ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 10, Issue 3, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 6.551

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

Movie Recommendation System

Ms.Dhanya Shenoy[1], Sukanya S Naik[2] , Prajna S[3], Swathi K[4], Sachitha[5]

Asst. Professor, Department of Computer Science & Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Karnataka, India [1],

Student, Department of Computer Science & Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Karnataka, India[2,3,4,5]

ABSTRACT: There is already enough content available on the movie recommendation system. Showing the movie recommendations is essential so that the user need not waste a lot of time searching for the content which he/she might like. Thus, movie recommendation system plays a vital role to get user personalized movie recommendations. After searching a lot on the internet and referring to a lot of research papers, we got to know that the recommendations made using Content-based Filtering are using a single text to vector conversion technique and a single technique to find the similarity between the vectors. In this research work, we have used multiple text to vector conversion techniques and manipulated the results of the multiple algorithm You can think of it as a hybrid strategy that only employs the Contentbased Filtering method.The internet has increased the ability of various domains to interact and share important information during the past few years. Since everything has benefits and cons, the expansion of a domain is accompanied with information overload and challenges with data extraction. The suggestion system is crucial in addressing this issue. With its quick and logical suggestions, it improves the user experience. This study outlines a method that provides users with generalised suggestions based on the popularity and/or genre of a film. The implementation of the Content-Based Recommender System involves several deep learning techniques.This study also provides a glimpse into the difficulties that content-based recommendation systems encounter, along with our efforts to address them.

I.INTRODUCTION

Movies are a part and parcel of life. There are different types of movies some for entertainment, some for educational purposes, some are animated movies for children, and some are horror movies or action films. A recommendation system or recommendation engine is a model used for information filtering where it tries to predict the preferences of a user and provide suggestions based on these preferences.Movie Recommendation Systems helps us to search our preferred movies among all of these different types of movies and hence reduce the trouble of spending a lot of time searching our favorable movies. These systems have become increasingly popular nowadays and are widely used today in areas such as movies, music, books, videos, clothing, restaurants, food, places and other utilities.These systems collect information about a user's preferences and behavior, and then use this information to improve their suggestions in the future.A large number of companies are making use of recommendation systems to increase user interaction and enrich a user's shopping experience. Recommendation systems have several benefits, the most important being customer satisfaction and revenue.

II. LITERATURE REVIEW

1. In content-based recommendation, items are recommended which are similar to those provided by the user, whereas in collaborative recommendation users whose tastes are similar are identified to those of the given user and recommends items they have liked. Later with the evolution of the recommender system hybrid method has been invented which merges two or more techniques. Before the invention of the recommending system, people had to read reviews and choose the movie that best suited their interest or had to randomly choose any movie based on some other criteria. This became difficult as the number of movies that are available online started increasing rapidly.
2. De Campos et al. also made an analysis of both the traditional recommendation techniques. As both of these techniques have certain setbacks, he proposed another system which is a combination of Bayesian network and collaborative technique.
3. Kużelewska proposed clustering as an approach to handle the recommendations. Two methods for clustering were analyzed: Centroid-based solution and memory-based methods. The result was that accurate recommendations were generated.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

4. Luis M Capos et al has analyzed 2 ancient recommender systems i.e. content based mostly filtering and cooperative filtering. As each of them have their own drawbacks he planned a replacement system that may be a combination of Bayesian network and cooperative filtering. This paper so presents a replacement Bayesian network model to subsume the matter of hybrid recommendation by combining content-based and cooperative options.[5]shortcomings of collaborative filtering approach like sparsity problem or the cold-start problem. In order to avoid this issue, the authors have proposed a solution to use category information.The authors have proposed a movie recommendation system which is based on genre correlations. The authors stated that the category information is present for the newly created content.Thus, even if the new content does not have enough ratings or enough views, still it can pop up in the recommendations list with the help of category or genre information. The proposed solution is unbiased over the highly rated most watched content and new content which is not watched a lot. Hence, even a new movie can be recommended by the recommendation system. [6]Muthurasu, Nandini Rengaraj, K.C. presented a research paper that used TF-IDF and Cosine Similarity in Content-based Filtering for Movie recommendation system. Recommendation engines are trained to produce fast and coherent suggestions to users. This paper describes a hybrid video recommendation system using Term- frequency,Inverse document frequency technique for vectorization. For similarity measures Cosine similarity method is used. The system is presented to the user through a web-hosted user-interface. Despite small dataset, the system gives efficient and correct recommendations.

