IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 7.301
Volume 3, Issue 1, March 2023

Video Recommendation System using Collaborative Filtering - A Systematic Review

Krupali Bhimani
Department of Computer Science
Mithibai College, Mumbai, Maharashtra, India
University of Mumbai, Mumbai

Abstract: *Recommendation System is basically an algorithm that recommends useful content like videos, images or information to people. Like, choosing which movie to watch, which item to purchase, which book to read, etc. based on user's actions on previous views. There are various algorithms used to perform this recommendation technique. This paper reviews different algorithms of recommendation system used in different applications where video/visual broadcasting is involved and it concludes how collaborative filtering is the most commonly used method out of all.*

Keywords: Recommendation System, Collaborative filtering

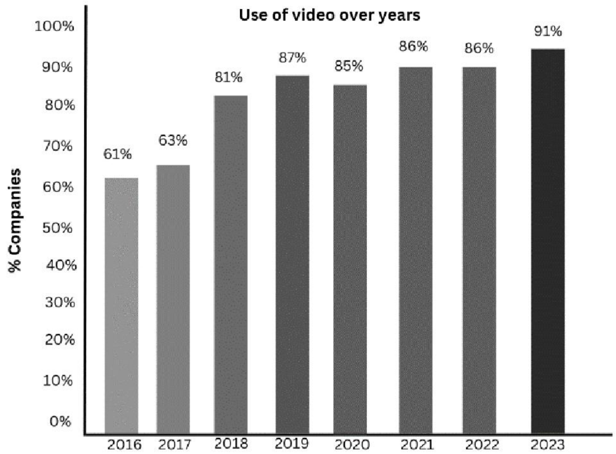
I. INTRODUCTION

Recommendation System acts like a Filter in Huge Content of Information available in any application. It fetches data from the user and performs according to user's preferences. It suggests user with some relevant content viewed by them. The applications such as YouTube, Instagram, Facebook and many other social media platforms deliver huge amount of data every second. Recommendation System plays a vital role in helping user in suggesting data that the user is seeking for. The data from collection can be based on video, audio, article, pictures and more. The well-known companies like Instagram, Google, Prime Video, YouTube, etc. uses recommendation system to offer data(videos) that the user is seeking for or would like to watch based on the algorithm that have learned about their likes and dislikes by their response to a particular video. The algorithms are based on Machine Learning. The most common and used technique is Collaborative Filtering. Algorithms based of Neural Network and Deep Learning are used to increase the efficiency and accuracy of the video suggestions. Over the web, there has been a notable change in the quantity of data. It is obvious that searching information in huge content will be difficult. Due to which the urge to develop algorithms to find expected results faster, recommendation system has come into the picture and has been growing since then. Hybrid recommendation system are widely used as traditional methods has their limitations. While building hybrid recommendation system, integration of NN and Deep learning algorithm has introduced. With the increasing growth of huge amount of user data generated and collected each second. How to make effective use of this data is a big question that arises. The data can be of different types like videos, images, information, audio, etc. The idea of videos that is moving pictures has taken over the whole online platform. Unto 2019, Video traffic over internet was 80%. Videos are everywhere and on every platform over internet in this era. It is the most preferred way of content by the people. As video makes tasks easier which doesn't require any sort of reading or scrolling. The idea of videos has become the heart of the content these years. Videos make audience know what they actually want or expect from the market. Here the recommendation system comes into the picture. The only reason of this popularity in videos is that it shows what you would have thought or expected. It helps to grow business creatively and effectively.

