Submitted By:

Donthagani Sai Rohan	2021BCS-025
Mekala Bhavana	2021BCS-041
Nelluri Pavithra Sai Lakshmi	2021BCS-049

TABLE OF CONTENTS

1. Introduction

1.1 Purpose

The Software Requirements Specification (SRS) document's purpose is to provide a complete overview of our software product, including its parameters and goals. By outlining the issue statement in detail, this paper intends to collect, analyze, and provide an in-depth understanding of the Hybrid Movie Recommender system. It focuses on the capabilities and needs of stakeholders while specifying high-level product features. This article describes the Hybrid Movie Recommender system's simple user interface, hardware, and software requirements.

1.2 Project Scope

With the assistance of our recommendation system, this software allows users of the customer platform to swiftly browse material. We are working on a Hybrid Recommendation System for Movies that combines collaborative and content-based filtering in the context of web-based recommender systems. We will specifically link the well-known TMDB data set. The system's content filtering is based on trained neural networks that represent unique user preferences. Through a variety of tests, we will show how supplemental user and item features affect the prediction accuracy of our proposed hybrid recommender. We will factorize our hybrid model using singular value decomposition to reduce system performance and identify latent user and item relations (SVD).

Because of the massive volume of information available online, the demand for highly developed personalization and filtering systems is increasing all the time.

Recommendation systems are a sort of information filtering that aims to provide items based on the interests expressed by the user.

1.3 Glossary

Stakeholder - Any person with interest in the project which is not a developer.

Collaborative filtering - It tries to find the correlation between users and recommend content from other categories that they might like.

Content - based filtering - Recommendation based on information available about the item but not about the user.

Singular Value Decomposition(SVD) - It is the factorization of a real or complex matrix that generalizes the eigendecomposition of a square normal matrix to any matrix via an extension of the polar decomposition.

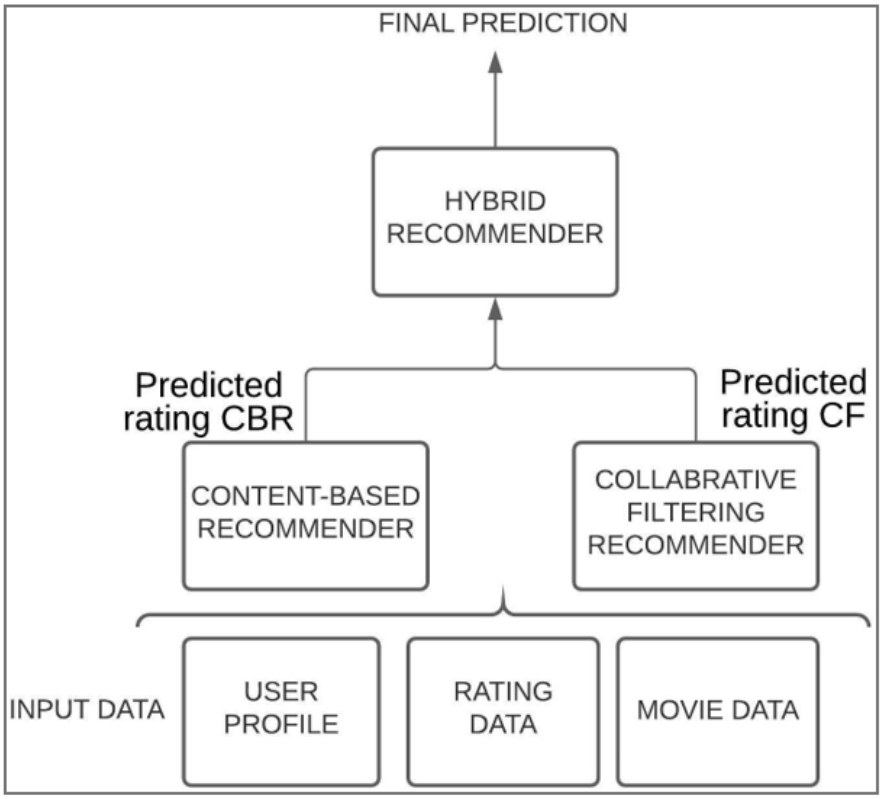
1.4 References

- Kunaver, Matevž & Perl, Tomaz & Pogacnik, Matevz & Tasic, Jurij. (2006). The evaluation of a hybrid recommender system for recommendation of movies.
- Abdollahpouri,Himan&Burke, Robin. (2019). Multi-stakeholder Recommendation and its Connection to Multi-sided Fairness.
- Rohan Nayak, Aniket Mirajkar, Jeetesh Rokade, Prof. Girish Wadhwa(2018). Hybrid Recommendation System For Movies