III. IMPLEMENTATION DETAILS

Content-based Filtering Systems (CBF based systems):

Items are suggested in content-based filtering based on comparisons between the item profile and the user profile. In the form of keywords, a user profile is content that has been determined to be pertinent to the user (or features). A user profile can be thought of as a collection of assigned keywords (terms, attributes), which an algorithm gathers from content the user deems relevant or interesting. The item profile is a collection of the item's attributes or keywords. Think about the situation where someone goes to a pastry shop to purchase his favourite cake, "X." Unfortunately, cake 'X' has already been consumed, so the shopkeeper advises the customer to purchase cake 'Y,' which uses components that are comparable to those of cake 'X'. A case of content-based filtering can be seen here. Advantages of content-based filtering are: λ They are able to suggest unrated things.λ By outlining the Content properties of an item, we can clearly illustrate how the recommender system functions. λ Users of content- based recommender systems simply need to rate the user in question; they are not dependent on any other system users. Disadvantages of content-based filtering are: λ As a content-based recommender analyses the user preferences and offers precise recommendations, it is ineffective for a new user who has not yet given any items a rating.
λ No suggestions for unforeseen goods. λ Limited Content Analysis: If the system is unable to discriminate between the items that a user loves and dislikes, the recommender will not function.

Collaborative filtering based systems(CF based systems): collaborative filtering system makes recommendations for things based on user and/or item similarity metrics. The system suggests products that similar users have previously preferred. This is based on a situation where someone asks his buddies with similar likes for movie recommendations. Advantages of collaborative filtering based systems: Based on user and/or item similarity metrics, a collaborative filtering algorithm suggests items. The computer proposes goods that users with comparable preferences have previously chosen. This is based on a scenario in which a person approaches his friends who share his interests for movie recommendations.

Disadvantages of collaborative filtering are: λ Early rater issue: Since there are no user ratings on which to build a prediction, collaborative filtering systems are unable to offer recommendations for new products. λ Gray Sheep: A group with similar traits is required for the CF-based system to function. It will be exceedingly challenging to promote people who do not consistently agree or disagree to these groups, even if such groups do exist. λ The "sparsity problem" is caused by the fact that there are typically far more items than users, which makes it challenging to locate items with a sufficient number of ratings.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

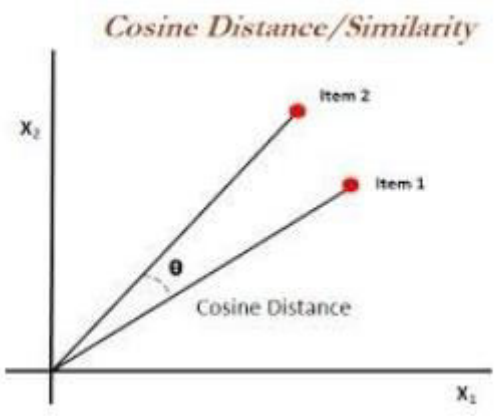
| Volume 10, Issue 3, May 2023 |

IV. RELATED WORK

The following are some typical methods used by recommender systems: Filtering based on content 2. Collective filtration

3. Using hybrid filters Content-Based Filtering, first Using the user's preferences as a filter, this method sorts the objects. It provides results based on previous user ratings. There are three ways to determine how similar two item vectors are: 1. Cosine similarity.

COSINE SIMILARITY: The angle of cosine between two objects is measured using cosine similarity. It makes a normalized comparison between two documents. Finding the dot product between the two identities will help.