Recommendation System acts like a Filter in Huge Content of Information available in any application. It fetches data from the user and performs according to user's preferences. It suggests user with some relevant content viewed by them. The applications such as YouTube, Instagram, Facebook and many other social media platforms deliver huge amount of data every minute. Recommendation System plays a vital role in helping user in suggesting data that the user is seeking for. Recommendation System is basically an algorithm that recommends useful content like videos, images or information to people. Like, choosing which movie to watch, which item to purchase, which book to read, etc. There are various algorithms used to perform recommendation techniques. This paper reviews different algorithms used in video recommendation system. In Huge amount of data, recommendation system plays a vital role and is very helpful to

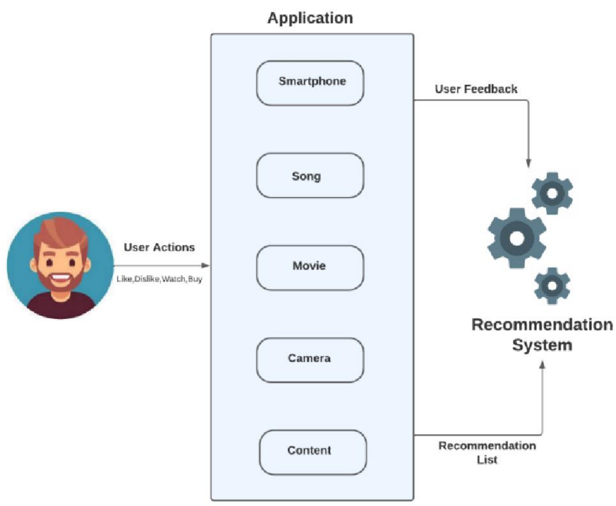
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

get the expected outcome of user's satisfaction. The data from collection can be based on video, audio, article, pictures and more. The well-known companies like Instagram, Google, Prime Video, YouTube, etc. uses recommendation system to offer data(videos) that the user is seeking for or would like to watch based on the algorithm that have learned about their likes and dislikes by their response to a particular video. The algorithms are based on Machine Learning.



Growth rate of video use over time

The most common and used technique is Collaborative Filtering. Algorithms based of Neural Network and Deep Learning are used to increase the efficiency and accuracy of the video suggestions. Over the web, there has been a notable change in the quantity of data. It is obvious that searching information in huge content will be difficult. Due to which the urge to develop algorithms to find expected results faster, recommendation system has come into the picture and has been growing since then. Hybrid recommendation system are widely used as traditional methods has their limitations. While building hybrid recommendation system, integration of NN and Deep learning algorithm has introduced.

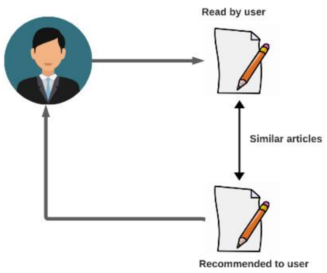


Content-Based Filtering

It employ's the contents of the objects to find similarities between different items. Once a significant number of products that one person has previously demonstrated favor for have been analyzed, the user interests profile is established. The recommender system may then use this profile to search a database and choose the right items. These algorithms fail to identify consumer preferences based on item contents. For instance, a recommender system would first compile a reader's whole library of previously read books before examining their contents and making new article recommendations. It makes suggestions for related products that a user would appreciate based on their input on prior jobs (including likes and dislikes). It does the same using features.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 1, March 2023
Impact Factor: 7.301

There are two types of Content-Based Filtering: one is cosine distance and classification approach. This model doesn’t require data of other users as it is specific to a particular user only. It captures interests of the user and recommend items accordingly.

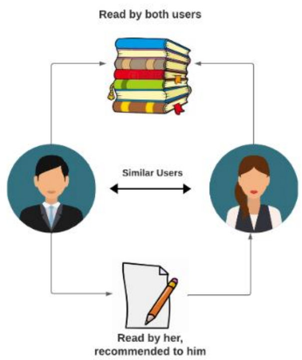


Collaborative Filtering

CF only takes into account each user's item ratings. CF is designed to make suggestions when accurate knowledge about the users and things is missing. Also, it successfully addresses the issue of over-specialization, these are very common drawbacks. Based on possible customer preferences for products from related user interests. It is founded on users' shared tastes. Also called social filtering. By comparing people's tastes, this system analyses user data to provide personalized recommendations. Since collaborative filtering doesn't require subject expertise, using this technique is simpler.

