IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Analytical Study On Recommendation Systems Using Collaborative Filtering: Ahp Perspective

Ismat Anjum[1],Anwar Ahmad Shaikh[2]

[1]CSE Integral University , Lucknow,India

[2]Department of CSE Integral University , Lucknow

Abstract: The recommendation system plays a crucial part in the current day and employed by many famous apps. The recommendation system has created the gathering of applications, producing a global community, and growing for plentiful knowledge. The recommendation system developed into \sCollaborative Filtering, Content-based, and hybrid-based techniques. It is essential to provide the user with movie recommendations so that the user does not have to spend a significant amount of time searching for content that they would like. As a result, the function of the movie recommendation system is quite important in order to acquire user-specific movie choices. After doing considerable research on the internet and consulting a large number of scholarly articles, we came to the conclusion that the suggestions generated by Collaborative Filtering only use a single method for converting text to vectors and only use a single method for determining the degree to which vectors are similar to one another. Our project's goal is to create a recommendation engine that responds to the user in order to obtain ideas for a movie.

Index Terms - **Recommendation Services, Collaborative Filtering, AHP, Comparative analysis**

I Introduction

A recommendation system is a form of knowledge filtering system that makes an effort to anticipate the preferences of a user and offers a recommendation based on the user's selections [1]. Systems may perform a range of functions. These have gained popularity in recent years and are now used by most online businesses. Such stages may include movies, music, books, and records, friends and tales shared on online networking media platforms, things sold on online business websites, persons featured on professional and matrimonial websites, and Google search results. Predictive systems are known as recommendation systems because they propose products to users, users to goods, and occasionally even users to users [2]. Recommendation systems also promote users to other users. Tech giants such as YouTube, Amazon Prime, and Netflix all employ similar strategies to propose video material to users based on the interests those users have shown. Because the internet includes vast quantities of data, locating your material may be very challenging and time consuming; hence, recommendations play a crucial part in reducing the amount of effort we have to spend [3]. The majority of the time, these systems will gather data on the decisions made by users and then utilise that data to enhance their recommendations in the future. Adapter, they are able to suggest the Adapter to a first-time customer who has just added a MacBook to their shopping basket. Users are consistently provided with useful suggestions as a result of the advancements made in recommender systems. If a user's preferred genres of music are not available via a music streaming app, the user will likely cease using the service [4, 5]. Because of this, the pressure on technology businesses to improve their recommendation algorithms has increased significantly. However, the problem is bigger than it seems. Every person has their own unique set of preferences and interests. They have to investigate uncharted territories to learn more about the user while also making the most of the information that is currently available on the user. There are three primary strategies that are implemented inside our recommender systems. One of these techniques is called demographic filtering, and it means that the company provides summed-up ideas to each customer according to the predominance of movies or their probable categorization [7]. Customers are given recommendations for films that are comparable to one another in terms of their section highlights. Since every client is diverse, this methodology is viewed as excessively straight forward [8]. This structure is based on the premise that popular films would be enjoyed by the typical audience member. Second, content-based separation, where we profile the client's interests using data and recommend items based on that profile.

II Background

Sang-Min Choi, et. al. [1] discussed the drawbacks of collaborative filtering, such as the cold-start issue or sparsity problem. The authors have come up with a solution to this problem by using category information. A movie recommendation system based on genre correlations has been suggested by the authors. According to the writers, the freshly developed material has category information included in it. Even if a new piece of content doesn't have a lot of ratings or views, it might nevertheless show up in the suggestions list because of its category or genre information. The suggested system is neutral in its treatment of both highly rated and less-watched new material. Because of this, even a newly released film might be suggested by the recommendation system. **George Lekakos, et. al.** [2] presented a hybrid approach to movie recommendation as a solution. According to the authors, both Content-based filtering and Collaborative filtering have advantages and disadvantages. A hybrid strategy has been devised by the authors, which takes into account both content-based and collaborative filtering techniques. A movie recommendation system known as 'MoRe' uses the solution. The Pearson correlation coefficient was not employed for the purpose of collaborative filtering. In place of this, a new formula has been used. However, there is a 'division by zero' mistake in this formula. In the case of a tie in movie ratings, this mistake happens. As a result, these users were overlooked by the writers. In the case of a purely content-based recommendation system, the authors have employed cosine similarity to take into account the writers, actors, directors, producers, and genre of the film. They used two methods to develop a hybrid recommendation method:'substituting' and'switching'. Collaborative filtering and content-based filtering are used in both of these systems, which provide results when particular criteria are satisfied. Because of this, the authors use the collaborative filtering methodology as their primary method. **Debashis Das, et. al.** [3] wrote about the different types of recommendation systems and their general information. An overview of recommendation systems was provided in this work.