The angle between v1 and v2 is shown in the diagram above as being. More similarity exists when the angle between the two vectors is less. In other words, if the angle between two vectors is small, they are practically identical to one another, and if the angle is big, the vectors are quite dissimilar to one another. Teamwork in Filtering It is dependent on users with comparable interests and outputs results depending on all users. User-based: In user-based collaborative filtering, it is assumed that a user will enjoy goods that other users who share similar tastes will enjoy. Item-based collaborative filtering is distinctive in that it anticipates users' preferences for things that are connected to earlier favorites. Blended Filtering Combining collaborative filtering and content-based filtering is known as hybrid filtering. It is currently the most often used and preferred method. It steers clear of any recommender technique's flaw. The suggestion mechanism is plagued by a few issues. As follows: Cold-start issue The unique feature of item- based collaborative filtering is that it anticipates users' preferences for items that are related to prior favorites. Mixture Filtering Hybrid filtering refers to the blending of collaborative and content-based filtering. It is now the approach that is chosen and employed most frequently. It avoids the drawback of any recommender system. There are several problems with the recommendation method. the following Cold-start problem The findings in this case don't quite match what was anticipated. Sometimes the system produces poor recommendations instead of successful outcomes. Additionally, coverage issues are always caused by data sparsity. Scalability: Scalability is another issue with recommendations. In this, the encoding of things proceeds linearly. When the data set is small, the system operates effectively. The recommendation algorithm struggles to provide reliable recommendations based on various movie genres as the data set grows. By using dimensionality reduction techniques like singular value decomposition, the scalability problem can be resolved (SVD). Additionally, it creates a fast and accurate result creation process. When several words have the same meaning, they are said to be synonyms. In this issue, the system occasionally is unable to distinguish between two words that are identical and cannot deliver the desired output. For instance, although the words "movie" and "film" have the same meaning, the recommendation system treats them differently and is unable to provide an accurate result. The synonymy problem can be resolved using Singular Value Decomposition (SVD), notably Latent Semantic Indexing. It became challenging to provide appropriate results in other recommendation systems because recommendations are based on the ratings and genres provided by other users. The term for this is collaborative filtering. This approach to creating recommendations is typical. We have used content-based filtering as an alternative recommendation mechanism to help users and get around this limitation. This approach makes use of ratings and genres provided by the user. Based on the movies the user has previously viewed, it makes recommendations. Because user ratings and genre selections are only based on the user's preferences and not on those of any other individual, this aids us in making appropriate choices. Because they rely on prior studies, they lack the accuracy of real recommendations. It is vital to note that suggestions can be based on a specific user and that they guarantee to provide recommendations that are on target because they depend on the preferences of a certain user. The suggested approach can produce effective outcomes and store a lot of data. Based on genre, our recommendation

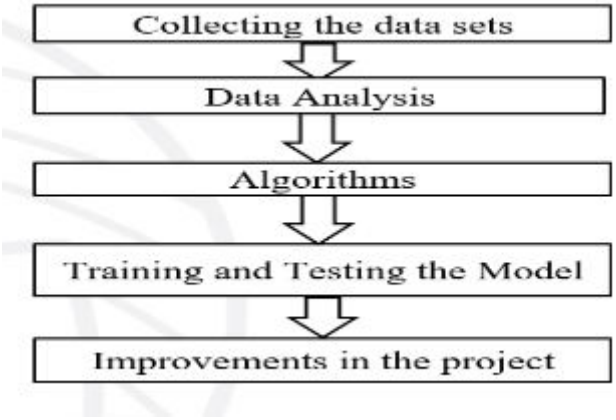
| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

engine looks for the finest films that are comparable to the one we just viewed. Data Collection: Collect a dataset of movie ratings, user reviews, and other relevant data.

Data Preprocessing: Clean and preprocess the data, including handling missing values, normalizing data, and removing outliers.