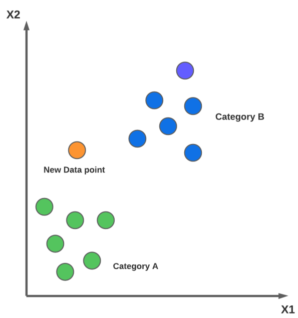
It contains two types: One is based on users and other is based on items.

Both user-based and item-based metrics assess similarity between two users and two items, respectively.



K-Nearest Neighbor (KNN) Algorithm

It is one of the simplest supervised algorithms that assumes similarity between new and available cases. Is is used for regression and classification problems. It is not parametric and also known as lazy learner algorithm. It helps to classify the category of new data point.



IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 7.301

Volume 3, Issue 1, March 2023

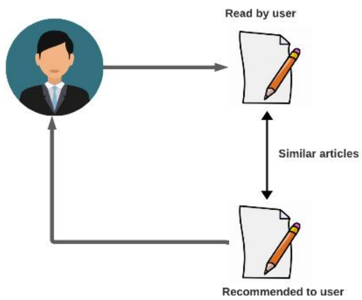
II. POPULAR RECOMMENDATION SYSTEM ALGORITHMS

Algorithms	Description	Advantages	Disadvantages
Content-Based Filtering	It examines the similarities in the content that item carries to return recommendation.	Makes easy recommendation for new items Based on similarity.	-Interest of Users are considered from their history to predict recommendations.
Collaborative Filtering	Based on ratings given by Users.	-Identifies User's Interest. -Simple Approach to handle complexity	- Restarts with regard to relatively newer items
Knowledge-Based Filtering	Based on information taken from User.	-Makes personalized recommendation for User's	-Can lead to inaccurate recommendations
Deep Neural Network	-Learns from human interaction and works accordingly.	-Can work for more than one task at a time. -Capable of machine learning models.	-It is dependent on Hardware.
Ontology-Based Recommendation	Based on properties, entities and attributes of an item.	-Enhances the performance of other algorithms. -Better technique.	-Complicate Algorithms

Content-Based Filtering

It employ's the contents of the objects to find similarities between different items. Once a significant number of products that one person has previously demonstrated favor for have been analyzed, the user interests profile is established. The recommender system may then use this profile to search a database and choose the right items. These algorithms fail to identify consumer preferences based on item contents. For instance, a recommender system would first compile a reader's whole library of previously read books before examining their contents and making new article recommendations. It makes suggestions for related products that a user would appreciate based on their input on prior jobs (including likes and dislikes). It does the same using features.

There are two types of Content-Based Filtering: one is cosine distance and classification approach. This model doesn't require data of other users as it is specific to a particular user only. It captures interests of the user and recommend items accordingly.

