iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.36377

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Movie Recommending System Using Collaborative Filtering

P. Rama Rao[1], V.V. Santoshi Laxmi[2], G. A. N. Murthy[3], G. Bhargav Naidu[4]
[1, 2, 3, 4]*Satya Institute of Technology and Management*

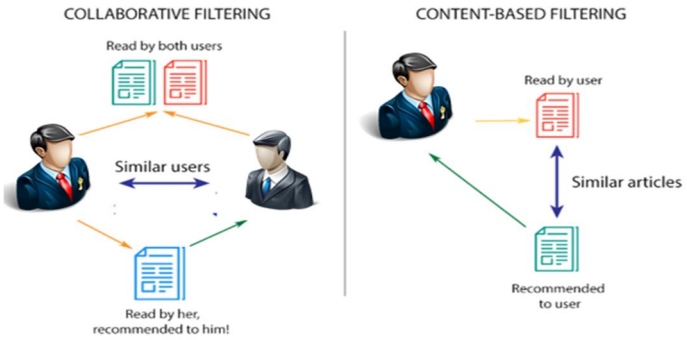
Abstract: ***Movies are one of the sources of entertainment, but the problem is in finding the content of our choice because content is increasing every year. However, recommendation systems plays here an important role for finding the content of desired domain in these situations. The aim of this paper is to improve the accuracy and performance of a filtration techniques existed. There are several methods and algorithms existed to implement a recommendation system. Content-based filtering is the simplest method, it takes input from the users, checks the movie and its content and recommends a list of similar movies. In this paper, to prove the effectiveness of our system, K-NN algorithms and collaborative filtering are used. Here, the usage of cosine similarity is done for recommending the nearest neighbours.***
Keywords: Collaborative filtering, Genre, Movie recommendation, Rating, Recommender system.

I. INTRODUCTION

Every one in the world loves movies. They are stress reliefers. Irrespective of their gender, age, interest all people have a positive opinion on movies. Some people loves drama, some people loves fantasy, some have interest in tragedy, thriller, fantasy etc. Some people choose movies based on genre and some watches movies based on their mood. Its very difficult to suggest movies because of their dynamic nature. Here our project comes into the picture, our system understand the interests of the different users and it will recommend based on the choices.

II. RELATED WORK

Recommendation System is nothing but a system which provides the domain item similarities based on rating or interest or any other criteria. This system considers all those factors and provides the similar type of domain items. In our case, the domain is movie. So the system recommends movies of the users choice.



1) Content-based Filtering

The name itself represents the way it filters the items. The system presents the movie based on the content of his previous searches. Here content may be of any type like Director, Genre, Rating, Cast, Year of releasing etc.

For example, let me take an user who likes movies ‘Titanic’ then the system can recommend certain movies based on ‘Romance/Drama’ or the system can recommend movies which was directed by ‘James Cameron’. So here the system recommend the movies based on data available. It can chose either cast or genre or other preferences for the movies like ‘Titanic’.

a. Downsides

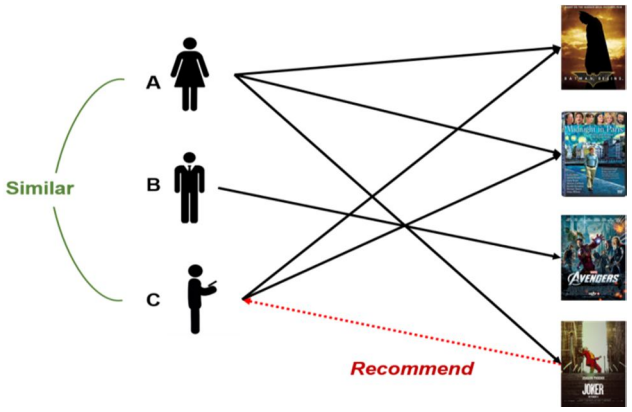
- Content dependency, if the information available is limited, then the efficiency of system to recommend movie is poor.
- Difficulty with new releases.