1.5 Overview

The remaining sections of this document provide a broad description, including user characteristics, product hardware, and functional and data needs. Section 2 of this document discusses a general description of the project. Section 3 describes the E-functional Store's requirements, data requirements and limitations, and assumptions. It also provides user feedback on the product. Section 3 additionally includes the product's particular needs. Part 3 goes through the external interface requirements and provides a full overview of the functional requirements. Section 4 is for more information.

2. Overall Description



2.1 Clients

Companies that offer video-on-demand services include Netflix, Amazon Prime, Hulu, and Hotstar. The company's business strategy, which offers movies on demand to its clients, makes it easy to find material on their website rapidly and boosts platform user happiness. The strength of its suggestions is key to its success. With the required information, this system may be expanded to any video-on-demand services.

We define similarity broadly, including how various movies, members, and genres are similar. It makes use of expressions like "Similar titles to watch quickly" and "More like...". One of the most important components of any OTT Platform is search.

2.2 Competitive Environment

There are several recommender systems available. Although all systems still base their recommendations on either a content-based or a collaborative recommender system, we are working to improve the platforms' accuracy of recommendations by implementing them and creating a new kind of recommendation paradigm that combines both collaborative and content-based principles.

Our recommendation algorithm stands apart from others thanks to this novel methodology. We are unable to file patent applications, however collaborative filtering enhances our system to a greater extent the more people engage with it and the better it functions. The secret to our system's superior performance over other recommender systems is our user data.

Companies like Amazon, Hulu, Disney+, Sony, HBO, and others are heavily pursuing this type of recommender systems.

2.3 Viewpoints of the Stakeholders

The subscribers and viewers of AHA are its main stakeholders. Those are the ones who would be immediately touched by the acts of this initiative. The AHA firm will benefit from an increase in subscribers and watchers thanks to our hybrid recommender system.

With regard to the work, the research team of AHA, who are actively involved in the creation and maintenance of the algorithm and the system, is one of the secondary stakeholders.

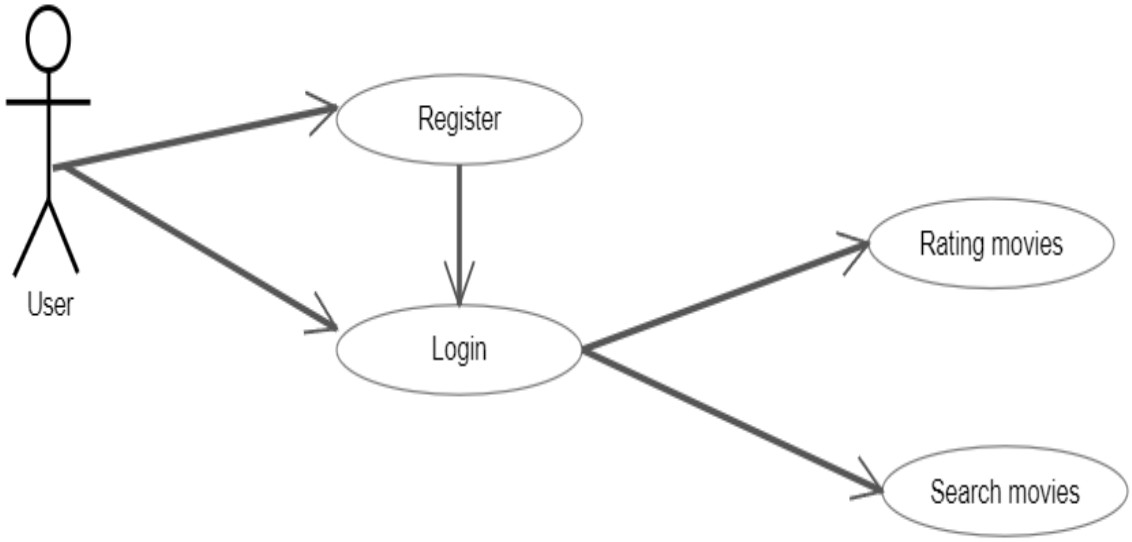
The system owner belongs to the third category of stakeholders. When movies are seen, the owner immediately benefits. But, the owner is equally worried about the company's long-term health. Studios may take their movies elsewhere and viewers can subscribe to other services that offer similar watching experiences. The utility function of the owner should take these factors into account. The goal of a logical system owner is to maximize personal usefulness.