The authors discussed customised and non-personalized recommendation systems. An excellent example was used to demonstrate the difference between user-based and item-based collaborative filtering. Various recommendation systems have also been discussed by the writers.

III Collaborative Filtering (CF)

It merely advertises the items to those with similar tastes by filtering out material based on user interest that is comparable to that of other users. It is also a well-known algorithm in a variety of fields. There are two primary filtering algorithms in memory-based techniques. Model-based techniques, on the other hand, are less reliable than memory-based tactics. In the event that sufficient data is available, collaborative filtering-based recommendation systems may provide an accurate prognosis since they are based on the user's preferences. When it comes to predicting consumer behaviour, the most critical part of a recommendation system, user-based collaborative filtering has proven extremely successful in the past. In spite of this, their widespread use has exposed certain real challenges, such as data sparsity and data scalability, as the number of users and items continues to grow. It is necessary to have a collection of items that are reliant on the user's previous choices in order to employ collaborative filtering. This approach does not need a substantial quantity of product features to function. An embedding or feature vector represents each item and User, and it sinks both the things and the users in a same embedding position. It develops enclosures for goods and users on its own. Other purchaser's responses are taken into account when offering a certain product to the principal user. It keeps track of the behaviour of all users before proposing which item is generally loved by people. It also links comparable consumers by similarity in desire and behaviour towards a similar product when suggesting a product to the core client. This chapter suggested item-based collaborative filtering as an application of dimension reduction in a recommendation system so that the prediction problem's execution time could be shortened and its accuracy could be increased. It illustrates that the suggested method is capable of achieving higher performance and execution time for the recommendation system in terms of the current problems, as shown by validation via AHP approaches. In figure 1, we have shown the flow structure of the research work behaviour from a number of different points of view.