Feature Extraction: Extract features from the data that can be used to calculate similarity between movies. For example, you can use the genres, directors, actors, and other metadata associated with each movie.
Vector Representation: Represent each movie as a vector in a multi- dimensional space. Each dimension represents a feature of the movie. Cosine Similarity: Calculate the cosine similarity between each pair of movies using their vector representations. Cosine similarity is a measure of the angle between two vectors in a multi-dimensional space. The higher the cosine similarity between two movies, the more similar they are in terms of their features.
Top-N Recommendations: For a given user, find the N movies that are most similar to the movies they have rated highly in the past. This can be done by calculating the weighted average of the cosine similarity scores between the user's rated movies and all other movies in the dataset.
Model Evaluation: Evaluate the performance of the model using appropriate metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Precision, Recall, and F1-score.



V. RESULT AND DISCUSSION

We used Cosine Similarity and Content-based filtering in the deep learning framework recommendation system to anticipate our results and recommend a movie to the user by running the algorithm in Python using NumPy and panda libraries. A. Experiment Findings The formula used to calculate how similar two films are is based on how closely they share several characteristics. It displays the cosine of the angle formed by two vectors when they are projected onto a multidimensional space. Finding related objects is made easier because to the cosine similarity.

$$CosSim(x,y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}}$$

Fig. The formula for calculating cosine similarity is used to suggest movies.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

Fig. To illustrate the concept of cosine similarity, we will use the examples of two films from the adventure and comedy genres.

Fig: Cosine similarity

Fig: Cosine similarity

The degree of similarity between the two movies will depend on their angle theta. Theta values range from 0 to 1. Theta values close to 1 and 0 are most similar and close to 0 are least similar,respectively. If the score is close to 1, the movie will be suggested because else there would be no similarity between them. The user will receive the finest movie recommendations based on their Cosine similarity. We used a normalized popular score after the cosine similarity to obtain our function for computing distance. Then, using KNN functionality, we discovered the user's closest neighbour, who will be suggested