2.4 User Characteristics

The user is expected to be Internet literate and be able to use a search engine. The main screen of the Recommendation Website will have the search function and a link to add new content for the content providers.

The Content provider is expected to be Internet literate, have login credentials to navigate through the website, and be able to log in and have full details about the content to be added so that the recommendation system works optimally.

2.5 User Use Case Diagram



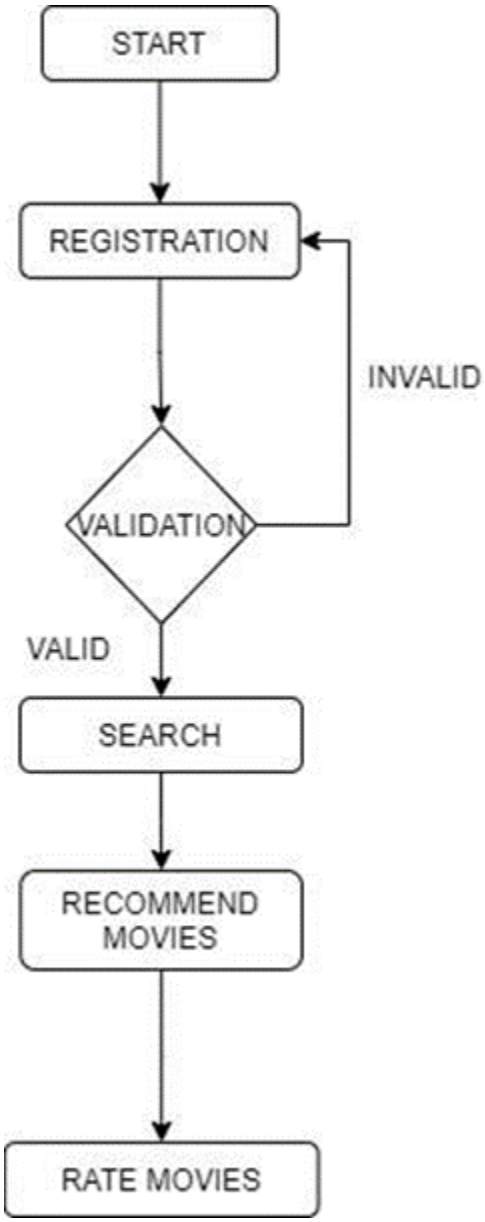
3. System Description

3.1 Functional Requirements Specification

The functions of each actor in this system are described in this part.

3.1.1 Registration For New User

Activity Diagram For New User



REQUIREMENT-1:

When a user first visits the website, he registers, looks for the movies he wants, and then selects the movies he likes or dislikes. This creates a profile for the user based on his preferences and the content of the movies he likes; the system then suggests movies to the user.

REQUIREMENT-1.1:

INPUT: The user chose the "REGISTER" choice.

OUTPUT: User is asked to provide personal information in order to register.

