



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VI Month of publication: June 2021

DOI: <https://doi.org/10.22214/ijraset.2021.35741>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Movie Recommendation System using Machine Learning

Ghanashyam Vibhandik¹, Siddhant Deore², Niketan Pawar³, Abhijit Gavali⁴, Prof. Reena S. Sahane⁵

^{1, 2, 3, 4}U.G Student, ⁵Professor, Computer Engineering Dept, SIEM, Nashik, India.

Abstract: *Movies are very significant in our lives. It is one of the many forms of entertainment that we encounter in our daily lives. It is up to the individual to decide whatever type of film they choose to see, whether it is a comedy, romantic film, action film, or adventure film. However, the issue is locating acceptable content, as there is a large amount of information created each year. As a result, finding our favourite film is really difficult. The goal of this research is to improve the regular filtering technique's performance and accuracy. A recommendation system can be implemented using a variety of approaches. Content-based filtering and collaborative filtering strategies are employed in this work. The content-based filtering approach analyses the user's history/past behaviour and recommends a list of comparable movies depending on their input. K-NN algorithms and collaborative filtering are also employed in this paper to improve the accuracy of the results. Cosine similarity is utilised in this work to quickly discover comparable information. The correctness of the cosine angle is measured by cosine similarity. People may quickly find their favourite movie content thanks to all of this.*

Keywords: *Movie recommender system, content-based filtering, collaborative filtering, cosine similarity, KNN algorithms, Hybrid recommendation system, SVD Algorithm.*

I. INTRODUCTION

Recommendation systems are a type of information filtering method that aims to predict the preferences of the user. These techniques are now extensively employed in a variety of fields, including movies, literature, music, cuisine, videos, and locations. The term "movie recommendation" is used in this study. Every year, a variety of films of various genres are released. Horror, action, animation, and adventure films, among others. As a result, individuals are having difficulties finding their favourite film material in these vast stores of film content. As a result, the Movie Recommendation System enables us to quickly find our favourite movies while also saving time. People used to go to a real store to shop. In which the number of goods available is limited. The amount of movies that may be displayed in a shop, for example, is limited. People nowadays, however, have access to a wealth of information via the internet. However, as time goes on, information becomes more readily available. As a result, a new issue developed. The issue is that more time is wasted trying to figure out what people's favourite movie is. This is where the system of recommendations comes into play. Recommendation systems can be implemented in a variety of ways. Content-based, hybrid recommendation system with collaborative filtering. In the collaborative technique, a vast quantity of data about user behaviours, preferences, and predictions of what users would enjoy based on their resemblance to other users has been collected and analysed. Users' information is used as an input in the Content-based approach to locate related movies.

The purpose of this article is to develop a Movie Recommendation System that analyses past movie ratings assigned by various users in order to make recommendations to the user.

II. LITERATURE REVIEW

Many recommendation systems have been made over the years. Authors developed and implemented these systems using various big data and machine algorithms.

- 1) *Movie Recommender System Using Collaborative Filtering, Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020):* Authors developed movie recommendation system using collaborative filtering. K-NN algorithms is used to enhance the accuracy of results, at the same time removing the drawbacks of the content-based filtering. Authors also compared the existing systems to show that the system is more accurate and reliable.
- 2) *Movie Recommender System Based on Percentage of View, 5th Conference on Knowledge-Based Engineering and Innovation, Iran University of Science and Technology, Tehran, Iran:* A percentage of view approach is used to recommend movies to the users. The approach is used in a recommendation system for a media service provider named Namava. An Implicit Opinion Measure (IOM) improves performance of system on implicit feedback. Proposed system is works 5 times better than a random recommendation

- 3) *Movie Recommendation System Using Collaborative Filtering, 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*: Apache mahout is used to implement the movie recommendation system. The system takes into consideration the ratings given to the movies and provides movie suggestions. Further Research indicates that hybrid systems are more effective and provide more accurate results. Apache Prediction 10 is a Machine Learning Server that uses Elastic Search and Apache Hbase
- 4) *A Content-based Movie Recommender System based on Temporal User Preferences, 2017 3rd Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*: A content based movie recommendation system is developed. Authors used a temporal preference model of the user to recommend movies. The method is then evaluated using Movie Lens dataset.
- 5) *An Improved Approach for Movie Recommendation System, International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017)*: Author combined the content based filtering and collaborative filtering to improve the quality of the movie recommendation system. Author used support vector machine as a classifier. Author proposed hybrid approach and also compared the results with the existing pure approaches.