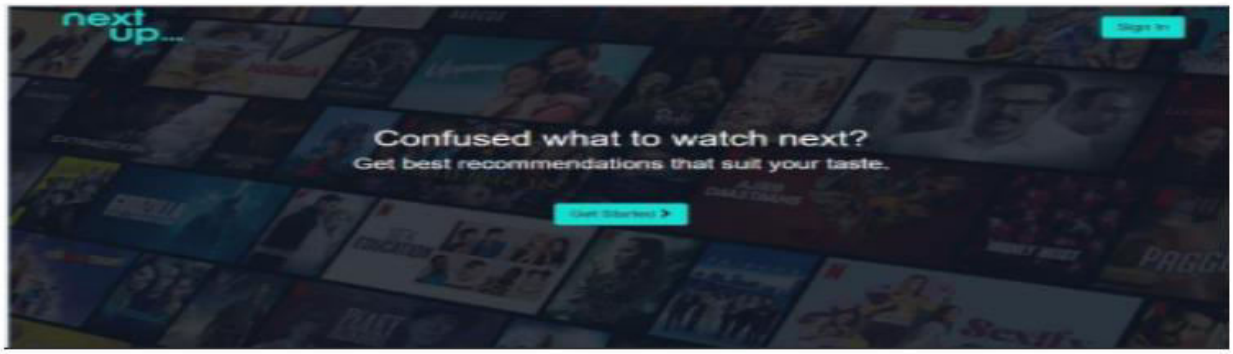


Fig 1: Main Page

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

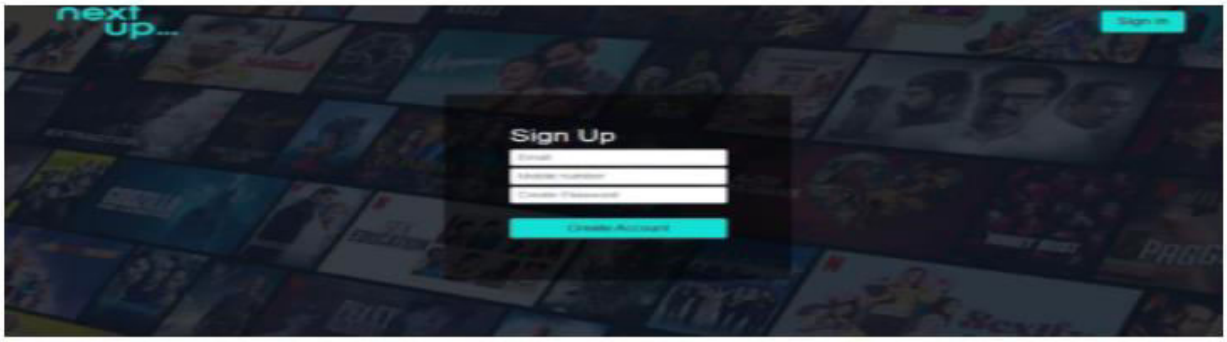


Fig 2:Registration Page

Fig 3:Login Page

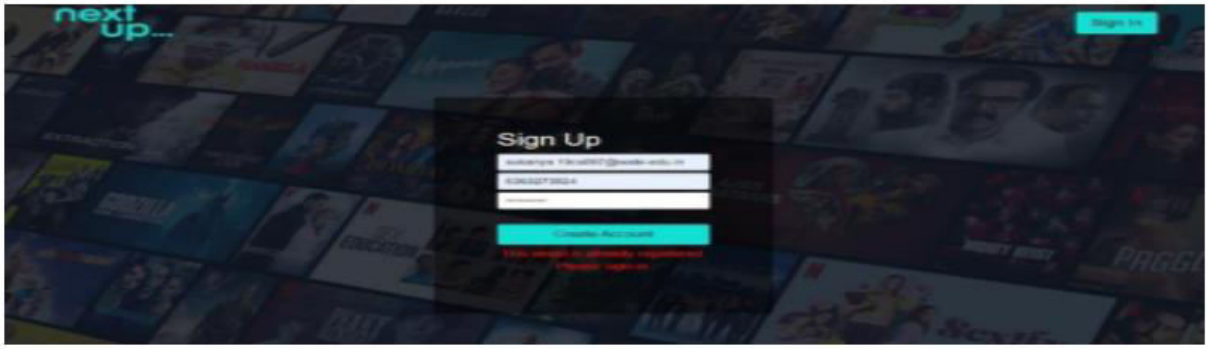


Fig 4:Error message showing this email id is already registered

Fig 5:selecting top 3 genres

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

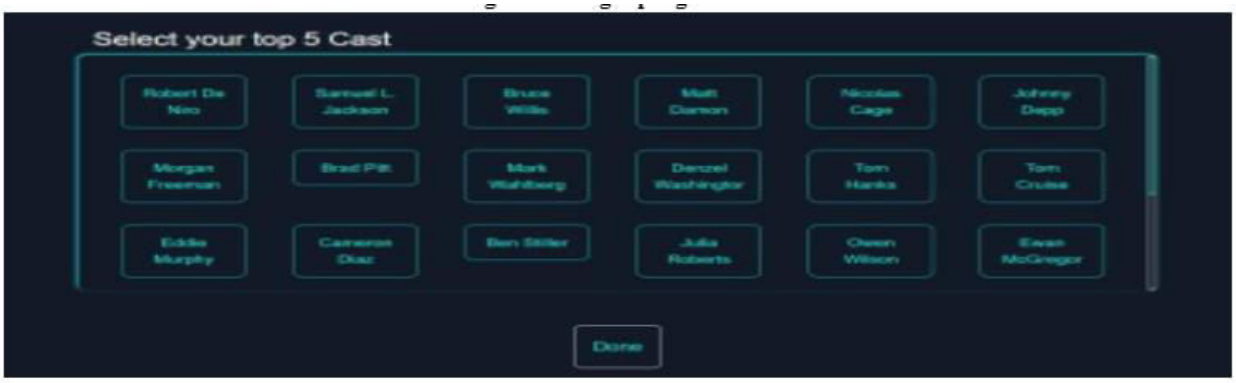


Fig 6:selecting top 5 cast

Fig 7:After the selection of a particular movie

Fig 8:Movie Review Page

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 6.551 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 10, Issue 3, May 2023 |