2) Collaborative Filtering

This filtering is entirely different with content filtering. Here Collaborative filtering takes consideration of users's past behavior as well as it searches the users of similar type, the system compares the behavior and then recommend movies. The main difference between the two filtering is data of the movie is more important in case of Content-based, similar users searches and interests are taken as a factor in case of Collaborative-based filtering.

Several methods existed to implement collab filtering but the actions and concerns of users changes the outcomes of the system.

There are 2 types of collaborative filtering algorithms:



A. User-based Collaborative filtering

The idea here is to find users that have indistinguishable choices and penetrating users choices, computing the likewise score and then endorse movies. For example, if the user say 'Amar' likes 'Titanic', 'Forrest Gump' and 'The Theory of Everything' while another user says 'Bon' likes 'Titanic', 'Forrest Gump' and 'The Girl Next Door' then the system found that they have similar choices because the above listed movies are Romance/Drama genres. So, there exists high chances that the user 'Amar' would like 'The Girl Next Door' and the user 'Bon' would like 'The Theory of Everything'.

Downfalls

- People are vigorous i.e., their choices differ from time to time so its hard to recommend movies.
- Movies has other features like year of release, so the other factors also effects the nature.
- This algorithm is not efficient incase of fake ids and users.

B. Item-based Collaborative Filtering

The concept here is to discover similar movies alternately of similar users and then nominating similar movies to that of 'A'. This can be done by searching every set of items that were rated or viewed or choosen by the same user, then computing the nature of those viewed or choosen across all users who wished or viewed both, and finally endorsing them based on alike scores.

Here, for example, let us take 2 movies 'Harry Potter' and 'The Orphan' and review their ratings by all users who have rated the above two movies and based on the closeness of these ratings, likes and based on this rating likeness by users who have rated both movies in that way the system find similar movies. So, if most of the familiar users have rated 'Harry Potter' and 'The Orphan' both similarly and there exists high chances that 'Harry Potter' and 'The Orphan' are similar, therefore if some user watched and liked 'Harry Potter' the system can recommended 'The Orphan' movie and vice versa.

Superiority over User-based Collaborative Filtering

1) People changes but movies won't change
2) There exists fewer movies than users
3) User may be faked but item cannot be faked

III. EXISTING SYSTEM

To evaluate the performance of the proposed system and to provide better results, experiments are conducted by comparing any random existing system with our proposed system based on terms of quality, accuracy, precision, recall, time computation. The accuracy-numeric value which determines the result of calculation fulfills to the precise or standard value. The existing system are compared and found that the proposed system is more reliable and accurate. It is also found that when the proposed methodology is applied to different larger datasets, both accuracy, and efficiency increase which proves that our system is both accurate and as well as efficient. This item-based filtering is more convenient than user-based.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

A. Pit Falls

1) Many wide varieties of recommending movies exists like Content-based, Collaborative (User-item, User-user), context based, hybrid methods and deep learning methods.
2) The hybrid approach provides very accurate results using both collaborative and content -based filtering
3) The content-based system uses data of users and interest and browsing history to calculate the results. This requires a lot of domain knowledge, thus becomes a drawback compared to collaborative filtering