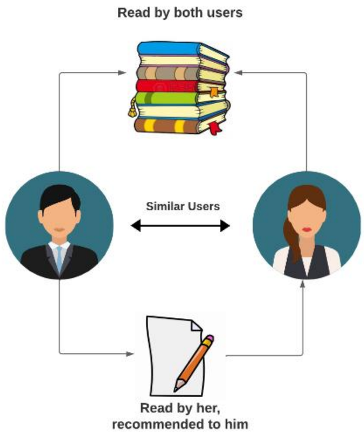


Collaborative Filtering

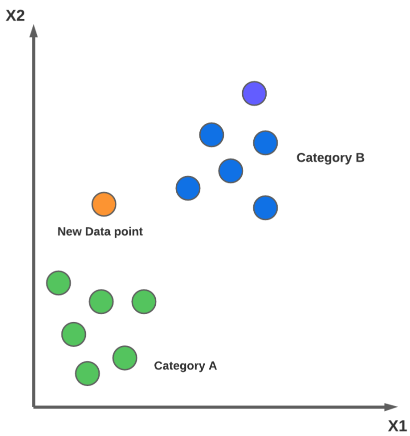
CF only takes into account each user's item ratings. CF is designed to make suggestions when accurate knowledge about the users and things is missing. Also, it successfully addresses the issue of over-specialization, these are very common drawbacks. Based on possible customer preferences for products from related user interests. It is founded on users' shared tastes. Also called social filtering. By comparing people's tastes, this system analyses user data to provide personalized recommendations. Since collaborative filtering doesn't require subject expertise, using this technique is simpler.

It contains two types: One is based on users and other is based on items.

Both user-based and item-based metrics assess similarity between two users and two items, respectively.



K-Nearest Neighbor (KNN) Algorithm – It is one of the simplest supervised algorithms that assumes similarity between new and available cases. Is is used for regression and classification problems. It is not parametric and also known as lazy learner algorithm. It helps to classify the category of new data point.