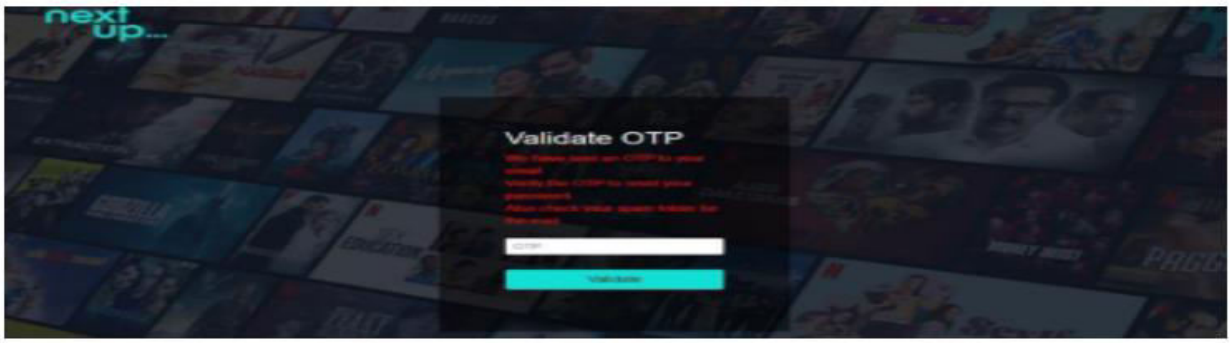


Fig 9:Incase if we forget the password it can be retained through OTP

VI. FUTURE SCOPE

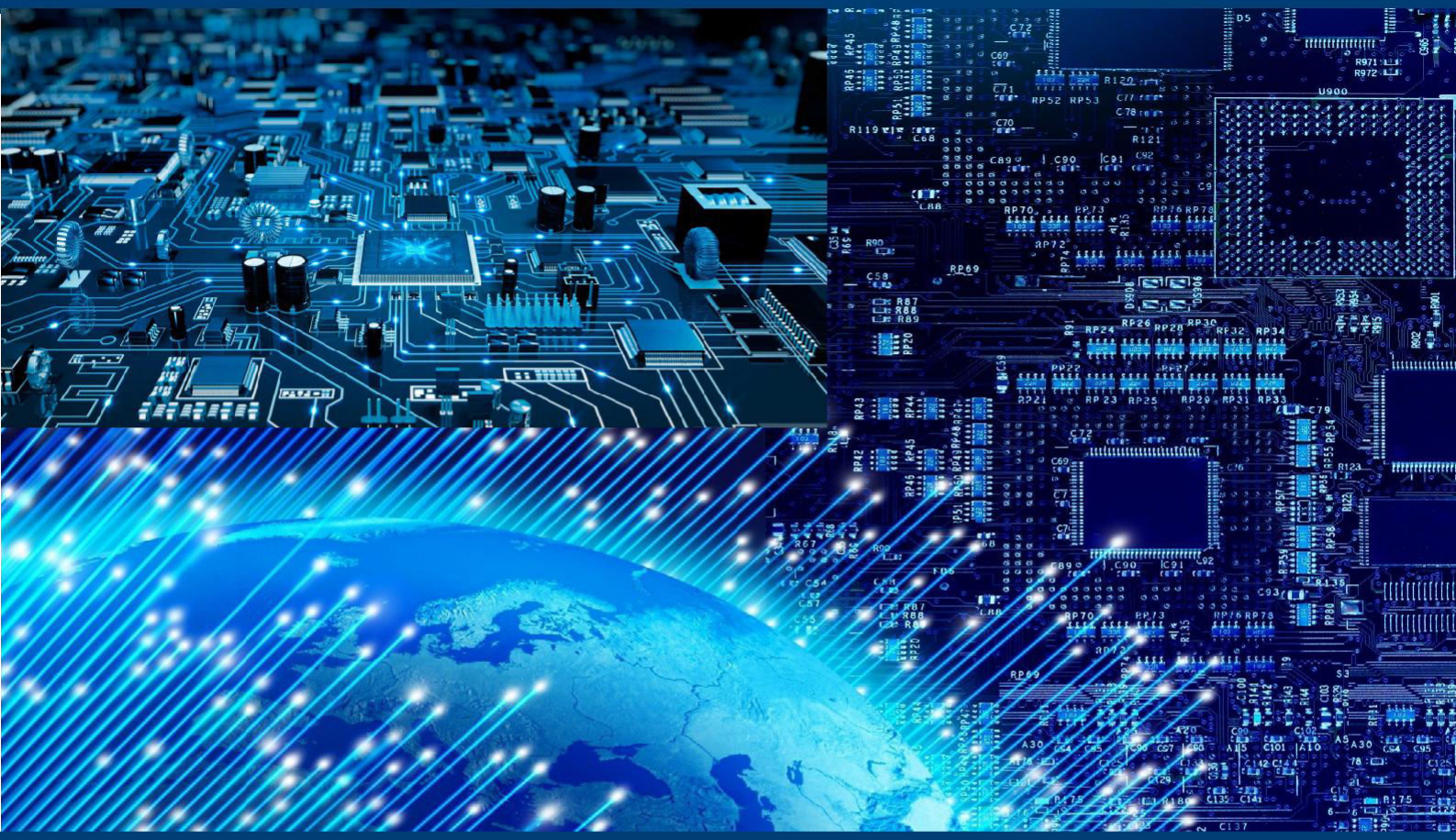
In the proposed approach, It has considered Genres of movies but, in future we can also consider age of user as according to the age movie preferences also changes, like for example, during our childhood we like animated movies more as compared to other movies. There is a need to work on the memory requirements of the proposed approach in the future. The proposed approach has been implemented here on different movie datasets only. It can also be implemented on the Film Affinity and Netflix datasets and the performance can be computed in the future.