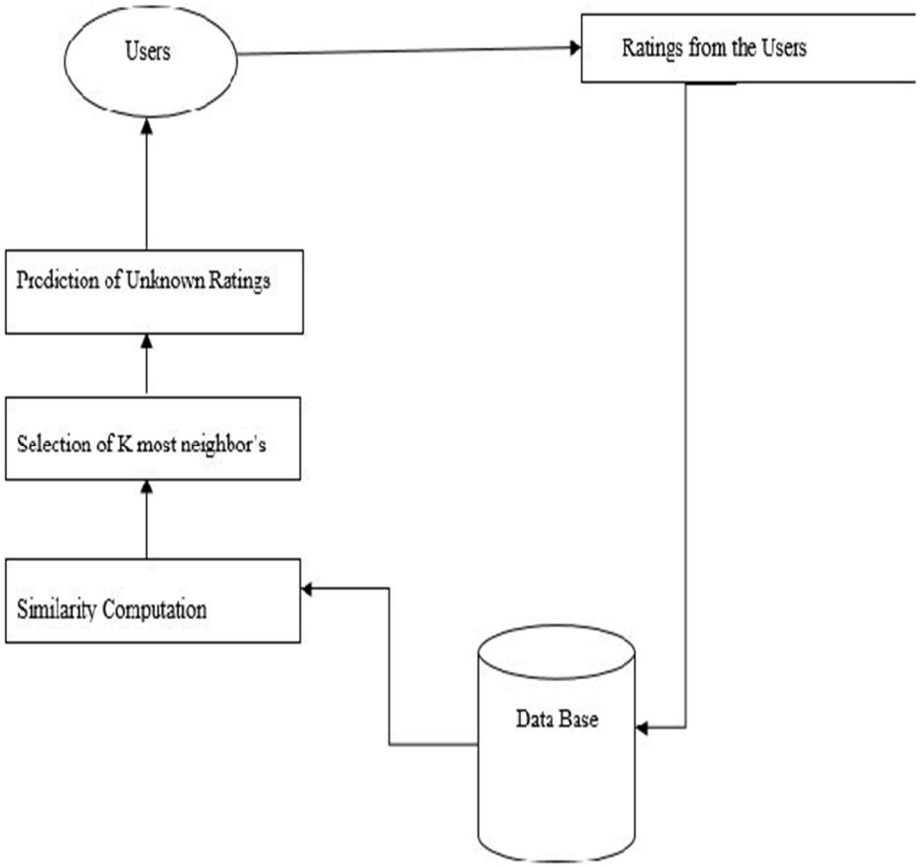
IV. PROPOSED SYSTEM

The proposed recommendation system used the collaborative filtering technique (item-based approach) which is far more accurate and more efficient to use, as the item-based method can be done offline and because of its non-dynamic nature whereas the user-based changes. The proposed approach uses the KNN algorithm to find the distance between the target movies with every other movie in the dataset and then it ranks the top k nearest similar movies using cosine angle similarity. The system results checked with some random recommendation system for its performance analysis.

A. Precedence

1) The objective here is to recommend movies using the item-based technique.
2) First, the extraction of the dataset to gather information about the target movie and the user's rating.
3) Second, the collaborative filtering begins with the formatting of the rating dataset so that it can be consumed by the KNN model, to remove the huge dataset handling problems. The dataset is reduced according to the popularity removing the noisy error pattern to get the sparse matrix.

B. Architecture For Movie Recommandation System



V. IMPLEMENTATION

A. There Is Three Input Database

1) Movies - containing movie-id, title, genres
2) Users - containing user-id, gender, age, occupation
3) Ratings – containing user-id, movie-id, rating

For data analysis, data filtering python libraries are used to analyze datasets. The use of pandas and numpy and scikit-learn libraries were given efficient results in the proposed system.
KNN will work on the processed data and uses cosine similarity to find nearest neighbours.

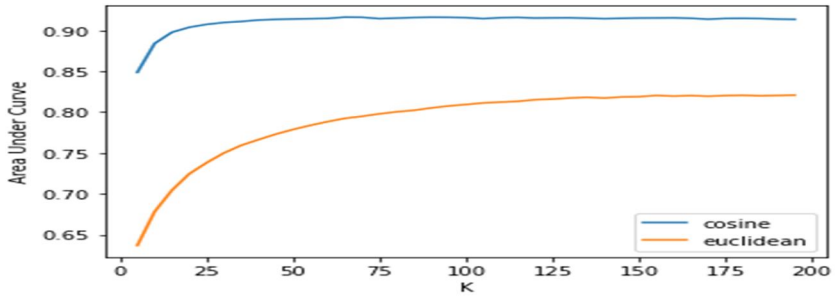
We can implement a KNN algorithm using following steps:

a) Loading the data
b) Set the value of k
c) For the desired result, repeat from 1 to count of training data points

- Calculate the distance between test data and each tuple of training data using cosine similarity metric.
- Arrange the computed distances in desired order based on computed merits
- Retrieve top k rows from the sorted array
- Get the most frequent class of these tuples
- Return the prevised class

Cosine similarity computes the affinity between two vectors of an inner product space. It is estimates by the cosine of the angle between two vectors and determines whether two vectors are pointing in roughly the same direction.

VI. RESULT ANALYSIS



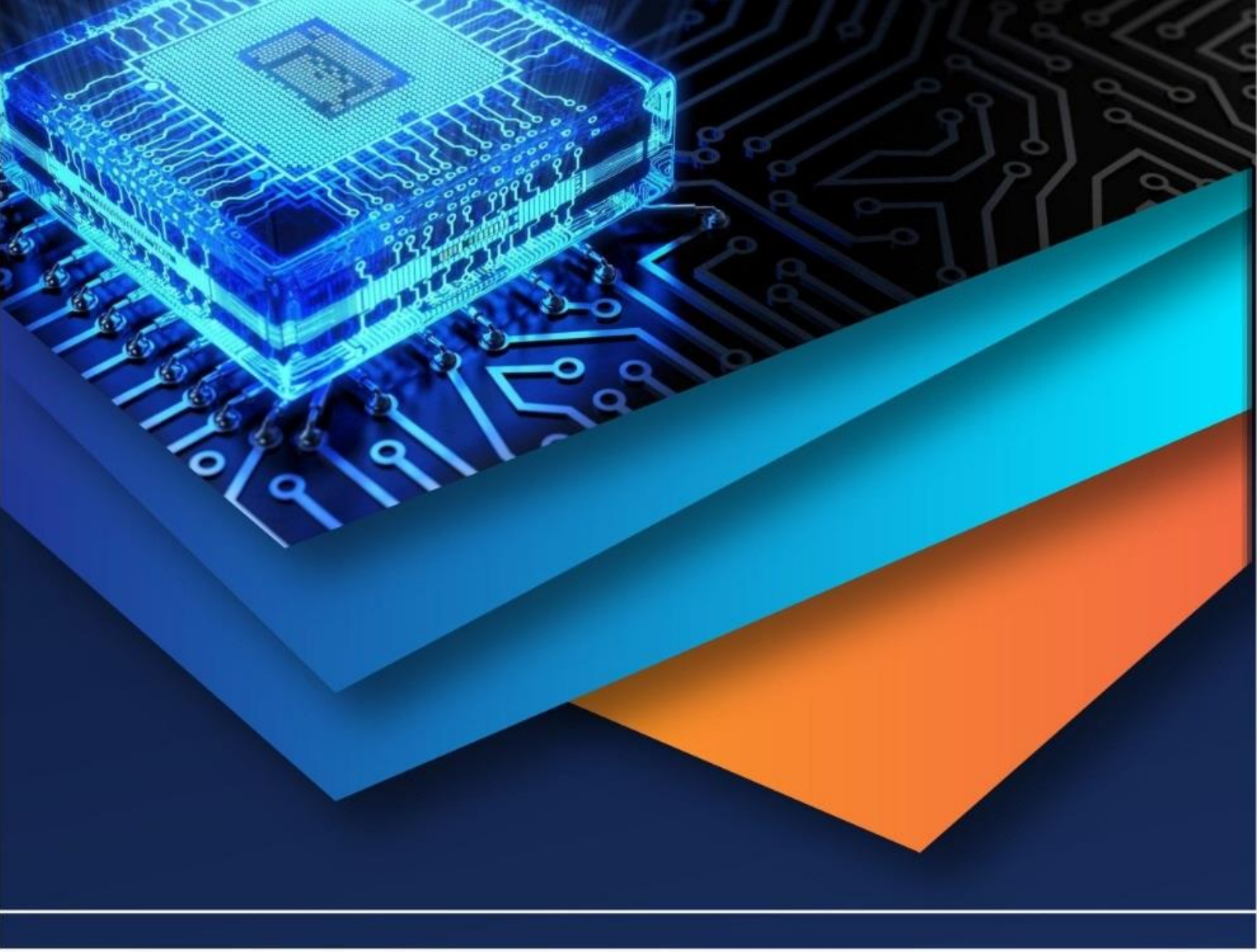

```
Console 1/A
dtype: int64
<class 'pandas.core.frame.DataFrame'>
Int64Index: 100021 entries, 419597 to 932927
Data columns (total 3 columns):
 #   Column    Non-Null Count   Dtype
---  ------    --------------   -----
 0   user_id   100021 non-null  int64
 1   movie_id  100021 non-null  int64
 2   rating    100021 non-null  int64
dtypes: int64(3)
memory usage: 3.1 MB
None
(80016, 3)
(20005, 3)
shape:  (6040, 6040)
predictions shape  (6040, 3952)
User-based CF RMSE:  0.11434636376977626
predictions shape  (6040, 3952)
Item-based CF RMSE:  0.1121515667609337

Recommendations for user {} based on user-based CF are the movies:
['Star Wars: Episode V - The Empire Strikes Back (1980)']
['Star Wars: Episode IV - A New Hope (1977)']
['Taxi Driver (1976)']
['Alien (1979)']
['Breakfast Club, The (1985)']

Recommendations for user 886 using SVD are the movies:
['Silence of the Lambs, The (1991)']
["Schindler's List (1993)"]
['Godfather, The (1972)']
['Usual Suspects, The (1995)']
['Forrest Gump (1994)']
```

