**A PROJECT REPORT**

**on**

**“MOVIE RECOMMENDATION”**

**Submitted to**

**KIIT Deemed to be University**

**In Partial Fulfilment of the Requirement for the Award of**

**BACHELOR’S DEGREE IN**

**COMPUTER SCIENCE**

**BY**

**N Pranav Reddy 2105637**

**B Bhanu Prasanth 2105881**

**Manish Kumar Singh 2105890**

**N Kaghuvahan Reddy 21052774**

**Vemana Sumanth 21052803**

**UNDER THE GUIDANCE OF**

**DR. PRADEEP KANDULA**

****

**SCHOOL OF COMPUTER ENGINEERING**

**KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY**

**BHUBANESWAR, ODISHA - 751024**

**May 2024**

KIIT Deemed to be University

School of Computer Engineering

Bhubaneswar, ODISHA 751024



CERTIFICATE

This is certify that the project entitled

“**MOVIE RECOMMENDATION**“

submitted by

**N Pranav Reddy 2105637**

**B Bhanu Prasanth 2105881**

**Manish Kumar Singh 2105890**

**N Kaghuvahan Reddy 21052774**

**Vemana Sumanth 21052803**

is a record of bonafide work carried out by them, in the partial fulfilment of the requirement for the award of Degree of Bachelor of Engineering (Computer Sci-ence & Engineering OR Information Technology) at KIIT Deemed to be university, Bhubaneswar. This work is done during year 2023-2024, under our guidance.

Date: 23/03 /2024

**DR. PRADEEP KANDULA**

Project Guide

**Acknowledgements**

We are profoundly grateful to **DR. PRADEEP KANDULA** of **KIIT University** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion

**N Pranav Reddy**

**B Bhanu Prasanth**

**Manish Kumar Singh**

**N Kaghuvahan Reddy**

**Vemana Sumanth**

# ABSTRACT

A typical recommendation system engine uses different algorithms to extract different types of data, examine the type of previous behavior of the user, and output the best matches of related items as needed. , and recommends the most suitable items based on that. Something that is currently in high demand. Another main reason for this could be to maximize profits. There are basically three ways he can recommend items to a user: Demographic filter that provides one or more general recommendations for the most popular movies based on their genre. Recommend movies to users if the demographics match. In general, popular movies are likely to be liked by more users. Another technique is content-based filtering, which takes into account people's interests and suggests movies, and his third technique used here is to search for the same people with the same characteristic interests, or the same pattern. It is a collaborative filtering method with movies are grouped together.

# INTRODUCTION

Recommendation systems are essentially filtering systems that predict a user's decisions and suggest more accurate results based on the user's previous preferences. We have been using various applications of these recommendation systems for many years, and they are now used on a variety of online platforms. The basic content of all these platforms is basically different types of movies such as: B. Romantic action thriller or e-commerce website: Social media platforms with professional websites like LinkedIn. For example, Instagram allows you to see previous stories in the feeds of people you follow. Here you can see that Instagram monitors your interactions and past activities with different people and can only infer other related stories about other accounts that have previously or currently done the same type of activity. Recommendation systems often improve the activities of a large number of users based on the activities they have tried through you. For example, if you buy a laptop or a mobile phone on Flipkart, you are advised to use a tempered glass case for your phone or buy a USB Type-C or Type-A adapter for your laptop. With steady improvements in recommendation systems, users will always receive better recommendations and will continue to improve throughout the 21st century to provide near-accurate solutions. If a dispute arises between electronic app music, music platforms, and educational institutions, you can simply refuse to use the app. Additionally, companies should focus on recommendation systems that are more complex than they appear. Whether you want to listen to music while gaming, traveling, running, or after a relationship argument, different users have different preferences and different choices based on different types of activities and moods.

## PROBLEM STATEMENT

## Recommendation systems are tools that aim to obtain user ratings, recommend movies from a large data set based on the user's matching interests, and classify them into different categories. The sole purpose of this entire recommendation system is to find content that suits the individual's personal interests and serves as a personal oasis. However, different factors are taken into account to create different content lists specific to different categories of individuals/users. Essentially, the AI-based algorithm used in the recommendation system creates a list of different possible device scenarios and finally matches all the interesting and relevant interests/choices in each category. All results are basically based on various activities you have done so far, such as: Previous searches that take into account things like what your profile looks like, what you've done in Chrome browser, Opera browser, and other browsers, and what you like your movies to be based on demographics and genre. A predictive model is created using available (large-scale) data, including history, and movies are protected by a list of 2000 movies. A large number of selected movies are recommended using different algorithms, different methods, and different similarity measures.

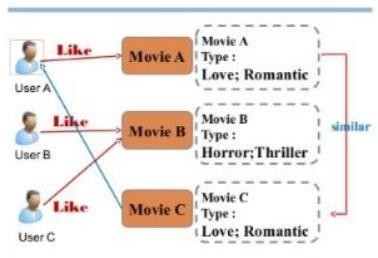
## OBJECTIVE

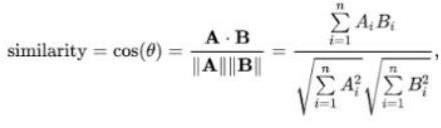
## Movie recommendation systems classify users with the same interests, provide a mechanism to search for content that may be of interest to different user groups, and create different types of lists to personalize based on content preferences. Provide interesting recommendations. The main goal of the recommendation system is to find a group of movies that each user prefers for a specific user group using approaches that suggest demographic filtering, content-based filtering, and collaborative filtering. Movies that are likely to be liked by the general user group are ultimately shown to the user by the recommender. It then uses different techniques to try to find users with different interests using information collected through different activities. Indian tests all users with the same interests with collaborative filtering to get a final set of movies that are individually recommended to the user. Therefore, we use different categories of recommended filtering techniques, compare the results obtained with different methods, and try to improve the results as the h dataset of a series of movies becomes increasingly large and exceeds the computational limit. This is generally a limitation for large data sets.

# METHODOLOGY

# **Content Based Filtering system**

# It uses content-based filtering techniques to compare different articles to users' interest profiles. Essentially, a user profile contains content suitable for use in the form of functionality. This takes into account previous actions or feedback, and typically takes into account descriptions of content edited by users of various selections. Suppose someone buys their favorite item "M" but it's sold out, so they need to buy item "N" on someone's recommendation. Because "N" has the same match characteristics as the first item. So this is simply content-based filtering as shown below.



Here, the number used to calculate the similarity between two types of films is the cosine similarity and calculate the score. Calculating the size of the score obtained by cosine similarity is very fast.

The steps involved in getting the movie recommendation are as below:

. Having the title find the index of that movie

. Calculate the cosine similarity scores for all the movies

. Arranging the scores in the order of highest priority first that is ascending order

. And then shorting the list based on the similarity scores.

. Getting the first 10 element of the list excluding the first one as it is the movie name in itself.

. Getting the top elements

By repeating the above steps, you will find the top movies that can get you the best possible recommendation based on distance. Ultimately, the recommender shows users movies that the general user base might like, and uses the information gathered through various activities to try to find users with different interests in other ways. Indian uses collaborative filtering to test all users with the same interests and get a final set of movies recommended individually to the user. Cosine similarity accounts for the angle between two vectors when the vectors are nonzero, and the dot product space is described as the scalar product of two vectors divided by the product of Euclidean quantities. In most cases, cosine similarity is used to obtain favorable recommendations for users.

This method simply uses the cosine distance between vectors, uses similarity to calculate a score, and then calculates the user's preference. For example, a movie where actors define the number of likes of users and only a small number of actors dislike a group of users. Therefore, we assume that a good sign angle is drawn between the user vector and the movie vector. This generally results in a large positive fractional angle close to zero, and there is a small cosine distance between the two vectors. In some sense, the metric is better because the distance between the film and the cosine is large, and the similarity of the cosine is lost. In this case, a new method call decision tree is used to improve the recommendation system. This method usually involves layers. Baby can apply some conditions to its classification approach to improve its recommendation system to determine which movies users want to watch or not at all.

The advantages of the content based filtering are:

. Can recommend the unrated items.

. Can recommend the movies based on the ratings of the user The disadvantages of the content based filtering are:

. Can't work on the new user hasn't red kidney movie act

. It can't make the user likes with the un -likes

**Collaborative based Filtering**

Content-based filtering has some limitations as it can only suggest movies that the user has a preference for and cannot recommend genres. However, systems based on collaborative filtering make it very complex to find datasets between users' similarities and their peers with similar interests. Cosine similarity or Pearson correlation is used to measure the similarity of user views. For example, in the following matrix, there is a user in each row of the column that corresponds to movies with the same similarity. There are also user ratings of various movies. Every movie has a target user.

For user-based filtering, all collaborative filtering is simple, but it also has drawbacks. The main challenge is that user choices change over time. Matrix precomputation solved the performance degradation issue. Therefore, item-based collaborative filtering can be used. This essentially considers items based on their similarity to the item and suggests that the same similarity coefficient can be used with Pearson correlation or cosine similarity when finding similar matches with the target user. Element-based collaborative filtering is the most static in nature. For example, he is the only user who has associated both The Matrix and Titanic, so there is only one similarity between them. These two different movies can also be very similar because they have millions of users and these two different movies rank the same for users who have rated both.

When performing collaborative filtering, try to find out which users are interested in names and similar settings. In this case, rather than using item characteristics to make recommendations, we classify users into clusters of similar characteristics and split each cluster in order of user preferences. You can also use cosine distance here. This considers users with the same interest, which is greater than a small cosine angle between two users. Here you can simply use a utility matrix and assign null values ​​to the sparse columns to make the calculation easier. Object-based collaborative filtering is generally preferred because it considers movies rather than the number of users, further simplifying the classification of movies and users. Therefore, user-based collaborative filtering is not recommended as it simply takes into account the users and ignores the sparse values ​​that cause problems in highlighting the recommender system's performance.

Now I would like to convert the recommendation problem into an optimization problem. The most preferred and common metric is the root mean square error (RMSE). The better the performance, the lower the RMSE value.

Advantages of collaborative filtering based systems are:

. It is simply content dependent

. It often reads the mind of people having same preferences

. Create real quality assessment of items.

Disadvantages of collaborative filtering based systems are:

. Early rater problem as the most common where the collaborative filtering method fails to provide ratings of the movie which has no user waiting.

. Sparsity problem is more common in this type of welding method where null values are in so much quantity that is difficult to find items which are rated by the majority of the people.

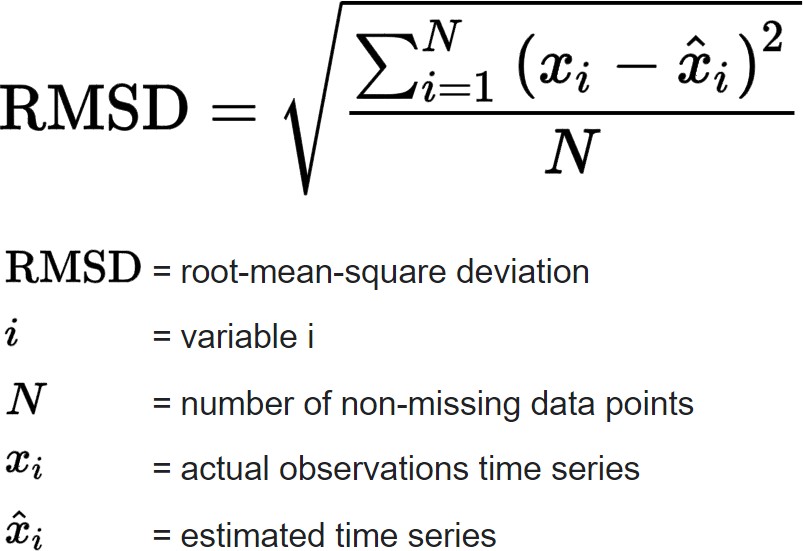
## Algorithms Used

### Cosine Similarity

Cosine similarity accounts for the angle between two vectors when the vectors are nonzero, and the dot product space is described as the scalar product of two vectors divided by the product of Euclidean quantities. The smaller the angle, the greater the similarity. Cosine similarity is much better than balanced distance because the angle is smaller in cosine similarity.

### RMSE (Root Mean Square Error)

### RMSE is essentially just the standard deviation of the prediction error. Residual error. The regression measure where the data points are located. However, it also shows the distribution of residuals within the data points and determines the best fit to the data. They are also used in prediction and regression analysis to find experimental residuals and validated results. The better the performance, the lower the RMSE value.



# CONCLUSIONS

# Therefore, when implementing a hybrid content and collaborative filtering technique, consider a hybrid approach to improve the overall system performance and recommend movies according to the user's selection in a much better way than the other two recommendation systems. To do. Reducing the average error further improves the accuracy of the recommender system, making it better usable in future applications. Although there are also systematic computational limitations and limitations for applying recommender systems to large datasets, we have made a good effort here to distinguish between various recommender systems and the final has a hybrid recommender system on top of everything. From this, we can conclude that hybrid-based filtering helps to make the system fragmentation more efficient and improve the overall system accuracy. And mixing both contents in a collaborative filtering technique ensures that even if one technique fails, the other technique maintains the overall accuracy of the system and simply improves the overall performance. There is no doubt that it means.