Movie Recommendation System Using Hybrid Model

Amit Nandi[1], VHarshaSai[2], Mohit Maity[3], Vinay Kumar A[4], Sridhar R[5]

[1,2,3,4] *Student, Department of Information Science and Engineering, Global Academy of Technology, Bangalore*

[5]*Assistant Professor, Department of Information Science and Engineering, Global Academy of Technology, Bangalore*

Abstract— **In this era of fast emerging technologies and fast paced life, an individual pays a heavy toll everyday making them exhausted to continue the same the next day. In such scenario entertainment plays the role of life support. Entertainment is the tool which rejuvenates our lost energy and will to continue. There are many sources of entertainment, movie is one among them. For watching movies of our choice we make use of recommendation systems (RS), since manual searching of similar movies of our taste might take ample time sufficient enough to exhaust a user. Therefore, to further elevate the efficiency of movie recommendation systems, a hybrid model is needed. Hybrid approach can help getting the advantages from both content-based and collaborative filtering methods and eliminate their short comings as well.**

I.INTRODUCTION

In general, a Recommendation System is a tool for making decisions on behalf of the user. A RS analyses the data given by the user during their first encounter with the system and recommends the most likely/relevant in their next interaction. Organizations including YouTube, Hotstar, Netflix, Amazon, etc. use recommendation systems to ensure that their users can identify and get the correct product (videos/movies in case of above mentioned platforms) what they are looking for.ARS deals with a high volume of data/information as it ingests, already existing data on the web along with the data preferences of the user to accurately make decisions that the user might like. RS, on the time of data input from users, extracts the most relevant information, analyses users' preferences and interests and searches for a match between the user and the item for recommendation. The sole purpose of a movie recommendation system is to predict the possibility that a user might wish to watch a particular movie. These recommendations may or may not be custom

for every user, depending on the objective of each platform. Already existing movie recommendation systems make use of content based filtering to provide recommendations which has been working quite good so far. But the changing era and expectation of instant and accurate results are putting an urge to improve the efficiency of the existing one sand a hybrid approach is the only way possible to do so. As the hybrid approach has the ability to incorporate both content-based and collaborative filtering techniques improving the efficiency and accuracy of the system by providing optimal results and thus its need.

II. LITERATURE REVIEW

After the analysis of all the below papers we've identified the methodologies they've used along with the key findings. Walk through of the papers.

The authors of this paper gives information about content and collaborative based filtering, explains the drawbacks of content based filtering and also how to counter it. They also do propose an item based CF that improves the accuracy of a traditional Recommendation system (RS) using KNN and cosine similarity algorithm. KNN catches the k nearest neighbours to a target item. Cosine similarity calculates the similarity between 2 entities based on cosine angle between the entities. The proposed RS engine is different from traditional CF. Firstly after loading the movie and the rating dataset, the rating dataset is redesigned and movie dataset is filtered using the item base model. Then KNN is used to get k nearest neighbours for a target item and using the cosine similarity similar items is found. And if the user watches that movie then stop or else the priority list is changed and the whole process runs again. This improves accuracy of RS engine. The MEA (Mean