VII. CONCLUSION

Recommendation systems are the best technology by retrieving other functionalities from the respective databases. These type of systems helps the users to discover the movies of their choices. These systems allows users to find the desired content. They are well-known for their efficiency and they became a necessary tool in E-commerce business. They have to satisfy users with the new released contents and should have the capacity to exhibit other additional features and additionaly they need to be attractive. New technologies should get involved here for efficient results.Therefore, to overcome this challenge, we proposed an approach to mix both explicit and implicit ratings to alleviate the data sparsity problem further in this aspect.

REFERENCES

[1] Schafer, J. Ben, Joseph Konstan, and John Riedl. 1999. “Recommender Systems In E-Commerce.” In 1St ACM Conference On Electronic Commerce, 158-166.
[2] Su, Xiaoyuan, and Taghi M. Khoshgoftaar. 2009. “A Survey Of Collaborative Filtering Techniques.” Advances In Artificial Intelligence 2009: 1-19. doi:10.1155/2009/421425.
[3] Melville, Prem, and Vikas Sindhwani. 2010. “Recommender Systems.” Encyclopedia Of Machine Learning.
[4] Xingyuan Li.2011 “Collaborative Filtering Recommendation Algorithm Based on Cluster”, International Conference on Computer Science and network Technology(ICCSNT), IEEE, 4: 2682-2685.
[5] Francesco Ricc, LiorRokach, Bracha Shapira.2011. Recommender Systems Handbook. NY:Springer, 1-35
[6] Sarwar, George Kaypi, Joseph Konstan and John Riedl.2000. "Application of Dimensionality Reduction in Recommender Systems -- A 6 Study." In ACM WebKDD Workshop.
[7] Sarwar, George Kaypi, Joseph Konstan, John Riedl.2000. “Analysis of recommendation algorithms for E-commerce.” In the Second ACMConference on Electronic Commerce, 158–167 42



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY