



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 [ijirset@gmail.com](mailto:ijirset@gmail.com)

 [www.ijirset.com](http://www.ijirset.com)

# AJAX Content Based Movie Recommendation System with Sentiment Analysis

Deepali Hiranman Gavhane, Rinki Chauhan, Rishabh Singh, Sejal Zarande,

Assistant Professor, Department of MCA, G.H.Raisoni College of Engineering and Management, Wagholi, Pune, India

P.G. Student, Department of MCA, G.H.Raisoni College of Engineering and Management, Wagholi, Pune, India

P.G. Student, Department of MCA, G.H.Raisoni College of Engineering and Management, Wagholi, Pune, India

P.G. Student, Department of MCA, G.H.Raisoni College of Engineering and Management, Wagholi, Pune, India

**ABSTRACT:** The basic objective of a recommendation system, which is a type of filtering software, is to predict a user's "rating" or "preference" for a domain-specific item. The recommendation system is a significant field that is well-liked and helpful in helping people make wise decisions. Therefore, our recommendation system's primary goal is to present consumers who looked for a movie they like a total of ten movie choices. One method for suggesting movies is content-based filtering. The system not only offers suggestions but also details on the movie that was searched for. The rating of the film, its release date, the cast, and its genres are among the extra facts. In addition, the system offers more cast-related details.

**KEYWORDS:** Recommendation System, Collaborative Filtering, Sentiment Analysis, Content-based Filtering, Cosine Similarity, Naive Bayes, AJAX Technology, Machine Learning, Natural Language Processing (NLP).

## I. INTRODUCTION

One of the most popular applications for information filtering is recommendation systems (RS). Recommendation systems seek to save users' time when surfing the internet by extracting relevant data and automatically generating suggested things (movies, books, news, music, and webpages) based on their past preferences. The main purpose of recommendation systems is to help users get results that are relevant to their interests. User must receive movie recommendations from web-based models. As Genres such as drama, comedy, action, drama, animation, and thrillers can be used to classify movies. In the last few years, a lot of research has been done on content-based movie recommendation systems. The first step in watching a movie is to choose one that the user enjoys.

This is where a recommendation system becomes necessary. The system may suggest well-liked films based on user preferences and ratings, but its real value lies in its capacity to suggest films based on personal tastes. A content-based movie recommender system utilizing movie ratings as the social information also we are using sentiment analysis for recommendation system. Sentiment analysis is the process of dissecting an emotive text, digesting it, summarizing it, and applying logic to it.

## II. LITERATURE SURVEY

1. Muhammad Abbas, Kamran Ali Memon, Abdul Aleem Jamali, Saleemullah Memon, Anees Ahmed In His paper "Multinomial Naive Bayes Classification Model for Sentiment Analysis" IJCSNS International Journal of Computer Science and Network Security, VOL.19 No.3, March 2019. presents a framework for sentiment analysis using the MNB classification algorithm along with the Term Frequency-Inverse Document Frequency (TF-IDF) method.
2. Ghanashyam Vibhandik, Siddhant Deore, Abhijit Gavali, Niketan Pawar, Prof. Reena S. Sahane. In "Movie Recommendation System Using Machine Learning" © 2020 IJCRT | Volume 8, Issue 11 November 2020. states content-based filtering over collaborative filtering to counter the issues like coldstart and low speed.
3. Parth Kotak, Prem Kotak in "Movie Recommendation System using Filtering Approach International Journal of Engineering Research & Technology (IJERT) Vol. 10 Issue 11, November-2021." Content based filtering focus on single user idea and predict on their interest unlike the collaborative based filtering.
4. ZSOLT NAGY in his "AJAX-Based Data Collection Method for Recommender Systems"

5. Implementation of ajax in recommendation system for asynchronous interaction between the user and the site.
6. Raghav Mehta ,Shikha Gupta in “Movie Recommendation Systems using Sentiment Analysis and Cosine similarity”
7. Problems which can arise and the solution for data cleaning and processing and gave us the insight and the blueprint for the development of the MRS.

### III. METHODOLOGY

#### 1. Data Collection

Movie tricks (title, category, duration, rating, banner, and so on) can be retrieved from the TMDB using the API. The largest data science community in the world, Kaggle offers robust tools that let users test and build models in a web-based credit marketplace for data science, as well as study and share data sets. Information on the response to different movies and character reviews may be found in Credit.csv on Kaggle. Two datasets will be used to build the recommender system: one is a Wikipedia database that has a list of movies from 2018 to 2022, and the other is a movie dataset that is available on Kaggle. The movie dataset comprises the top 5000 movies up until 2017 as well as the movies metadata dataset. An API is used by TMDB (The Movie Database) to obtain the

#### 2. Data Pre-processing

For training and testing reasons, we pre-process the data after it has been imported and visualized. Data collection needs to be labelled or put through a feature extractor test. The Data Frame contains selected data features. Line 5043 and 28 columns of the movie\_metadata.csv file include the following details: movie title, category, duration, actor name, director name, and rating. 45476 rows and 8 columns make up the Credit.csv dataset that was acquired from the Movie Database on Kaggle.

#### 3. Web scraping from Wikipedia using Python

Web removal is a copying method that involves gathering and copying particular data from the internet, usually posting it to a local website or spreadsheet for further analysis or retrieval. A web page's component must be downloaded and uninstalled in order to be removed. Wikipedia's removal of movie traits and character names from its database is justified by the dearth of data on Kaggle from 2018 to 2022.

#### 4. Data Processing

Following the generation of the datasets, we prepare them by creating several data frames using the Python "pandas" and "NumPy" libraries. The datasets will be split up into the various groups and frames needed for the following stage, which is the Flask framework (a Python web framework) data transmission to the front end. In order to make processing the dataset easier and maintain information order, it will be consolidated into a multiple list in the form of a dictionary that can be supplied to the HTML file.

A pair of values with KEY equal to [Movie, Cast, title, category, duration, rating, and so on]

### IV. MODELLING

The project's goal is to create a platform that can make movie recommendations, give users a thorough synopsis of the films they've searched for, and analyse movie reviews sentiment. There is no doubt that the time spent choosing a movie to watch will be reduced by the information offered. The primary goal of creating a movie recommendation system is to give people recommendations based on their favourite movies rather than only on popularity or rating. As a result, the recommendation will be extremely tailored, improving the recommendation system's accuracy. The user will be assisted in choosing a movie by the sentiment analysis and the extra details about the movie that they searched for. The user won't have to do any online browsing.

Cosine Similarity, a straightforward technique, is used to suggest movies. One metric for assessing how similar two items are to one another is cosine similarity. It can be calculated mathematically as the cosine angle in a three-dimensional plane between two vectors. To find out how distinct or similar the two vectors are from one another, we may also calculate their Euclidean distance. The movie that is searched in our instance is one of the vectors; the other movies in the database are checked as the second vector. Recommendations for the top 10 films with the lowest Euclidean distance for the searched film are displayed. One sort of content-based filtering technique is cosine similarity.

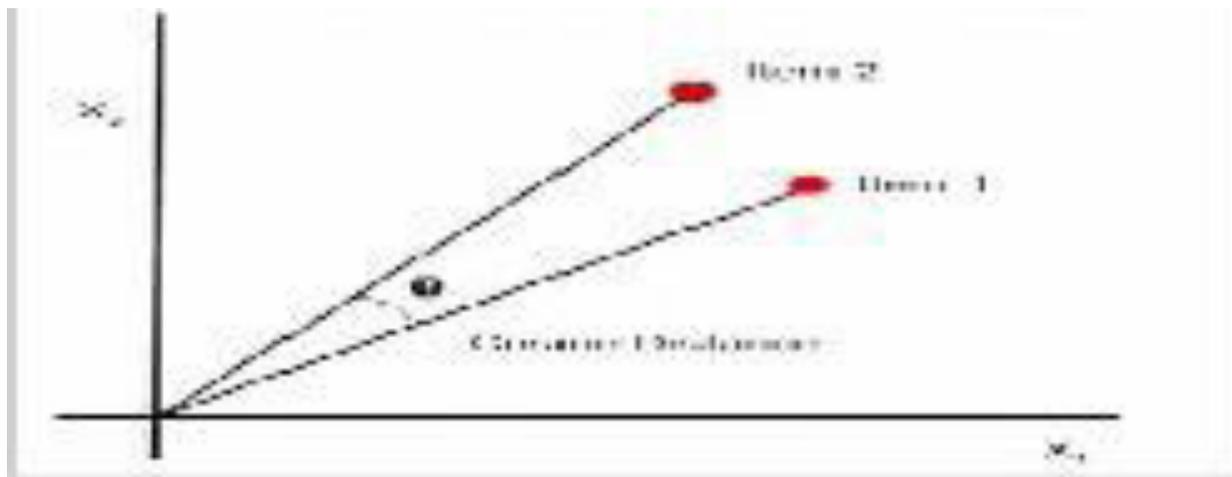


Fig.1 Cosine Similarity

Content-Based Filtering for Suggested Movies The recommendation algorithm is based on how similar the products are to one another. The main idea is that it is acknowledged or very likely that if you like something, you would like a similar product as well. A product functions well when its characteristics are easily determined. When a user provides explicit information, a content-based recommendation system for movies may suggest a movie and then develop a user profile based on that information. Ideas are also generated from this profile, and these ideas improve over time. "Concepts of Term Frequency and Inverse Document Frequency are utilized for filtering systems and information retrieval" in systems based on content. These terms are mostly used to evaluate a film's significance.

### Sentiment Analysis

A natural language processing (NLP) method called sentiment analysis, often known as opinion mining, is used to identify whether data is positive, negative, or neutral. Sentiment analysis is frequently used on textual data to assist organizations in understanding customer demands and tracking sentiment towards their brands and products in customer feedback. In this method, we want to use Natural Language Processing, or NLP, for Sentiment Analysis.

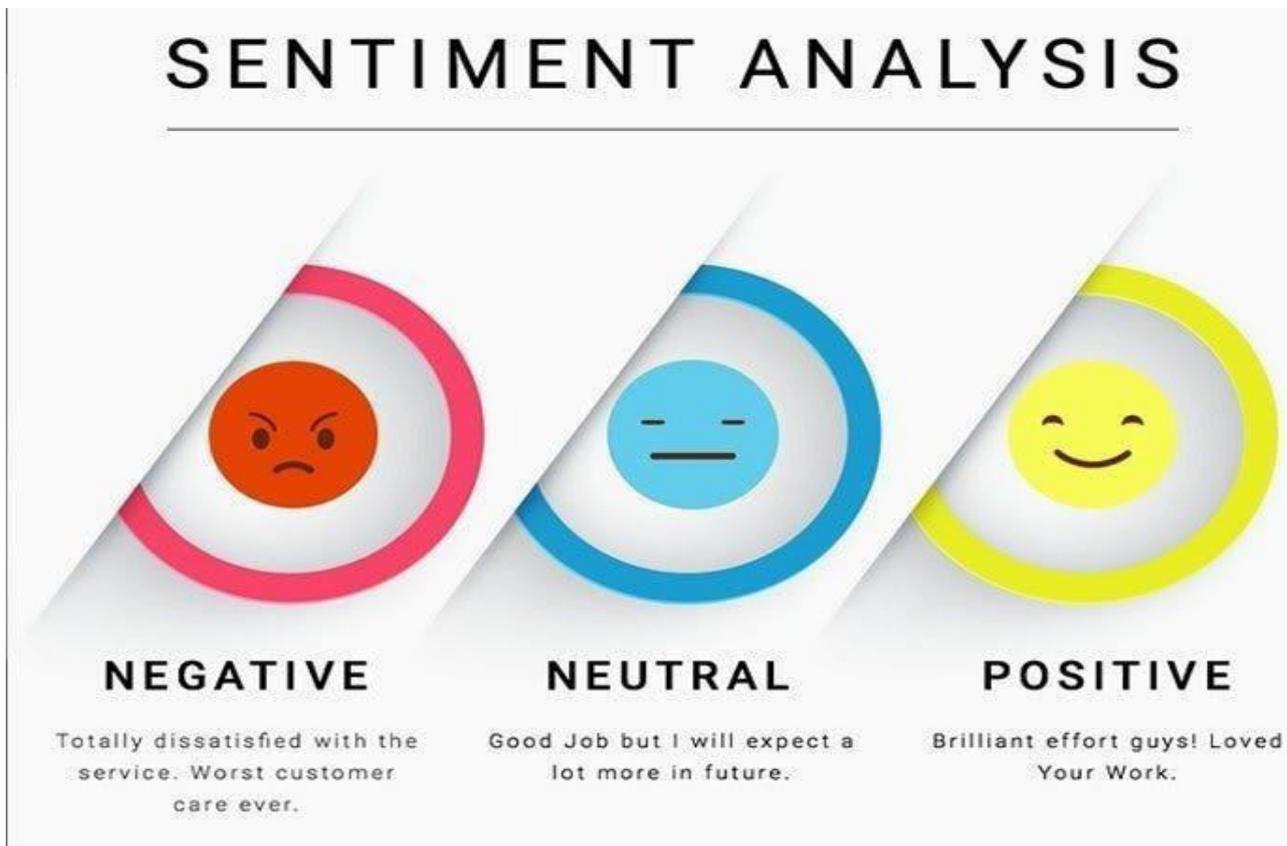
Fig.2 Sentiment Analysis

The feedback is read from the.txt file into an ipynb file. We obtain user reviews from the IMDB website through web scraping. The text is processed by NLTK (Natural Language Toolkit) and Tfidf (Term Frequency – Inverse Document Frequency) vectorizer. It aids in assessing a word's significance within a corpus or collection of texts. Since it indicates which terms are most frequently used in the text.

One well-liked approach for text classification is Naive Bayes. Despite being straightforward, it frequently works just as well as much more intricate ones. The way a naive Bayes classifier operates is by calculating the likelihood that various data qualities belong to a given class. This relies on the Bayesian theorem.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \text{ P(A)P(B)P(AB) = P(B|A), P(A)P(B) is the theorem.}$$

Thus, "the probability of A given that B is true equals the probability of B given that A is true times the probability of A being true, divided by the probability of B being true."



### V. PROBLEMSTATEMENT

The recommendation system on which this study is built suggests various items to users. Users will be able to choose movies from this system. Comparing this system to the current ones, the findings will be more accurate. The current system is based on the ratings of individual users. Since each user may have different tastes, this could occasionally be unhelpful for users whose preferences differ from those suggested by the algorithm. This technology finds commonalities across users and then suggests movies to them based on ratings from other users with similar interests. The user will receive an accurate recommendation as a result.

### VI. RESULT

1. Landing Page: This is the main webpage for the project. The input field is a form that receives a movie data string and sends it to the database's back end for preparation. The results are then returned to the client side via AJAX.

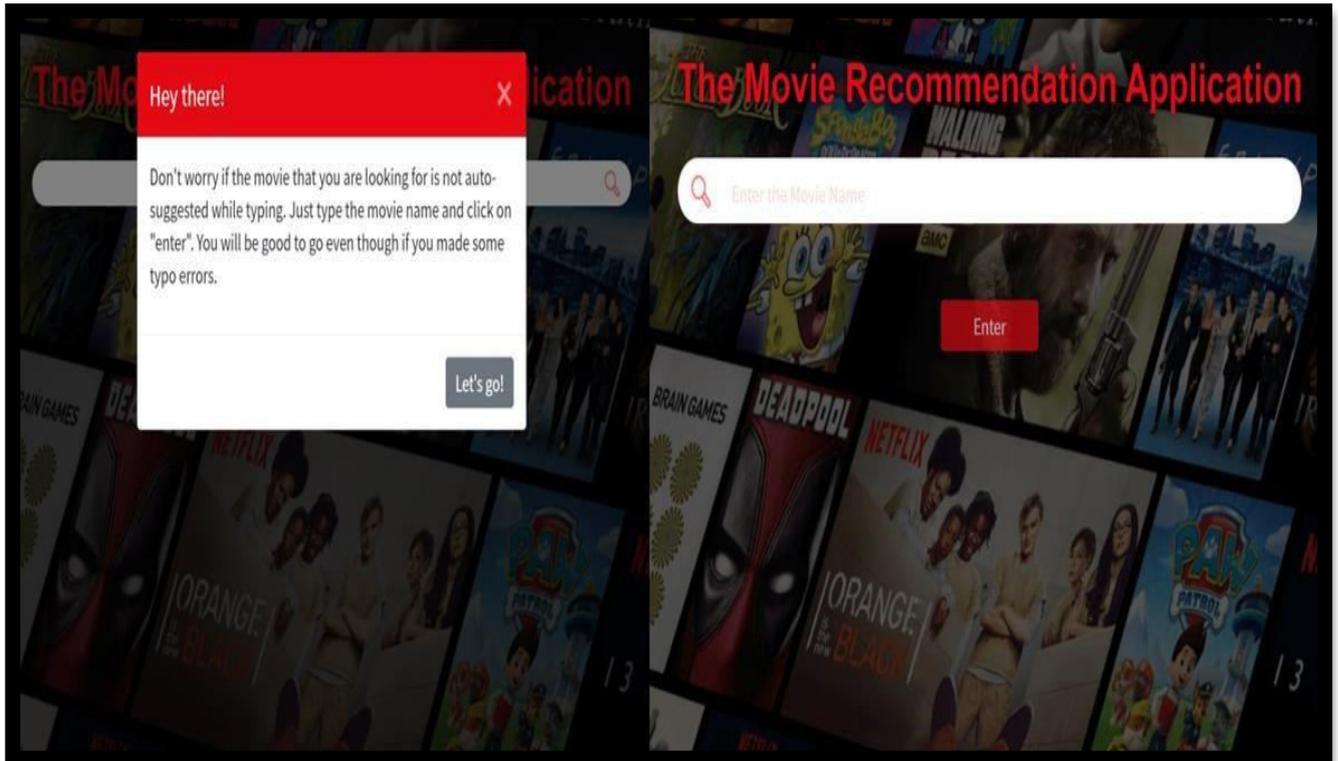


Fig.3 Home page

2. Search result:

The movie title that the user input will be the outcome that the search engine produces. This is the outcome of our project, in which the client sees the movie's details, including its genre, synopsis, runtime, rating, status, leading cast, name, year of release, etc., once the string has been processed at the back



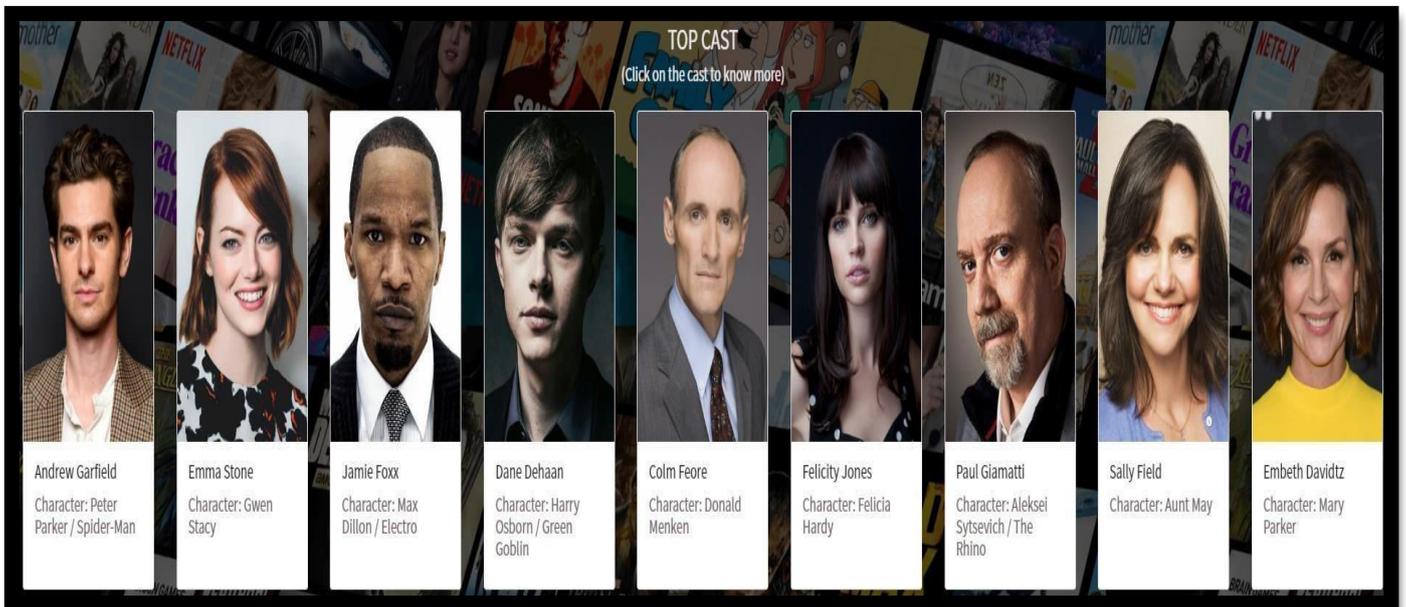


Fig.4 Search Result

### 3. Web Scraping:

Web scraping is used in the cast section of movies to display the cast members' bios. This is the result of the web scraping phase, where Wikipedia was used to gather details about each actor who starred.

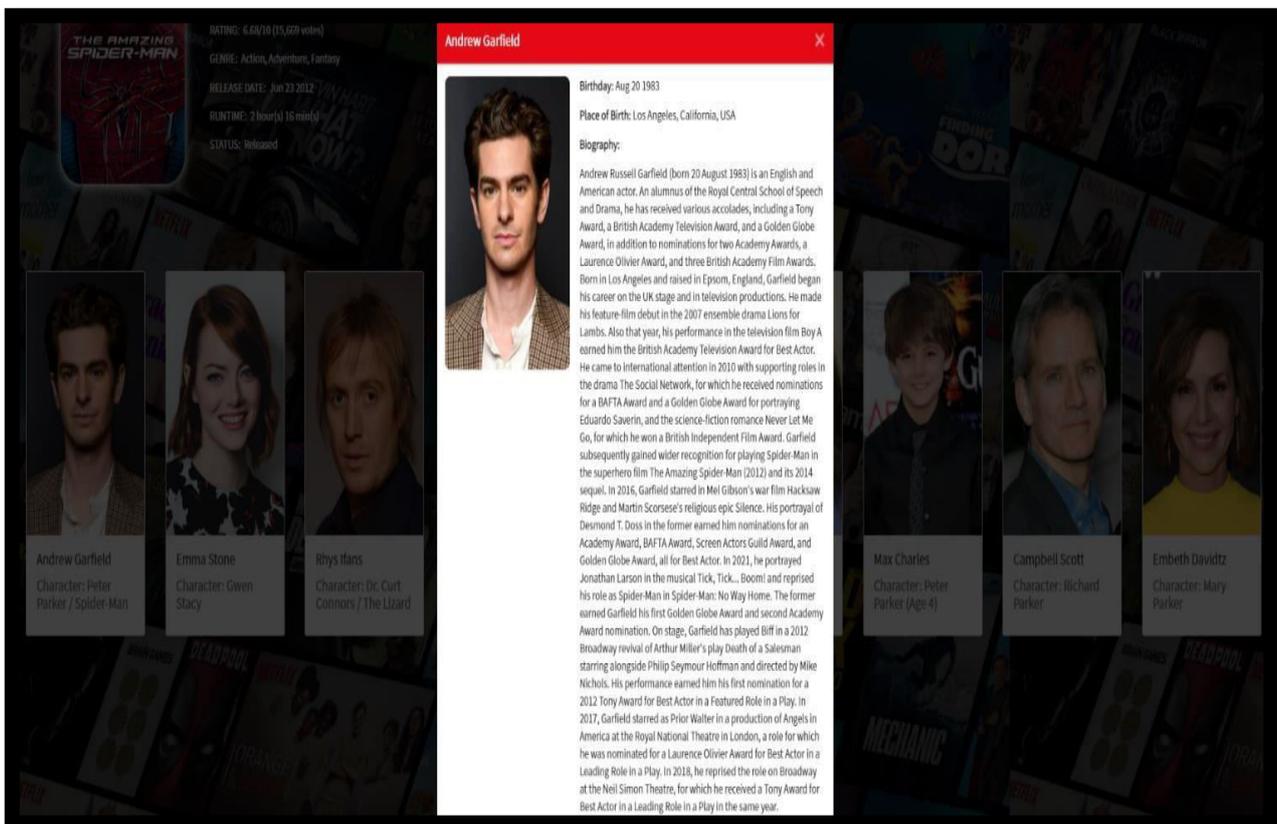


Fig.5 Actor's Bio

4. Sentiment analysis:

This is the project's sentiment analysis area, where user reviews are shown and utilized to categorize them as neutral, positive, or negative.

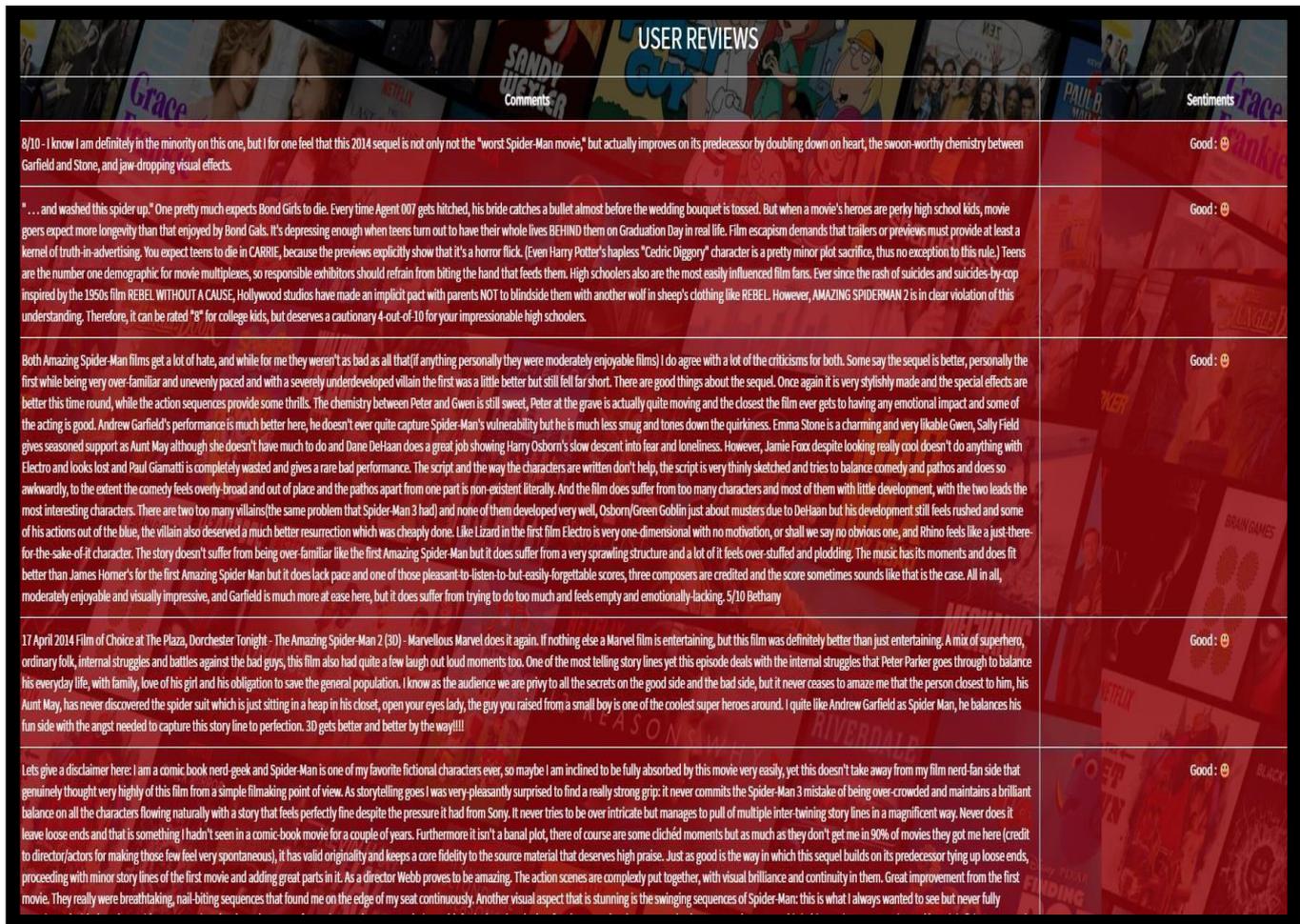


Fig.6 User's Reviews

5. Movie recommendation: Sentiment analysis is used to make movie recommendations based on the director, the genre, and the same actors. The films that have been rated highest by content-based filtering are displayed. This is the result of content-based screening.

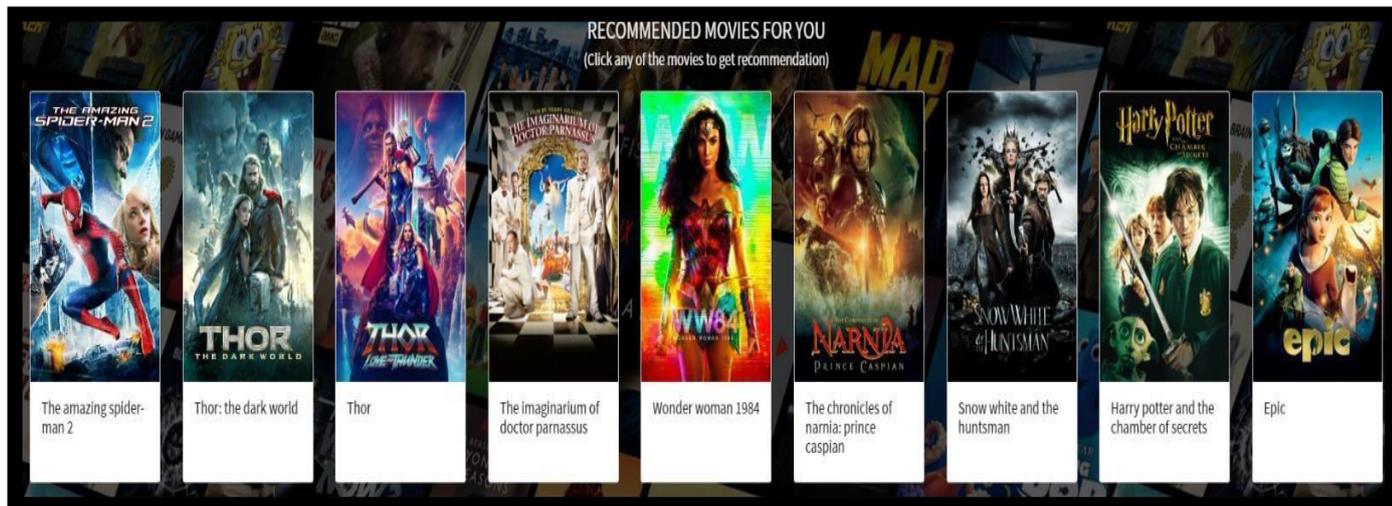


Fig.7 Recommend Movies

## VII. CONCLUSION

Recommendation systems can be a great tool for information filtering, providing consumers with just relevant content. Here, we show a recommender system that utilizes sentiment analysis and the TMDB data set. A social network and metadata are crucial components that are used to suggest movies.

Sentiment analysis was primarily used in the suggested methodology to examine user reviews of a certain film. Sentiment analysis is useful for obtaining information about how a given movie is received by the audience; this information is then used for creating more recommendations. By compiling more important information about user preferences, opinions, and other details, we enhance the suggestion process.

## REFERENCES

1. In "Movie Recommendation Systems using Sentiment Analysis and Cosine Similarity," Raghav Mehta and Shikha Gupta (2021) vol. 7(01), pp. 16–22.
2. Sentiment analysis using AJAX for movie recommendation system, Aarushi Sharma, Kandyala Lekhana, Keerthana M, Keshav Anil Rathi, Dr. Sumithra Devi K A, Vol. 03, Issue 04, 2021.
3. Surithra Devi KA, M. Keerthana, AJAX, sentiment analysis, Sharma, Aarushi, Kandyala Lekhana, and Keshav Anil Rathi suggest a film
4. "Movie Recommendation System Using Naive Bayes Algorithm with Collaborative Filtering," Dubey, A., and Ranjan, R. International Journal of Science and Research, 2018, 9(7), pp. 2018–2020.
5. John Le, "The 4 recommendation engines that can predict your movie tastes." (2018).
6. Neurocomputing, 2016, vol. 210, pp. 164–173. J. Ma and B. Shen, "An intelligent movie recommendation system through group-position sentiment analysis in microblogs."
7. Ravi, Kumar, and Vadlamani Ravi. "A check on opinion mining and sentiment analysis tasks, approaches and operations." Knowledge-based systems 89(2015)
8. Vader, Eric Gilbert, and Clayton J. Hutto a simple, rule-based method for evaluating the sentiment of social media material. This collection contains the papers given at the Eighth International Conference on Weblogs and Social Media, which took place in 2014.
9. Using Bayesian adaptive algorithms, stoner profiling exercises both explicit and implicit input. P. Zigoris and Y. Zhang are the authors of this work. Delivered in 2006 (27), at the Information and Knowledge Operations Conference
10. Information filtering with a nonparametric hierarchical Bayesian [frame(10) Yu, K., Yu, S., and Tresp, V. On runners 353 – 360 of SIGIR' 04, you may find the proceedings of the 27th periodic international ACM SIGIR conference on disquisition and advancement in information recovery. ACM Press published this book in 2004. (25)
11. Kaggle Your Machine literacy and Data Science Community
12. The Movie Database (TMDB), Themoviedb.org
13. <http://imdbws.com/datasets>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details