III.SYSTEM ARCHITECTURE

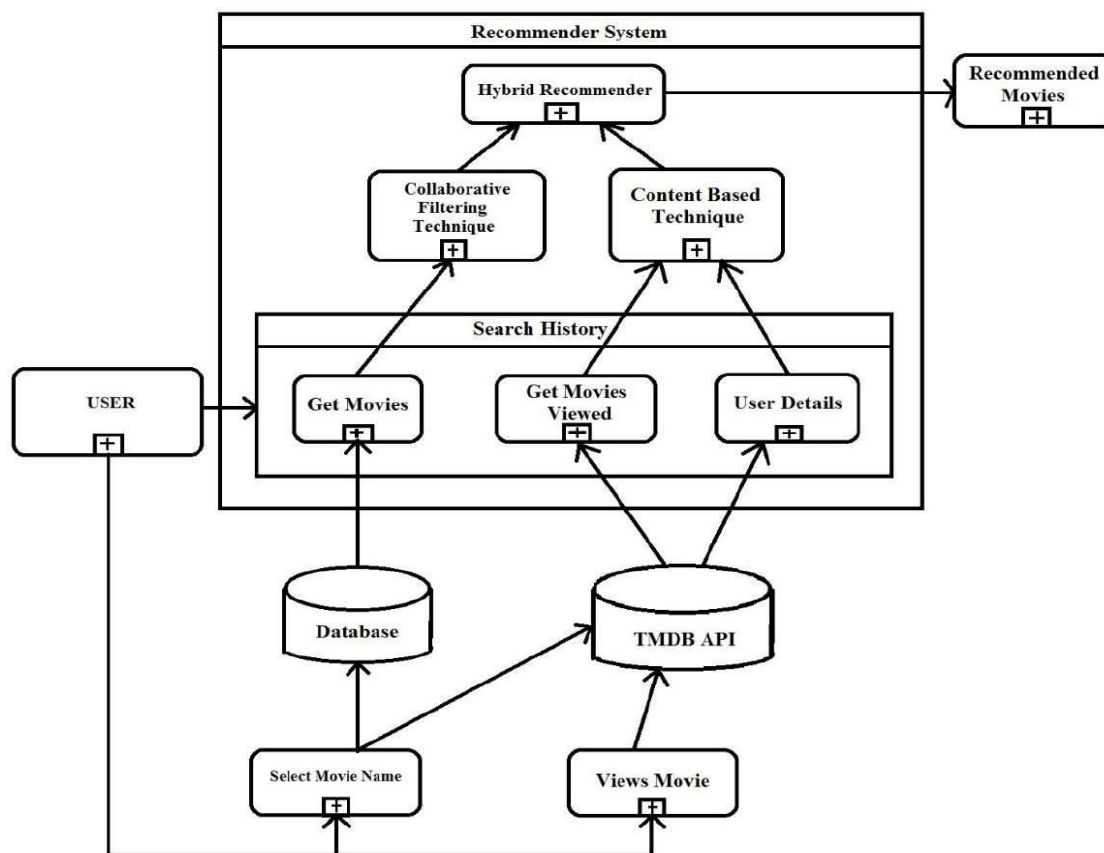


Figure 1: System Architecture

The movie suggestion mechanism is depicted in the diagram above. This system allows users to log in and view the many movies in the database. Additionally, the user may store his or her favourite movie genres to receive more tailored movie material. Several machine learning algorithms and approaches, such as Collaborative Filtering and Content-Based Filtering, are used to process the user's preferences in order to provide more tailored suggestions. The user may also examine movie data such as the cast, ratings, and reviews. In the collaborative technique, a huge quantity of data about user behaviours, preferences, and anticipating what users would enjoy based on their resemblance to other users has been collected and analysed. Users' information is used as an input in the Content-based approach to locate related movies. Because we are utilising the MovieLens dataset of 25 million ratings given to 62,000 movies by 162,000 viewers, the Singular Value Decomposition (SVD) algorithms provide more accurate findings. To generate comparable recommendations, the KNN collaborative approach with Cosine similarity is also used.

IV. IMPLEMENTATION

- 1) Importing the Libraries for Preprocessing of data : To analyze the data we need to import the libraries like pandas, matplotlib, etc.

```
In [1]: import pandas as pd
import numpy as np

In [2]: data = pd.read_csv('movie_metadata.csv')

In [3]: data.head(10)
```

| | color | director_name | num_critc_for_reviews | duration | director_facebook_likes | actor_3_facebook_likes | actor_2_name | actor_1_facebook_likes | gross |
|---|-------|-------------------|-----------------------|----------|-------------------------|------------------------|-------------------|------------------------|-------------|
| 0 | Color | James Cameron | 723.0 | 178.0 | 0.0 | 855.0 | Joel David Moore | 1000.0 | 760505847.0 |
| 1 | Color | Gore Verbinski | 302.0 | 169.0 | 563.0 | 1000.0 | Orlando Bloom | 40000.0 | 309404152.0 |
| 2 | Color | Sam Mendes | 602.0 | 148.0 | 0.0 | 161.0 | Rory Kinnear | 11000.0 | 200074175.0 |
| 3 | Color | Christopher Nolan | 813.0 | 164.0 | 22000.0 | 23000.0 | Christian Bale | 27000.0 | 448130642.0 |
| 4 | NaN | Doug Walker | NaN | NaN | 131.0 | NaN | Rob Walker | 131.0 | NaN |
| 5 | Color | Andrew Stanton | 462.0 | 132.0 | 475.0 | 530.0 | Samantha Morton | 640.0 | 73058679.0 |
| 6 | Color | Sam Raimi | 392.0 | 156.0 | 0.0 | 4000.0 | James Franco | 24000.0 | 336530303.0 |
| 7 | Color | Nathan Greno | 324.0 | 100.0 | 15.0 | 284.0 | Donna Murphy | 799.0 | 200807262.0 |
| 8 | Color | Joss Whedon | 635.0 | 141.0 | 0.0 | 19000.0 | Robert Downey Jr. | 26000.0 | 458991599.0 |
| 9 | Color | David Yates | 375.0 | 153.0 | 282.0 | 10000.0 | Daniel Radcliffe | 25000.0 | 301956980.0 |

10 rows x 28 columns

- 2) After Analyzing the data we need to apply some preprocessing methods to extract some useful information.

```
In [41]: final_df
```

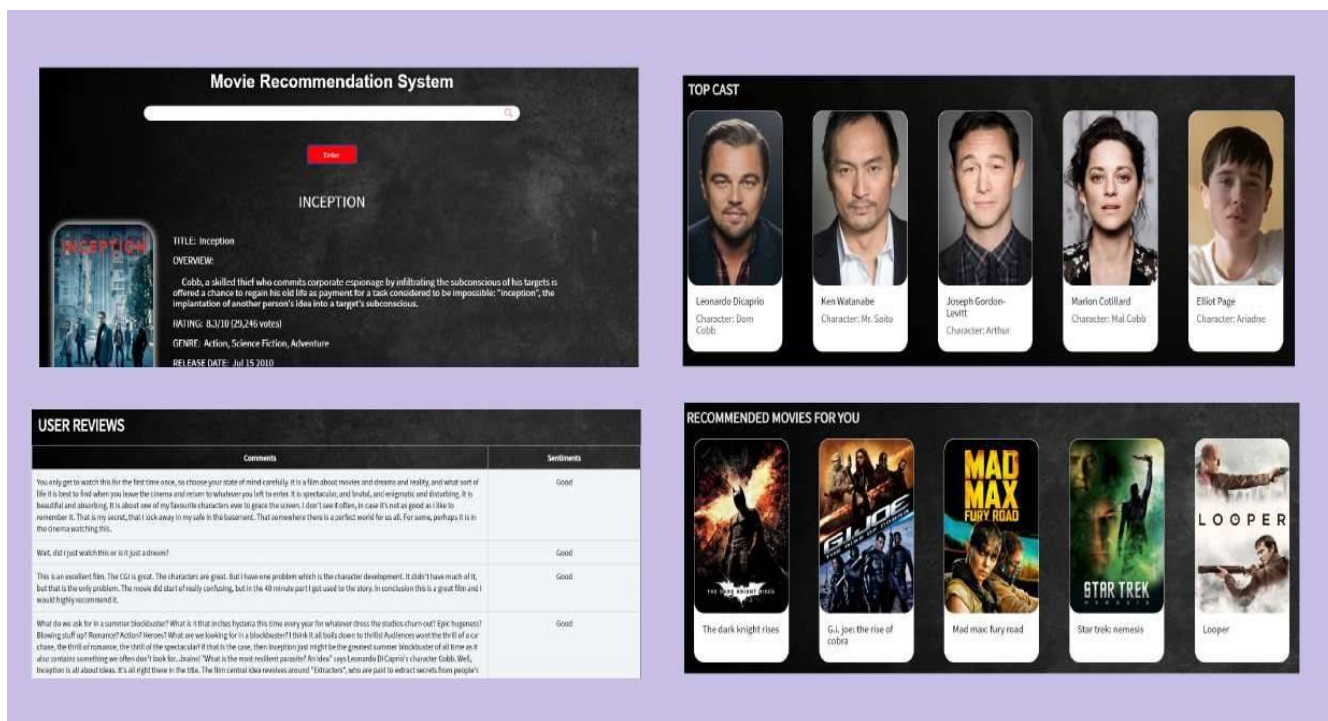
| | director_name | actor_1_name | actor_2_name | actor_3_name | genres | movie_title | comb |
|------|-------------------|-----------------|------------------|----------------------|---------------------------------|--|---|
| 0 | James Cameron | CCH Pounder | Joel David Moore | Wes Studi | Action Adventure Fantasy Sci-Fi | avatar | CCH Pounder Joel David Moore Wes Studi James C... |
| 1 | Gore Verbinski | Johnny Depp | Orlando Bloom | Jack Davenport | Action Adventure Fantasy | pirates of the caribbean: at world's end | Johnny Depp Orlando Bloom Jack Davenport Gore ... |
| 2 | Sam Mendes | Christoph Waltz | Rory Kinnear | Stephanie Sigman | Action Adventure Thriller | spectre | Christoph Waltz Rory Kinnear Stephanie Sigman ... |
| 3 | Christopher Nolan | Tom Hardy | Christian Bale | Joseph Gordon-Levitt | Action Thriller | the dark knight rises | Tom Hardy Christian Bale Joseph Gordon-Levitt ... |
| 4 | Doug Walker | Doug Walker | Rob Walker | unknown | Documentary | star wars: episode vii - the force awakens | Doug Walker Rob Walker unknown Doug Walker Doc... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 6005 | Joseph Kosinski | Tom Cruise | Miles Teller | Jennifer Connelly | Action Drama | top gun: maverick | Tom Cruise Miles Teller Jennifer Connelly Jose... |
| 6006 | Joel Crawford | Nicolas Cage | Emma Stone | Ryan Reynolds | Animation Adventure Family | the croods 2 | Nicolas Cage Emma Stone Ryan Reynolds Joel Cra... |
| 6007 | Liesl Tommy | Jennifer Hudson | Forest Whitaker | Marlon Wayans | Music Drama | respect | Jennifer Hudson Forest Whitaker Marlon Wayans ... |
| 6008 | Ridley Scott | Matt Damon | Adam Driver | Jodie Comer | Drama | the last duel | Matt Damon Adam Driver Jodie Comer Ridley Scot... |
| 6009 | Paul Greengrass | Tom Hanks | Helena Zengel | Neil Sandilands | Drama Western | news of the world | Tom Hanks Helena Zengel Neil Sandilands Paul G... |

6010 rows x 7 columns

- 3) After completing the preprocessing steps we get the final data set. After getting the final dataset we feed this data to the different machine learning algorithms.
- 4) Flask framework is used to embed the machine learning with the web interact with the system.

Flask is a Python-based microweb framework. It is referred to as a microframework since it does not necessitate the usage of any specific tools or libraries. It doesn't have a database abstraction layer, form validation, or any other components that rely on third-party libraries to do basic tasks. Extensions, on the other hand, may be used to add application functionality as if they were built into Flask itself. Object-relational mappers, form validation, upload handling, several open authentication methods, and numerous framework-related utilities all have extensions.

V. RESULT ANALYSIS



VI. CONCLUSION

The approaches of Collaborative Filtering and Content-Based Filtering are integrated in this article, resulting in a hybrid system that minimizes the error of discovering excellent suggestions. Because the system is given a dataset of 25 million ratings applied to 62,000 movies by 162,000 people, the KNN collaborative technique utilizing Cosine similarity produces comparable findings, while SVD produces more accurate results. These approaches improve the system's accuracy and efficiency. More features might be added in the future to improve the accuracy of the system. Additionally, a variety of alternative algorithms may be used to improve the recommendation system.

VII. ACKNOWLEDGMENT

We would like to offer our heartfelt gratitude to Asst. Prof. Reena S. Sahane, Department of Computer Engineering, Sandip Institute of Engineering and Management, for her unwavering support, prompt assistance, direction, and sincere cooperation throughout the duration of my work. We are thankful to her for providing all of the required facilities during the project's development, as well as for her assistance at various stages.

REFERENCES

- [1] Meenu Gupta, Aditya Thakkar, Aashish, "Movie Recommender System Using Collaborative Filtering" in Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020) IEEE Xplore Part Number: CFP20V66-ART; ISBN: 978-1-7281-4108-4
- [2] Ramin Ebrahim Nakhli, Hadi Moradi, Mohammad Amin Sadeghi, "Movie Recommender System Based on Percentage of View" in 5th Conference on Knowledge-Based Engineering and Innovation, Iran University of Science and Technology (2019)
- [3] Ching-Seh (Mike) Wu, Deepti Garg, Unnathi Bhandary, "Movie Recommendation System Using Collaborative Filtering", pp. 978-1-5386-6565-7118 (2018)
- [4] Bagher Rahimpour Cami, Hamid Hassanpour, Hoda Mashayekhi, "A Content-based Movie Recommender System based on Temporal User Preferences" in 3rd Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), 2017
- [5] Shreya Agrawal, Pooja Jain "An Improved Approach for Movie Recommendation System" in International conference on I-SMAC (2017)



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)