Absolute Error) values showed that item based CF had lowest MEA (0.248), compared to the content based and model based CF and user based CF when tested for performance. [1] The authors of this paper focus on item based collaborative based filtering to build a RS. They discuss the strengths of the user based CF model and also their disadvantages like sparsity and scalability of data and how item based collaborative is used to counter it. A brief introduction to Grouplens made recommendation and Ringer is given. For the implementation of proposed item based CF movielens and ratings dataset are used which is provided by Grouplens. For user-item based collaborative filtering, RS is based on the ratings given by the user and if the user doesn't give the ratings the weight matrix formula is used to predict the user ratings. For item-item based CF, cosine similarity matrix is used. The results are compared and the authors of the paper found out that accuracy is lower compared to the other models but the focus had been providing real time recommendations in which the model stood out. [2] There exist many algorithms to determine the neighbourhood for collaborative filtering like KNN and Grouplens. This paper proposes use of fuzzy clustering algorithm for getting the neighbours. In this paper the performance of fuzzy clustering for spherical data is compared with the grouplens. For the grouplens algorithm the weight of similarity is found using Pearson coefficient. Three fuzzy clustering for spherical data are discussed i.e. BFCS, QFCS, and KLFCS for the research. However proposed RS engine involves using grouplens after the clustering. For a movielens dataset the proposed algorithm was used with 3 fuzzy clustering algorithms and based on the AUC values it was found that they all outperform the traditional methods in terms of accuracy. [3]In this paper the focus is on the item based collaborative filtering and association rule mining and to build a baseline recommender system using them. Amazon review data for mobile and accessories is used for implementation. Item based collaborative filtering is the most relevant form of filtering technique used, which is used in leading e commerce websites like amazon. The proposed approach involves using both apriori algorithm and correlation matrix on the pre-processed dataset to obtain accurate recommendations. The authors also discuss about use of recommendation system in the COVID 19 angle for recommending masks, PPE kits etc. [4]This paper proposes a book recommender system using deep learning technique and k-Nearest Neighbours (k-NN) classification. This paper considers application of Machine Learning Technology and presents a different approach based recommender system. This paper also explores different dimensions of a recommendation system by analysing multiple literature review papers. To compute the similarity of users' cosine similarity algorithm is used. This model can be used to test various recommender algorithms for performance evaluation. The knowledge gained from this paper can be used to solve other recommendation problems and to improve quality of prediction and to take knowledgeable decisions in future. [5]The paper discusses the use of genetic algorithms in recommendation systems. The GAs have been incorporated into RS in three ways: clustering, hybrid models, and using the GAs without relying on additional information obtained by hybrid models. In general, the GA provides better results and faster results than other traditional metrics due to the selection of the optimal similarity function. A traditional recommendation system is based on one criterion or one utility function. This research work developed a multi-level filtering technique, BLIGA. BLIGA provides recommendations based on the GA approach. The BLG analyzes the best recommendation list rather than looking for the best items to comprise the recommendation list. It is possible to address two limitations of this study in future research. One limitation arises from the specifications of the machine used in the experiments, which made it impossible to test the method on large datasets. The second limitation is caused by the limited use of the genre feature of the items for a semantic feature. As a result, other features may be considered and tested in the future.[6]The recommendation system for patients/dieticians is implemented for the purpose of encouraging the patients to take nutritional supplements, to consume food that are considered better to meet their healthy needs, tastes, and dietary preferences. This paper discusses the nutritional information modeling from a cloud system and its implementation into diet recommendation systems based on patients' nutritional status, patient's health assessment of foods, contradicts current nutritional theories. Therefore, future food/diet-based

recommendation systems must address different types of nutritional errors. It describes patients relying on medicines instead of resorting to precautionary nutritional measures in obtaining their nutritional needs. Selecting healthy diet based on nutritional needs has been a serious issue. Insufficient and incorrect food consumption has been identified as leading causes of several health challenges and ailments. The main reason for this is the vast variety of products and information about healthy diets. Studies have shown that a robust diet recommended by a Dietician or by a medical diet based Artificial Intelligence automated cloud system can increase longevity, protect against further diseases, and improve quality of life overall. [7] Based on online binary purchase data, the paper proposes an approach to help e-commerce companies to select the good collaborative filtering algorithms for automated recommendation generation. In order to construct the framework, an experimental design is applied to binary purchase data sets with distinct input data characteristics characterized by distinguish reduction of data techniques, CF techniques and particular measures for several CF configurations. Based on their binary purchase data sets, e-commerce companies can determine the optimal configuration of this framework. As a result, they are able to assess how modified input data sets affect the preferred algorithm configuration. Prior studies have not addressed the combination of characteristics of both input and required algorithmic configurations. A preprocessing step measuring the efficiency of a data reduction technique is determined by the input characteristics. The best data reduction technique in terms of accuracy is CF. [8] The paper uses clustering as a method to develop a recommender system. The algorithms used by recommendation systems were usually developed based on content-based filtering, associative rules, multi-model ensembles, and collaborative filtering. In order to train the prediction model using multi-model ensemble algorithms, a lot of real-world data is necessary to create a personalized recommendation system. Finding association rules and determining levels of dependence among rules is done using the Apriori algorithm. For multi-model RS, multiple classifiers are usually used. Based on clustering, this paper proposes a book recommendation system that uses a variety of approaches, such as hybrid,