III. LITERATURE SURVEY

The field of recommender systems is the subject of many studies. The project was inspired by some recent developments in the field of research. In a published research article, Tan Nghia Duong et al. The idea that features such as word vectors or video genres are effective indicators of item content is explored. Video word vectors and genome tags can effectively represent videos. These “genomic markers” may have different genomic scores and share many semantic similarities. The researchers attempted to reduce this complexity by using a "word2vec" algorithm to create new raw genomic markers with composite scores. The previous process eliminated some tags, but a large portion of the tags are still connected but not conceptually persisted. To find these tags, the authors used an auto-encoder. Using this procedure, the complexity of the space is greatly reduced, and then the output is generated by matrix factorization techniques to generate propositions. The authors compared the two models and discovered that they perform better in terms of RMSE and MAE (mean absolute error). Use the RMSE and MAE measurements to examine accuracy. One of the issues raised above is the usage of hard clustering in practically all clustering techniques. I attempted a different tactic in this study by Remigio Hurtado et al. To further elaborate on what was previously stated, practically all early research articles used clustering algorithms that typically dealt with "hard" clusters, where members obviously belonged to a single cluster, not even a cluster. This study assigns the likelihood that an item belongs to the so-called "soft" group rather than assuming such rigidity. Generates the probability that a user belongs to each group in the user membership matrix. Two soft clustering techniques, BNMF and CFM, are applied. Further-more suggested is a third tactic that blends soft clustering methods with conventional CF methods. A few of the quality indicators that are taken into account by these soft clustering methods are the coefficient Fpartition, which shows how much one cluster overlaps another, and soft separation, which is the bare minimum distance that should exist between clusters. The study also provides an extensive and in-depth analysis of soft set theory, highlighting how it deals with conflict situations more effectively than approximate set theory by reaching conclusions in about 8.05% fewer seconds of computation. The subject of recommender systems has been the subject of many studies. This work was inspired by some recent advancements in the realm of study. The authors believe there is a lot of room for soft set theory in the field of recommender systems. The nearest neighbours of objects or video users can be found through clustering depending on strength, support, cover, and certainty. A number of methods are developed to estimate the computing time of a video recommendation system while taking the system's performance and scalability into consideration. as an instance, the authors of the paper have created a simple yet effective real-time video recommendation system. They start by carefully examining certain features of each user profile before categorizing users using a simple but incredibly effective algorithm. Second, for each cluster, a virtual opinion leader is picked to speak on behalf of all other users in the cluster. This reduces the original dimensionality of the user-item matrix. To generate final recommendations, a weighted single-slope algorithm was developed and applied to the virtual opinion leaders item matrix. With this strategy, the system's complexity can be decreased and outcomes can be attained more quickly. The expected accuracy in this situation isn't the ideal, unfortunately, because utilizing just one user in the cluster can result in the loss of some crucial dataThe extensive collection of content provided by online streaming services like Netflix, Amazon Prime, Disney+, Hotstar, etc. is a particularly time-consuming challenge from an IT perspective. By comparing the efficacy of recommender systems that included and those that did not include low-rated films, Muppana Mahesh Reddy et al. looked into this matter. In their research report, they contrasted MRS performance where videos ranked below average (for example, 2.5 on a scale of 0 to 5) were never recommended with and without an upper limit of recommendation. Any videos with ratings below the median are first eliminated using the research paper method. The Pearson correlation coefficient was then used to identify the related videos. Finally, offer a recommendation. Run the identical programme in parallel with no upper bound and contrast the outcomes of the two. We discovered that just 10% of the films scored below average on the particular dataset, thus we advise disregarding those that do when making suggestions to save money. The work mentioned above is the most pertinent to the goals of this research work. Several research, however, have adopted a totally different strategy. This can be seen in the work of XIAOJIE CHEN, who describes videos by isolating aesthetic elements from video stills and postersTo extract CNN (Convolutional Neural Network) characteristics from videos, a programme called "VGG16" is employed. 'VGG16' has undergone some modifications to eliminate the Softmax layer (used for image classification, not for image description). An aesthetically conscious Unified Visual Contents Matrix Factorization (UVMF-AES) model was created to give suggestions to the user using the output and assessment capabilities of VGG16. By the use of techniques like negative sampling, this process was optimised, and recommendations were then given. Another technique was to employ user-submitted written reviews, or, to put it another way, a new strategy was devised to leverage user-submitted written assessments. Sumaia Mohammed Al-Ghuribi used recommendation algorithms that run on websites that accept written comments. If handled correctly, this feedback can be more insightful and significant. Such methods solicit user evaluations as well as a quick overall assessment. These reviews are then used in conjunction with NLP algorithms to gather relevant data and improve recommendations. Three fundamental strategies are used here, beginning with the use of stacked autoencoders. (a) using stacked-autoencoders (b) Clustering users around cluster nodes and creating recommendations based on groups of users with similar characteristics clustered around the same cluster node. (c) Using a utility-based approach. Three metrics are utilized to calculate utility scores between objects: PCC score, Euclidean distance, and cosine similarity. Furthermore, point ranking, pairwise ranking, and list ranking are employed to ascertain user expectations. These characteristics are used by the utility function. For "handwritten" user evaluations, NLP techniques are used to find terms that fit into the categories of opinion words (polarizing, outrageous), sentiment words (friendly, insightful), and comparison words (better, worse). Obtaining relevant feedback is critical for developing an accurate MS. It is also critical to put this feedback data to good use by extracting relevant data from it to supplement your recommendations. In recent years, many sentiment-based systems that focus on the emotional tone of reviews provided in a dataset have been developed with the same idea in mind. The authors attempted to account for user sentiment in tweets on the microblogging site Twitter. They used Twitter to learn about people's general reactions and impressions of the film. They started by performing a sentiment analysis on the tweets. Finally, the hybrid system was integrated into the sentiment analysis results, improving them and allowing them to outperform most standard methods. Similarly, the authors built a hybrid model with the Spark platform and sentiment analysis. They started by creating a hybrid referral system to generate a preliminary list of referrals. They then used sentiment analysis to refine and optimize the list. A sentiment analysis and hybrid recommendation system is eventually developed using the Spark platform. Our method generated recommendations quickly while retaining a respectable level of accuracy. Writing reviews was the focus of Zhou Zhao, another researcher. This study examines how users' written input might be combined with their emotional state to generate suggestions. We used WordNet emotion lexicons. The first step was to develop a vocabulary for the six main emotions, including love, fury, fear, joy, sadness, and surprise. Corresponding WorldNet keywords for primary, secondary, and tertiary emotions were added to their respective lexicons and given different values, along with their synonyms. They created a system based on the fusion of opinion mining and similarity analysis in the paper. They first gathered the data and preprocessed it using a few key preprocessing techniques. The preprocessed data is subsequently sent for attribute extraction, both implicit and explicit. Classes are used to further categorize aspect words. Lastly, the user is given the top-k suggestions. Eventually, a model that considers the novelty and variety of recommendations was constructed in the paper. Here, the authors suggested a brand-new recommendation strategy based on covariance that largely utilizes correlation coefficients. Despite not knowing the item distribution in advance, it first expresses the positive and negative correlations from the experimental samples. Hence it is suggested to use CovH, which is created by limiting the favourites in the list of suggestions and can increase the originality and variety of recommendations compared to Cov.

IV. CONCLUSION

In this survey, it is discussed what video suggestion is, why we need it, and where it is used and required in reference to all of these research publications. This essay outlines historical and current research in the area, as well as how authors have attempted to address issues with video recommendation system algorithms in various ways. Video Recommendation System follows collaborative filtering than any other methods. In terms of recommendation algorithms, collaborative filtering (CF) has gained significant traction. As they frequently combine the best elements from the earlier methods. However, there are some drawbacks even after advanced techniques like NN, DL, Knowledge discovery over the recent years. Still it is preferred by most of the authors as it gives accurate and effective results. To the best of our knowledge, however, relatively few studies have shown the common characteristics of the numerous CF algorithms for mobile Internet apps.

ACKNOWLEDGEMENT

Thank you to the Computer Science Department of SVKM's Mithibai College, University of Mumbai. Thankful to the entire staff for providing support and facility for carrying out the research.

REFERENCES

[1]. "A New Challenge on Content-Based Video Recommendation", Mennat Allah Hassan, Youssef Roshdy, Foad Osama, Aly Mohamed, Walaa H. El Ashmawi, and Amr Sherif., IEEE Access 2019.

[2]. "Utilizing User Attributes as the Base for Collaborative Filtering Recommendation Algorithm", Mengdi Ma, Yi Tian, and Ruonan Ji, IEEE Access 2021.

[3]. "Utilizing User Characteristics as the Foundation for Collaborative Filtering Recommendation Algorithm", Mengdi Ma, Yi Tian, and Ruonan Ji, IEEE Access 2021.

[4]. "A STUDY ON CONTENT-BASED VIDEO RECOMMENDATION" by Hanjie Wang, Liu Bo Chen, and Hailong Yan Li, IEEE Access 2017.

[5]. "A Big Data-based Micro-video Recommendation System", Songtao Shang, Zhiguo Hong, Wenqian Shang, and Minyong Shi, IEEE Access 2016.

[6]. "Using a Novel User Similarity Measure, a Modified Memory-Based Collaborative Filtering Algorithm", Ramil G. Lumauag, IEEE Access 2021.

[7]. "The design and implementation of online video syndication and programme recommendation system", Jiechao Chen and Huan Wang, IEEE Access 2012.

[8]. “Assessment of Context-Sensitive Recommendation System using Hybrid Collaborative Filtering Method Bambang” Dwi Wijanarko, Rosilah Hassan, Dina Fitria Murad, Silvia Ayunda Murad, Riyan Leandros, IEEE Access 2022.

[9]. “Algorithm for Recommending Videos Based on Human Interest”, Shainee Jain, Omkar Morye, Tejaswi Pawar, and Heth Shah.

[10]. “Open Video Project Video Project”, Michael L. Nelson, Gary Geisler, Johan Bollen, and Raquel Araujo, IEEE Access 2005.

[11]. “A Content Based and Collaborative Filtering Recommender System”, Li Zhang.; Vignesh Thannimalai, IEEE Access 2021.

[12]. “Collaborative Filtering Algorithm Based on Sliding Window”, Xuejun Yu.; Yukang Liu, IEEE Access 2022.

[13]. The decision tree-based "Design of Intelligent Film and Television Advertisements Recommendation System" Huawen Zhao, Xiaohong Ling, and Yanhua Liu.