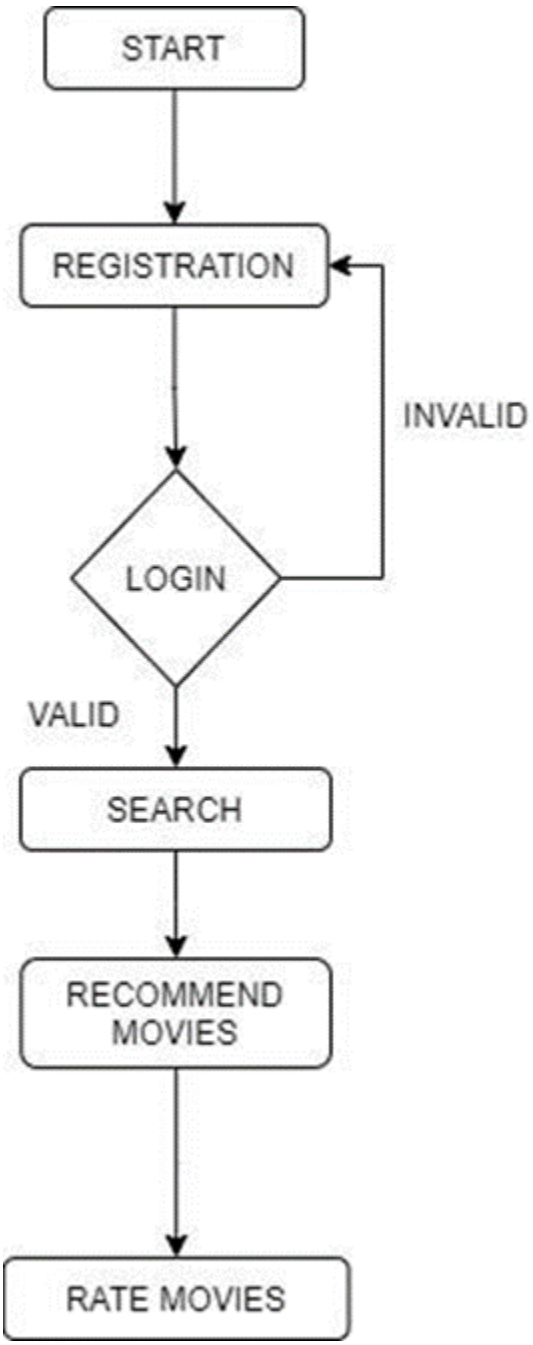
REQUIREMENT-1.2:

INPUT:The user inputs registration information.

OUTPUT: The user is taken to the search engine's home website. Check the input numbers' format as necessary during processing.

3.1.2 Recommendations For Existing User

Activity Diagram For Existing User



REQUIREMENT-2:

The User enters into the website if he already has an account there.We will suggest movies to the user based on what he types in the search box, his likes and dislikes, and the substance of the movies he enjoys.

REQUIREMENT-2.1:

The user inputs registration information.

OUTPUT: The user is taken to the search engine's home website. Check the input numbers' format as necessary during processing.

REQUIREMENT-2.2:

INPUT: enter USER NAME and PASSWORD OUTPUT: Redirect to the search engine's main page if the credentials are valid.Incorrect credentials are indicated by a prompt that prompts the user to enter them again.

REQUIREMENT-3.1:

INPUT:"search" option;

OUTPUT: Request for "MOVIE" name from person

REQUIREMENT-3.2:

INPUT: "MOVIE NAME"

OUTPUT: Information about the film, if it is included in the dataset, and suggestions for similar films.

PROCESSING: If the user is logged in, new movies are suggested based on collaborative and content filtering; otherwise, recommendations are made solely on the basis of content filtering.

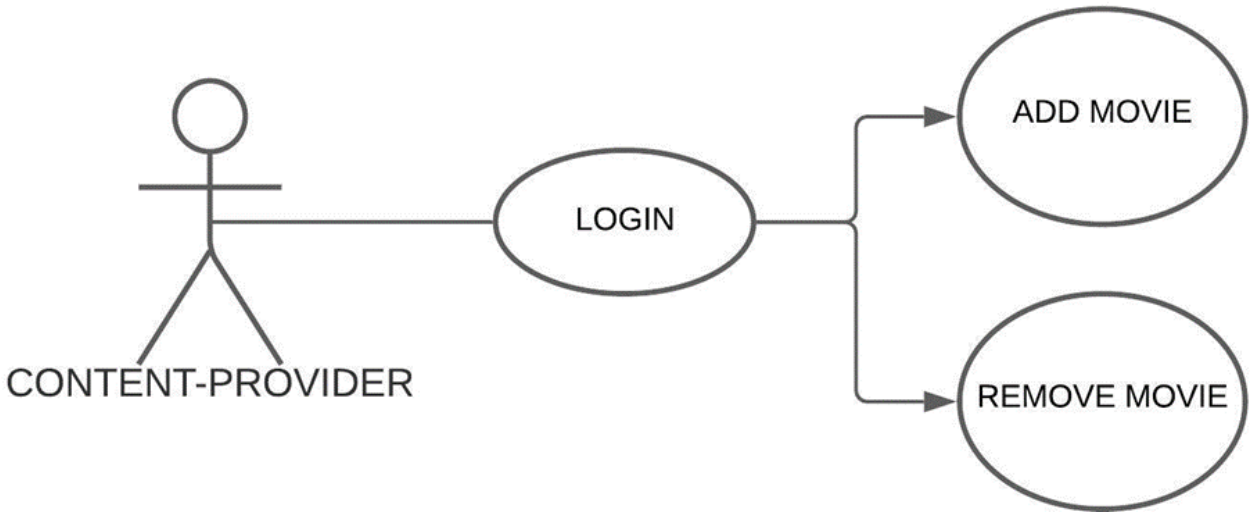
REQUIREMENT-4:

INPUT: LIKE / DISLIKE MOVIE

OUTPUT: The user profile is updated with the movie rating.

3.1.3 Content Provider

Content Provider Diagram:



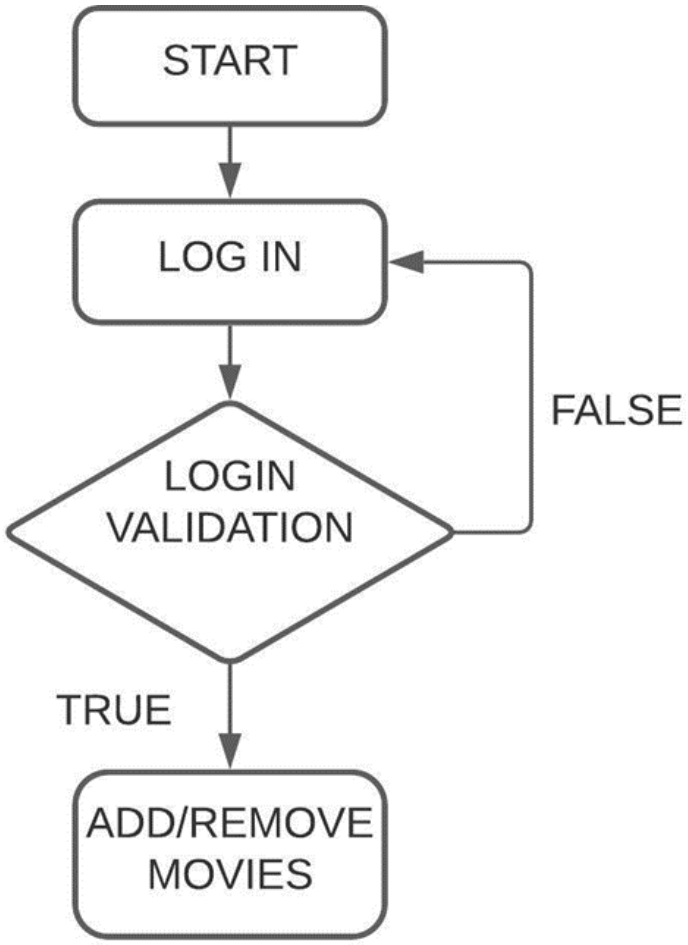
BRIEF-DESCRIPTION:

The network will be able to receive movies and the data that goes with them.

Initial Step-By-Step Description:

The content creator accesses the main recommendation website and then

Activity Diagram For Content-Provider



Requirement-5:

The necessary credentials were already in the possession of the content provider, who then logs in and uploads new movie data.

Requirement-5.1:

INPUT: “ADD/REMOVE MOVIES” option selected by the user. OUTPUT: user prompted to enter USER NAME and PASSWORD

Requirement-5.2:

INPUT: enter USER NAME and PASSWORD

OUTPUT: If the login information is accurate, it will direct you to the ADD/REMOVE page with a search engine. If not, it will notify you that your credentials are incorrect and require you to log in again.

Requirement-5.3:

INPUT: If Add movie is selected then Enter MOVIE DATA

OUTPUT: If the DATA is in the correct format, it is added to the database

Requirement-5.4:

INPUT: The choice to search movies is chosen if Remove data is chosen.
OUTPUT: Search box is prompted

Requirement-5.4.1:

INPUT: Proper keyword is entered to choose the movie.

OUTPUT: If the movie is present in the data set then delete the movie from the dataset.If movie is not present in the dataset show it to the content provider.
PROCESSING: The movie is deleted from the database based on the inputted keyword for the film.

3.2 Non-Functional Requirements

The online Recommendation system will be hosted on Heroku, which offers free servers for hoisting our sites.

To view the website, both users and content providers must have access to the required internet bandwidth.

PROCESSES AND TECHNOLOGY :

The recommender system employs unsupervised methods like dimensionality reduction and clustering/compression with subject modeling in addition to supervised methods like classification and regression. Factorization machines, matrix factorization, Singular Value Decomposition, connections to probabilistic graphical models, and techniques that are easily extended to be customized for various problems.

3.2.1 Performance

The application must be web-based and operated from a web server. Initial loading of the product will take some time, based on the speed of the internet connection and the media source. The hardware of the consumer or customer will have an impact on performance.If the Internet connection is strong, the program should update the user interface after 2-4 seconds of interaction.To avoid redundant data and boost speed, the database should be normalized.

3.2.2 Security

The computer that houses the website is protected by the hosting service provider.

Data Transfer: In all operations involving sensitive customer data, the system will use secure sockets. After a certain amount of inactivity, the system will instantly log off all users. All transfers must be verified by the system and the customer's web browser. The software won't save any cookies on the user's device that include their password or any other private information.

Data Storage: A customer's password will never be visible in their online browser or on the system's back-end servers. Back-end files for the system will be encrypted.

3.2.3 Safety

To avoid data loss, redundant databases should be used.Database backups ought to be performed every hour.

3.2.4 Software Quality Attributes

Availability: We should include as many movies as we can based on prioritization, ratings, and other factors. However, not all of the movies may be available.A user should be able to find a movie if they look for one.

Correctness: If a user looks for a movie using a keyword, such as a genre, the title of the movie, etc., the correct and related movies should be suggested in the event that the movie they are looking for is not available.

Maintainability: The application should make use of continuous integration to enable the quick and trouble-free deployment of new features and problem fixes.

Usability: The user experience should be simple to understand without the need for a tutorial and enable users to successfully complete their tasks.

3.3 External Interface Requirements

The HTTP interface will be applied. The machine must have a logical address in IPv4 format.

3.3.1 User Interfaces

The user interface for the software shall be compatible with any browser such as Internet Explorer, Mozilla, and Google Chrome by which users can access the system.

3.3.2 Hardware Interfaces

All of the hardware needed to connect to the internet is a hardware interface for the system because the application must operate over the internet. For instance, modems, WAN-LANs, and Ethernet Cross-Cables.

3.3.3 Software Interfaces

The best support and user-friendliness of the Windows operating system are the reasons we selected it.We have selected a SQL database to save information about the movie, customer preferences, etc.We will use Python libraries, Javascript, HTML, and other tools to carry out the assignment.

3.4 Risks and Mitigations

RISKS:

1. It's possible that not enough people will use the recommendations engine, which would make data more scarce and lower the yield on investment.
2. If the items pool is not under control, the recommender algorithm might suggest subpar items.
3. Users' worries about data privacy may be sparked by the recommender system.

MITIGATIONS:

For Risk 1: All users must be aware of the recommender engine's presence. Users must also be specifically asked for their feedback on the suggestions.

For Risk 2: We recommend that item pools be pre-filtered in order to retain only desirable items. For instance, editors must pre-filter news articles before adding them to the collection of news articles. As a result, the recommender engine wouldn't recommend them.
For Risk 3: In order to protect the privacy of user data, precautions must be taken to prevent user connections from showing up during the recommendations step.

3.5 Evaluation metric

The following measures can be used to assess recommendation engines.

Precision:

$$\text{Precision} = \frac{tp}{tp + fp}$$

Where,

tp = movies that user like

tp + fp = Total movies recommended to the user Evaluation metrics for recommendation engines

Recall:

how many of the products a user's recommendations they like. It comes from:

Here, tp stands for the number of items that a user has suggested that he or she likes, and tp+fn stands for the total number of items that a user likes.

The suggestions are better the higher the recall

RMSE (Root Mean Squared Error):

It calculates the deviation from the forecasted ratings:

In this case, Actual is the initial rating, and Predicted is the rating that the model predicted.

The recommendations are superior when the RMSE value is lower.