knowledge-based, and utility-based. Rearrangement of books based on user choice and review ratings are done through clustering. The ultimate target of this approach is to provide personalized recommendations with improved accuracy. [9] An assessment of the film recommendation system based on a literature review is presented. Distinguishing different collaborative filtering techniques is done by comparing their performance in multiple contexts and not only under favorable conditions. It analyses the usage of precision and accurate based algorithms with minimal computational time. The combination of both content and collaborative filtering techniques are used and measured. Future studies in various fields could benefit from the use of new algorithms and the applications of Machine learning algorithms for recovering the conceptual knowledge on the website could be particularly valuable. The paper discusses collaboration as the most common approach, which is based on a dataset divided into two subsets: training and assessment. This technique seems ideal for handling large quantities of data in this field. [10]So according to the growing demand for accurate results or needs, Recommendation system came into picture, but we generally don't get exactly what we want as the sources are also of various type. It is very difficult to address the particular needs of every individual, so to fulfill the need of particular search the discovery of recommendation system happened in the middle of 1990's. And in that period of time it was one of the major steps toward research on the field of recommendation. Recommendation system basically recommend individual what fits into their system of thoughts and is very much probable to their needs and requirements. Recommendation is software which can be separate helping software or integrated with websites or any other software which needs suggest probable things every individual. This paper is focused on collaborative filtering technique. Which also part of recommendation system where it suggests individuals their subject to interest products to them by their own rating and according those whose rating are at strong similarity to them. But the major limitation is the cold start problem as collaborative filtering is purely focused on the similarity of interests of the neighbors. One way we can resolve it by is taking a step toward hybrid approach. [11] To overcome the challenges that were

related to collaborative filtering is discussed in this paper. Basically, the main problem regarding collaborative filtering was accuracy of the prediction and the recommendations made to the users. To resolve this issue researchers come with Memory-based-model methods which collect their data to provide personalized user experience and it further divided into two categories one is UCF (user based collaborative filtering) and other one is ICF (item based filtering) UCF recommend product on the basis of similarity of user and ICF recommend product on the basis of similar product user had used before. Somehow it took great step but on the later phase researcher got to know that it also fulfilling the accuracy prediction criteria. So, they came with a great idea of hybrid design where UCF and ICF and linear regression model were combined which known as rUICF which later shows great performance. [12] This paper discusses about the recommendation system which is basically made and personalized for the fashion retail e-commerce users which basically recommend the fashion product to the user which is specific only to them. So this recommendation system was applied to the one of the biggest Korean company that is basically fashion based company that deal with fashion products both on online and offline market. The company focuses on two main things first is, on online and offline the product is same and second is product changes according to the demand and season. As there is so much complexity in the system a new method was introduced which basically the combination of both item based collaborative filtering algorithm and fulfilling the seasonal demand characteristics later called as K-RecSys which is extension of collaborative filtering. And records

show that it has given boost in selling product and made huge profit. [13] As research has always been an integrated part of these recommendation systems just to make sure that user always get a satisfactory experience and should get suggested whatever they want to see or get. This paper is highly focused on enhancing the previously existing system such as collaborative filtering technique which basically depends on user based similarity and item based similarity to review based collaborative recommendation. So in the new system the comments sentiments and reviews are processed through text mining and exploit them and according to that recommendation are made especially for the particular user. Although results are good from review based collaborative filtering, but further things are going to get more complex and vast so for that more work on research area is needed. [14] This paper is proper study on recommendation systems, as we know recommendation system recommend users through their previous data or through similar user experiences and through relevant items data. But we know the speed at which internet is growing and the volume of data that is present on the internet and different types of data. So, there is so much data present on the internet which confuses the users, and they go unsatisfied which leads to information overloading. And in this point of time Recommendation system comes and resolves this issue. And provide users their personalized experience. This paper also discusses about positive and negative of different kinds of recommendation system and serves as compass in the field of research. [15]

III. LITERATURESUMMARY

Table 1:Summary of Literature Survey

Sl.No.	Author Names	Year of Publication	Methodology	Findings
[1]	Meenu Gupta, Aditya Thakkar, Vishal Gupta, DhruvPratap SinghRathore, Aashish	2020	Collaborative filtering	The MEA (Mean Absolute Error) values showed that item based CF had lowest MEA compared to the content based, model based CF and user based CF when tested for performance.
[2]	MukeshKumarKharita, Atul Kumar,PardeepSingh	2018	Collaborative filtering	For user-item based collaborative filtering, RS is based on the ratings given by the user and if the user doesn't give the ratings the weight matrix formula is used to predict the user ratings. For item-item based CF, cosine similarity matrix is used. The results are compared and it is found that the accuracy is lower compared to the other models but the focus had been providing real time recommendations in which the

				model stood out.
[3]	Tadafumi Kondo, YuchiKanzawa	2018	Collaborative filtering and Clustering	For a movielens dataset the proposed algorithm was used with 3 fuzzy clustering algorithms and based on the AUC values it was found that they all outperform the traditional methods in terms of accuracy.
[4]	Jessica Lourenco and AparnaS. Varde	2020	Collaborative filtering	Reviewinganddiscussingaboutassociationrulesintextminingan dapriorialgorithm.
[5]	Rajesh Kumar Ojha and Dr.BhagirathiNayak	2018	Collaborative filtering	Analysis of KNN algorithms for recommendations.
[6]	Bushra Alhijawi, YousefKilani	2020	Collaborative filtering	Analysis of multi-filtering levels of CF techniques.
[7]	Celestine Iwendi, SulemanKhan, JosephHenry Anajemba, AliKashif Bashir, Fazal Noor	2019	RNN,LSTM,GRU	Analysis of different MLalgorithms-RNN, LSTM, GRU.
[8]	StijnGeuens, KristofCoussement, and KoenW.DeBock	2018	Collaborative filtering	Reviewing and comparing CF ondifferentmethodsmodel-basedCF, demographic-basedsystems.
[9]	DhimanSarma, TanniMittra, MohammadShahadatHos sain	2021	Clustering, K-means	Reviewing and analysis ofClustering&K-meansalgorithm.
[10]	Sri Hari Nallamala, UshaRani Bajjuri, SarvaniAnandarao, Dr. D. Durga Prasad and Dr.Pragnaban Mishra	2020	Collaborativefiltering & Contentbasedfiltering	Reviewing and comparing ContentandCollaborativefilteringte chniques.
[11]	Sura I. Mohammad AliandSadiqSahip Majeed	2021	Collaborative filtering	ReviewingandanalysisofKNNalgori thm.
[12]	ManjaiahDoddaghattaHu chaiah, Mohammed FadhelAljunid	2020	HybridModel	UCF and ICF algorithm are used toovercome the limitations ofCollaborativefiltering.
[13]	HyunwooHwangbo, Yang Sok Kim, KyungJinCha	2018	Collaborative filtering	K-RecSys which is extension of collaborative filtering has given boost in selling product and made huge profit.
[14]	Mehdi Srifi, AhmedOussous, AyoubAitLahcen and SalmaMouline	2021	Collaborative filtering	Reviewinganddiscussingcollaborati vefilteringanditscategories: Memory based & modelbasedCF.
[15]	F.O.Isinkaye, Y.O.Folajimi,B.A.Ojoko h	2018	Collaborativefiltering & Contentbasedfiltering	ReviewingandcomparingtheConten tandCollaborativetechniques andmemorytechniques.

IV. METHODOLOGY

The main objective to build a movie recommendation system using a hybrid approach is to enhance the recommendation metrics of existing systems and provide optimal results. The proposed movie RSis expected to generate a finer quality and accuracy than the existing ones. The system emphasizes on reusing the information and preferences provided by the users that can be used in the prediction of future recommendations. Through this project we are proposing a RS which provides recommendation based on the gathered information from the web and the information given by the users. It can be achieved by analyzing user's psychological profile, their watching history, movie ratings, etc.

V. CONCLUSION

To ensure the scalability, elevate the accuracy and enrich the quality of movie recommendations, we are proposing a hybrid model which unifies both content-

based and collaborative filtering methodologies. This model can be used to optimize the decision making ability of the system. Hybrid recommendation system merges the preferences of the users with similar users who rated the movies along with the data available on the web to accurately recommend movies.

REFERENCE

[1] MeenuGupta, AdityaThakkar, VishalGupta, DhruvPratap Singh Rathore, Aashish, "Movie Recommender System Using Collaborative Filtering", 2020.

[2] MukeshKumarKharita, Atul Kumar, Pardeep Singh, "Item-Based Collaborative Filteringin Movie Recommendation in Realtime", 2018.

[3] TadafumiKondo,YuchiKanzawa,"PerformanceComparisonofCollaborativeFiltering Using Fuzzy Clustering for SphericalData", 2018.

[4] Jessica Lourenco and Aparna S. Varde, "Item-BasedCollaborativeFilteringandAssociationRules foraBaselineRecommenderinE-Commerce", 2020.

[5] Rajesh Kumar Ojha and Dr. Bhagirathi Nayak, "Application of Machine Learning in Collaborative Filtering Recommender Systems", 2018.

[6] BushraAlhijawi, YousefKilani, "A Collaborative Filtering Recommender System using genetic algorithm", 2020.

[7] Celestine Iwendi, SulemanKhan, JosephHenry Anajemba, AliKashif Bashir, FazalNoor, "Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System through Machine LearningModel", 2019.

[8] StijnGeuens, KristofCoussement, and Koen W.DeBock, "A framework for configuring collaborative-filtering based recommendations derived from purchase data", 2018.

[9] DhimanSarma, TanniMittra, Mohammad ShahadatHossain,"PersonalizedBookRecommen dation System using Machine Learning Algorithm",2021.

[10] SriHariNallamala,UshaRaniBajjuri,SarvaniAnan darao,Dr.D.DurgaPrasadandDr.PragnabanMishra ,"ABriefAnalysisofCollaborativeandContentBas edFilteringAlgorithmsusedinRecommenderSyste ms",2020.

[11] Sura I. Mohammad Ali and SadiqSahipMajeed,"AReviewofCollaborativeFilt eringRecommendationSystem", 2021.

[12] ManjaiahDoddaghattaHuchaiah,MohammedFadh elAljunid,"Anefficienthybridrecommendationmo delbasedoncollaborativefilteringrecommendersys tems",2020.

[13] HyunwooHwangbo, Yang Sok Kim, Kyung JinCha, "Recommendation system development for fashion retail ecommerce", 2018.

[14] Mehdi Srifi, Ahmed Oussous, AyoubAitLahcenandSalmaMouline,"Recommen derSystemsBasedonCollaborativeFilteringUsing ReviewTexts",2021.

[15] F.O.Isinkaye,Y.O.Folajimi,B.A.Ojokoh,"Recom mendationsystems:Principles,methodsandevaluat ion", 2018.