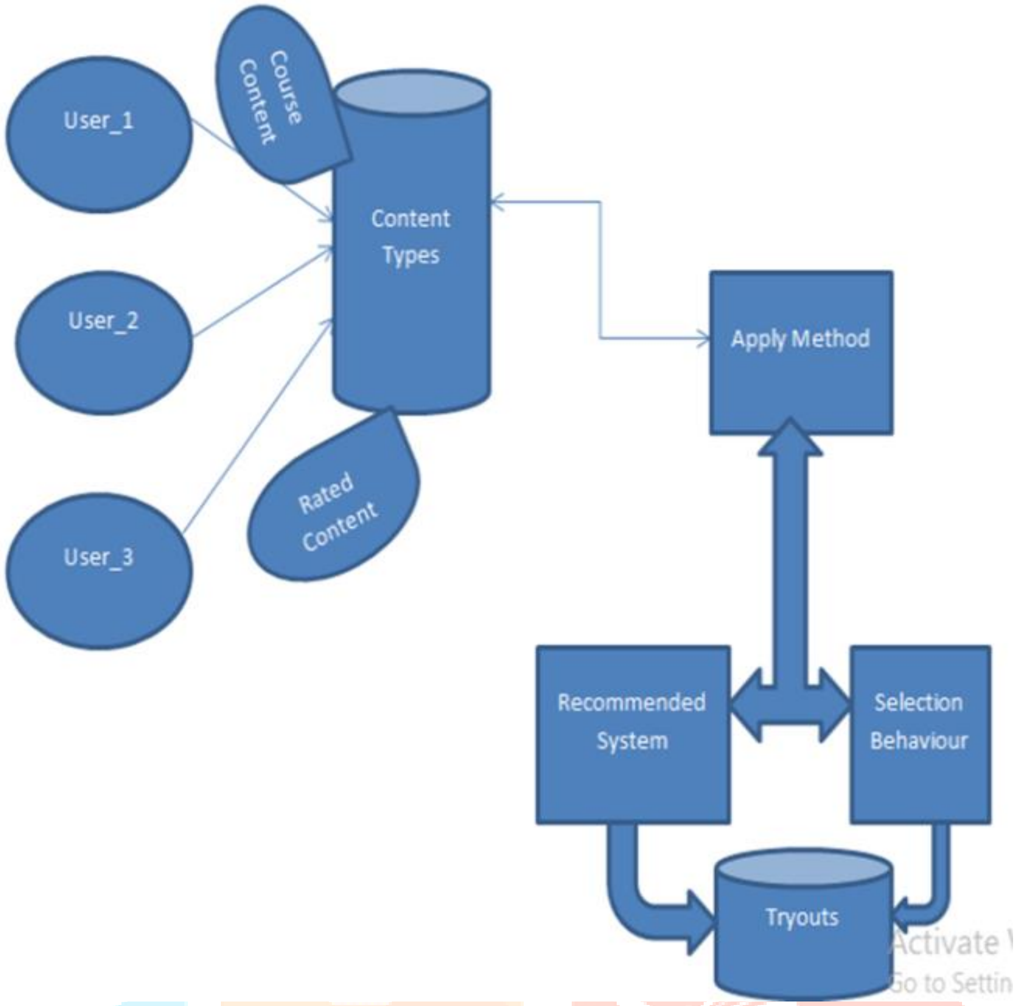

Figure 1: Flow Structure of Research

3.1 Computing Similarity among Users

In order to determine the degree of similarity between users, we use the idea of Pearson correlation, as well as the cosine similarity and user-based prediction computing formula.

3.1.1 Pearson Correlation

This method computes the statistical correlation (Pearson's *r*) between two user's common ratings to determine their similarity. The correlation is computed by the following:

$$\text{sim (i,j)}=\frac{\Sigma u\in U(Ru,i-\bar{R}u)(Ru,j-\bar{R}u)}{\sqrt{\Sigma u\in U(Ru,i-\bar{R}u)^2}\sqrt{\Sigma u\in U(Ru,j-\bar{R}u)^2}}.$$

(Eq. 2.1)

Where,

i: Set of items rated by the user.

R_u, i: Is the rating given to item I by user (u).

$\bar{R}_u$: Is the mean rating given by user (u).

Cosine Similarity

Nearest-Neighbor CF algorithm:

A measure of the degree of similarity between two non-zero vectors of an inner product space is referred to as the cosine similarity, and it is determined by the cosine of the angle that separates them. The cosine of an angle of 0 degrees equals 1, but the cosine of every other angle is less than 1.

For N-dimensional vector of items, measure two customers A and B

Similarity $(\vec{A},\vec{B})= \cos(\vec{A},\vec{B})= \frac{\vec{A}.\vec{B}}{||(\vec{A})||*||(\vec{B})||}$ (Eq. 2.2)

When calculating cosine similarity, negative ratings are not allowed, and unrated things are given a rating of zero regardless of whether or not they have any ratings at all.

2.3.3 User Based Prediction Computing Formula

$p_{C,e}= \bar{r}C + \frac{s(C,A)(rA,e-\bar{r}A)+s(C,D)(rD,e-\bar{r}D)}{(|s(C,A)|+|s(C,D)|)}$ (Eq. 2.3)

(pC,e): User C's prediction for Equilibrium

$\bar{r}C$: Mean rating of user C

$\bar{r}A$: Mean rating of user A

$\bar{r}D$: Mean rating of user D

s(C,A): similar user A and C

s(C,D): similar user C and D

From above formula we find the prediction of user in unknown item. Due to this, then we recommend this item to that user whose was not rated this item. Now we consider a table of ratings of different users.

Table 1 Ratings matrix

	Item 1	Item 2	Item 3	Item 4
User A	4	?	3	5
User B	?	5	4	?
User C	5	4	2	?
User D	2	4	?	3
User E	3	4	5	?

Take a look at the ratings matrix in Table 1, and keep in mind that we need to locate User C's estimate for the value of Equilibrium (pC,e) using the following configuration:

• Pearson correlation.

• Neighborhood size of 2.

$\bar{r}C = \frac{(5+4+2)}{3}$

$\bar{r}C =3.667$

C's average score is 3.667. There are only two users who have given the game a rating, and as a result, there are only two people who are candidates for inclusion in the neighbourhood:

$$s(C,A)=\frac{(4-4)*(5-3.67)+(3-4)(2-3.67)}{\sqrt{((0)^2+(-1)^2)}*\sqrt{((1.33)^2+(-1.67)^2)}}$$

$$=0.784$$

A and D, $s(C,A) = 0.784$ and $s(C,D) = -0.518$ from Equation 1.1.The prediction pC,e is therefore computed as follows:

$$p_{C,e}=\bar{r}C + \frac{s(C,A)(rA,e-\bar{r}A)+s(C,D)(rD,e-\bar{r}D)}{(|s(C,A)|+|s(C,D)|)}$$

$$= 3.67+ \frac{0.784*(5-4)+ -0.518*(2-3)}{0.784+0.518} =4.667$$

Here we calculated the value of user rating of C. On the behalf of this rating we can recommend items to the users.

IV Impact Analysis

An accurate assessment of the relevant data is one of the procedures that is considered to be among the most critical in several decision-making processes. This is a difficulty that is not exclusive to the AHP technique alone; rather, it is essential in a wide variety of other approaches that need the decision-maker to provide qualitative information. There is a high likelihood that qualitative data cannot be known in terms of absolute values and are not unique. As a result, a significant portion of AHP is dedicated to figuring out the relative relevance, or weight, of the various options in terms of each criteria that is associated with a specific decision-making issue. With the Analytic Hierarchy Process (AHP), pairwise comparisons are conducted to establish the relative relevance of each choice in relation to each criteria of the django framework. In this method, the individual in charge of making decisions is tasked with providing commentary on the significance of a single comparative pair at a time. In most cases, the person making the decision will be required to choose his response from a number of distinct options. In order to put this validation into action, we must begin by taking the parameters, whose numeric values represent the heading and velocity. The prediction model provides an explicit estimate of the heading, while the point model delivers an implicit estimate of the heading depending on the selection. After that, we take a model of the movie system that was suggested to us. In order to compute, we make advantage of the many domains that the dynamic paramerts provide. The next thing we do is form the hypothesis that the dynamic model will go on with its scenario, and that it will most likely keep the same offset that it has now. The researchers will get information on the fundamental steps involved in the recommendation system.

The recommendation system operates mostly based on the specifics of the product as well as the data of the user. We are required to get them from the system or from the database, and based on the ratings, we will make our selections. If an item that is comparable to what was searched for was discovered, then a recommendation system will be developed; otherwise, no recommendation system will be formed.

4.1 AHP Methodology:

When it comes to making decisions, the Analytic Hierarchy Process (AHP) was developed by Saaty (1977 and 1994). Many scholars are intrigued by the AHP, in part because of its appealing mathematical features and the relative ease with which it may be used with readily available input data. Using the AHP as a decision support tool, you may work through more difficult issues of decision-making. Objectives, criteria, sub-criteria, and options are organised into a multi-level hierarchical framework. A series of pairwise comparisons yields the relevant information. For each individual choice criterion, these comparisons are utilised to determine how important that criterion is, as well as how well each option performs in contrast to the others. In cases when the comparisons aren't exactly the same, this is a way to make things more consistent going forward. In table 5.1, Let C = {Cj| j = 1, 2... n} be the set of decision criteria. The data of the pair wise comparison of n sub-criteria can be summarized in an (n× n) evaluation matrix A in which every element aij (i, j = 1, 2 ... n) is the quotient of weights of the criteria. A square matrix and a reciprocal matrix may be used to illustrate this pair-wise comparison. In this matrix $a_{ij} = 1/a_{ji}$, for all experts, we would have (n× n) matrices (see table 5.1).

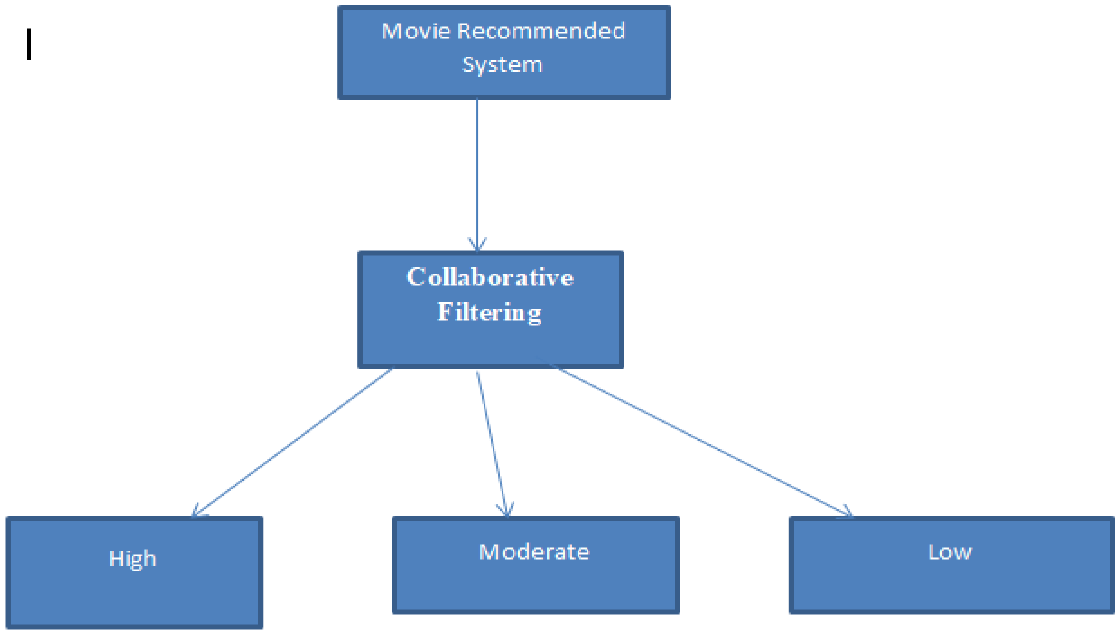

Figure 2: Conceptual Behaviour

Table 1 Assign Weight (Collaborative Filtering)			
	High	**Low**	**Moderate**
High	1	3.9	5.6
Low	0.2564	1	3.2
Moderate	0.1785	0.3125	1
	1.435	5.213	9.800

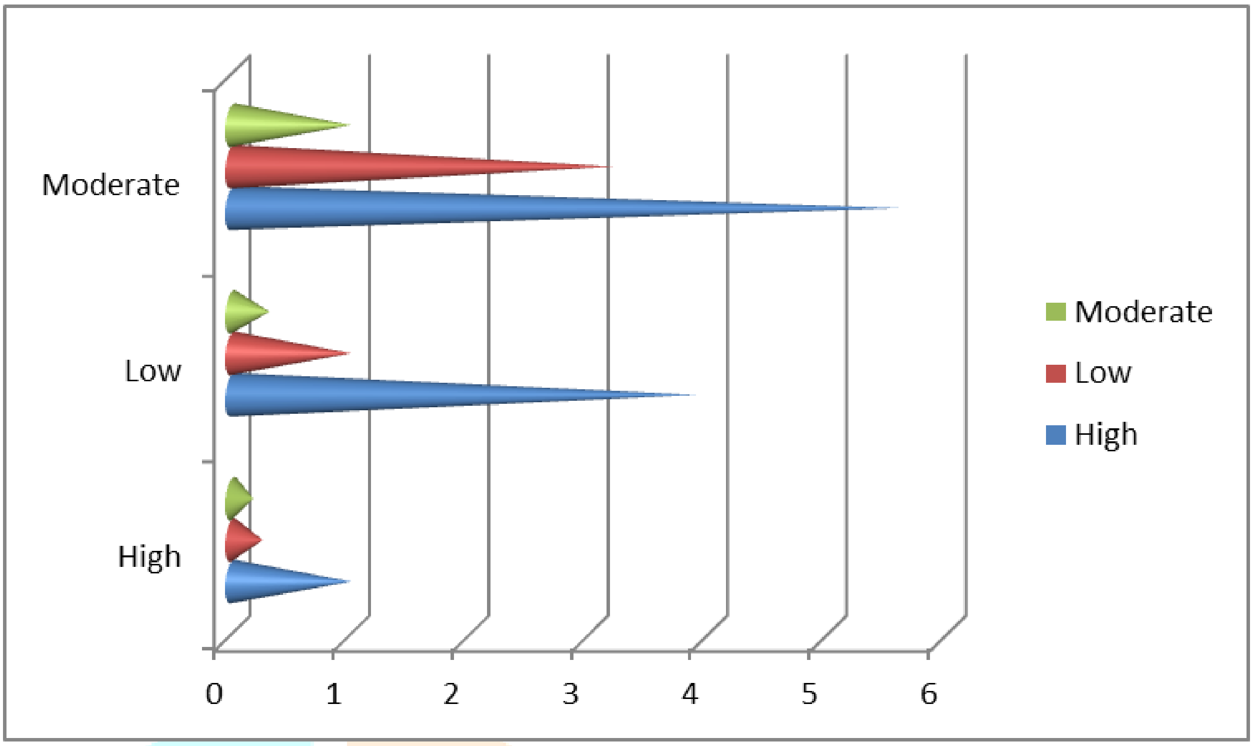

Figure 3: Graphical Valuation

In Table 2, the Accuracy resulted in the best value with an ideal priority Vector 0.6722, mainly because it was the highest evaluated in the two metrics: 0.6969 and 0.7482 (figure 3). Nonetheless, there are three alternatives for model that also stand out. In table 2, it is possible to see the average consistency index obtained from the output of the alternatives in the test phase.

As a result of the curse of dimensionality, it is possible to use the AHP to calculate the options among different models and justify the model's accuracy. The new approach has been introduced to solve the most important alternatives, and the details of this approach are provided in Table 2 with consistency index for verifying the stage calculations.

Table 2 Normalized Metrics

	High	Low	Moderate
High	0.6969	0.7482	0.5714
Moderate	0.1787	0.1918	0.3265
Low	0.1244	0.0600	0.1020

Eigen Vector	Priority
0.9645	0.6722
1.2111	0.2324
0.9357	0.0955

Eigen Value 3.114	0.0557

In table 3, we have calculated to overall priority of each criteria respect to model weight. We are observed that high values of very effective in this research work. Table 3 are given the finalize metrics in the form of priority High Low Moderate context. Accuracy value is the maximum effective constraints which provides verifying and validated of research.

Table 3 Calculate overall priority

	High	**Low**	**Moderate**
High	0.6722	0.653	0.600
Moderate	0.2324	0.251	0.200
Low	0.0955	0.096	0.200

Table 4 Finalize Metrics

High	**0.6252**
Moderate	0.2137
Low	0.1611
Highest Priority = Highest Score	

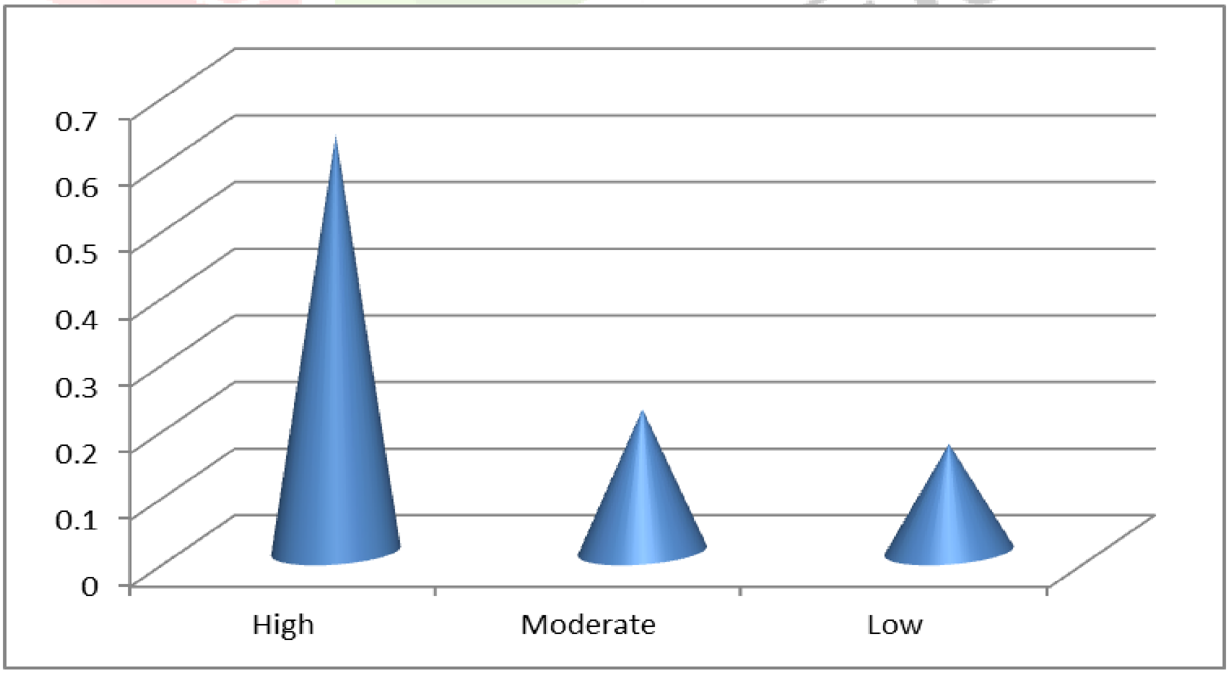

Figure 4: Final Structure

The inconsistency of the AHP pairwise comparison matrix is addressed in this study using three different models, depending on three different levels of inconsistency: high, low, and moderate. Initially, both methodologies are compared using simulations that include training, validation, and testing. Compared to High Value in CR reduction, the Collaborative Filtering approach has a similar behaviour but is better at anticipating previously unknown inputs supplied to the network, giving it a major advantage over the former. It's possible to observe the ultimate AHP weight in Figure 4 (right).

5. Comparative analysis

After we have obtained the findings, it is strongly advised that you compare the work that was proposed with other recent work that has been done in the same area. Manoj et al. (2015) have implemented a method for a movie recommendation system that is based on the weighting of criteria, and this method is quite similar to the one that we have suggested here. In the future, we want to concentrate on improving its user interface as well as its weaknesses.

Table 1 Comparison of ranks between proposed work and existing work (Manoj et.al(2015))

	Weight of Proposed work	Ranking of Proposed work	Weight of Existing work	Ranking of Existing work
Aggregate	**0.6252**	1	0.60	1
	0.2137	2	0.54	2
	0.1611	3	0.48	3
Moderate	0.653	1	0.42	3
	0.251	2	0.36	2
	0.096	3	0.30	1
High	0.6722	1	0.61	3
	0.2324	2	0.48	2
	0.0955	3	0.42	1
Low	0.600	1	0.45	3
	0.200	2	0.36	2
	0.200	3	0.06	1

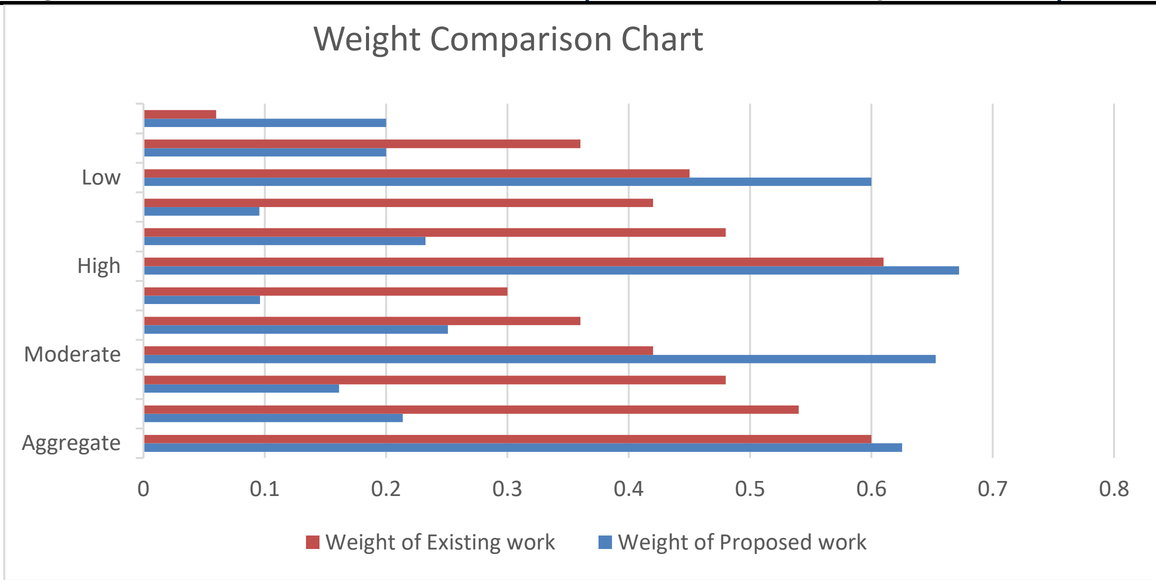

Figure 5: Weight Comparison Chart

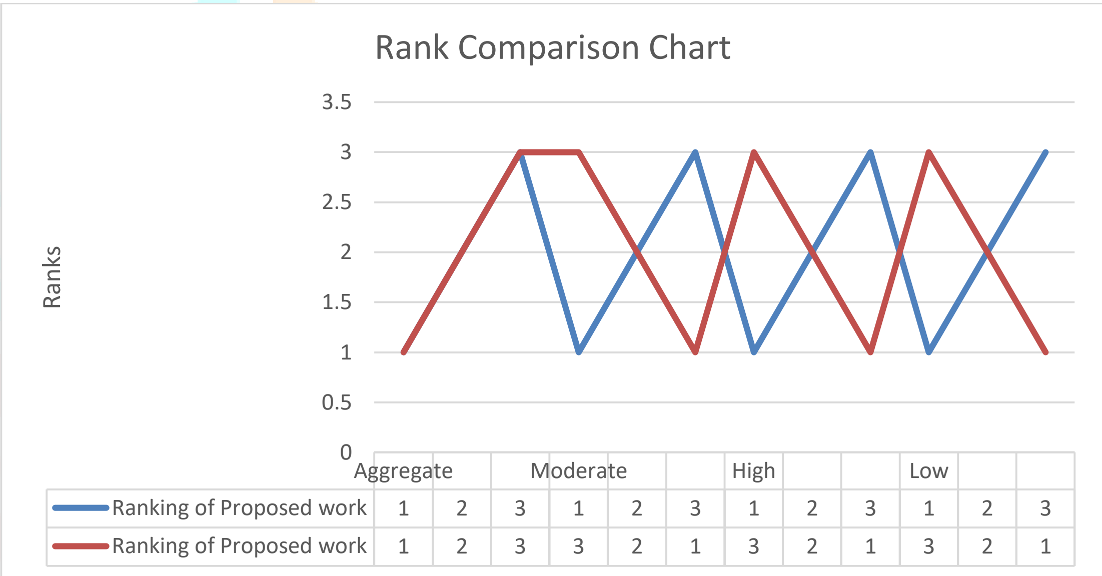

Figure 6: Rank Comparison Chart

6. Conclusions

There are new routes for obtaining personal information online thanks to the advised solutions. A common problem with information retrieval systems is an overabundance of information, and this solution helps to relieve that problem while simultaneously providing customers with quick, convenient access to products and services that may otherwise be unavailable. Based on the input that customers have given us in the past and our strategy to respond to their unique needs and interests, we recommend movies to them. As a result of this method, known as AHP, the recommendations model becomes more accurate. In addition, the system is more responsive and the data collecting procedure is more accurate as a result of this. In addition to confirming the model, this approach also improve the system's performance. When it comes to discovering new things, the Recommendation system may be a lifesaver. Certain restrictions prevent these systems from suggesting effectively to their consumers. Even though Collaborative Filtering is the most effective and powerful method, this algorithm has a high runtime and confronts challenges like as data sparsity, which may be eliminated by employing a Hybrid movie recommendation system. However, there are advantages and disadvantages to each option. In our suggested validation method, we can deal with quite a large quantity of data with ease. We plan to address its weaknesses and improve its user interface in the future.

References

1. Choi, Sang-Min, Sang-Ki Ko, and Yo-Sub Han. "A movie recommendation algorithm based on genre correlations." Expert Systems with Applications 39.9 (2012): 8079-8085.
2. Lekakos, George, and Petros Caravelas. "A hybrid approach for movie recommendation." Multimedia tools and applications 36.1 (2008): 55-70.
3. Das, Debashis, Laxman Sahoo, and Sujoy Datta. "A survey on recommendation system." International Journal of Computer Applications 160.7 (2017).
4. Dietmar Jannach, Gerhard Friedrich; "Tutorial: Recommender Systems", International Joint Conference on Artificial Intelligence, Beijing, August 4, 2013.
5. Gaurangi, Eyrun, Nan; "Movie GEN: A Movie Recommendation System", UCSB.
6. Harpreet Kaur Virk, Er. Maninder Singh," Analysis and Design of Hybrid Online Movie Recommender System "International Journal of Innovations in Engineering and Technology (IJIET)Volume 5 Issue 2,April 2015.
7. Manoj Kumar, D.KYadav, Ankur Singh, Vijay Kr. Gupta," A Movie Recommender System: MOVREC" International Journal of Computer Applications (0975 –8887) Volume 124 – No.3, August 2015.
8. Prerana Khurana , Shabnam Parveen; 'Approaches of Recommender System: A Survey'; International Journal of Computer Trends and Technology (IJCTT) Volume 34 Number 3 - April 2016.