VII. CONCLUSION

In this project, to improve the accuracy, quality and scalability of movie recommendation system, a Hybrid approach by unifying content based filtering and collaborative filtering; using Singular Value Decomposition (SVD) as a classifier and Cosine Similarity is presented in the proposed methodology. Existing pure approaches and proposed hybrid approach is implemented on three different Movie datasets and the results are compared among them. Comparative results depicts that the proposed approach shows an improvement in the accuracy, quality and scalability of the movie recommendation system than the pure approaches. Also, computing time of the proposed approach is lesser than the other two pure approaches.

REFERENCES

1. R. Vishnu Prasad, Ram Dantu, Aditya Paul and Paula Mears, “A Decentralized Marketplace
2. Application on The Ethereum Blockchain” in 2019 IEEE 4th International Conference on Collaboration and Internet Computing
3. Shangping Wang, Xu Wang, and Yaling Zhang, “A Secure Cloud Storage Framework with Access Control based on Blockchain”.
4. Constantinos Patsakis, Fran Casino, “Hydras and IPFS: a decentralised playground for malware” in the International Journal of Information Security.
5. Steichen, M., Fiz, B., Norvill, R., Shbair, W., & State, R. (2018), “Blockchain-Based,Decentralized Access Control for IPFS” in 2022 IEEE Confs on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics
6. Krishnendu Chatterjee, Amir Kafshdar Goharshady, and Yaron Velner, “Quantitative Analysis of Smart Contracts”
7. Trupti Bomble et al,”Android Based Complaint Management System for Municipal Corporation” in International Journal of Engineering Research and Applications ISSN :2248-9622, Vol. 5, Issue 4, (Part -3) April 2020, pp.64-66 .
8. Devika Radhakrishnan, Nisarg Gandhewar, Ruchita Narnaware,Prayas Pagade, Arpan Tiwari and Pooja, “Smart Complaint Management System” in International Journal of Trend in Research and Development, Volume 3(6), ISSN: 2394-9333 Nov-Dec 2019.
9. Fibina F.Maheen, Sumithra M.D ,“Development of Smart Complaint Portal based on Geotagging and Proximity Search”.
10. Sudeep J, Abhiram R, Adithya U S Vaidya, Rajath R Urs, Vallabh Joshi,” Smart Application for Complaint Registration” in International Research Journal of Engineering and Technology (IRJET) Volume: 06 Issue: 05 | May 2019.



International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |