Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 19, no. 2, February 2023, pp. 105-114
DOI: 10.23940/ijpe.23.02.p3.105114

Quality Enhancement of Recommendation using Improved Triangle Ratings

Devendra Gautam[a], Anurag Dixit[a], Latha Banda[b], Harish Kumar[c], Purushottam Sharma[d,*], and Chaman Verma[e]

[a]*Noida International University, UP, 203201, India*
[b]*ABES Engineering College, UP, 201009, India*
[c]*Mangalmay Institute of Engineering & Technology, UP, 226021, India*
[d]*Amity University, Noida, 201301, India*
[e]*Eötvös Loránd University, Budapest, 1053, Hungary*

Abstract

Recommender Systems are a potent technology used in many social networking sites. Personalized recommender systems are an added method for improving the quality of recommendation and customer's requirements. There are many kinds of techniques available to get personalised recommendations such as Content based, Collaborative filtering and Hybrid filtering. In these mentioned techniques, the most popular CF technique is used to enhance the accuracy of RS with some shortcomings such as sparsity, scalability and cold start user problems. To enhance the quality of collaborative filtering using tagging, the proposed approach IUGT-Jaccard-ITR used may target the issue of cold start user or item problems in recommendation.

Keywords: recommender systems; collaborative filtering; collaborative tagging; tagging systems; jaccard distance; improved triangle ratings

1. Introduction

When information grows in website applications, managing content is very difficult and finding relevant information as per user's interest from these applications is also a challenging task. To manage these scaling data, the Recommender Systems (RS) [1] is used to provide personalised information to users. RS is classified into three categories such as (i) Content Based Filtering (CBF) which is based on the past behaviour of the user. (ii) Collaborative Filtering (CF) technique recommendations are done using similar users behaviour and (iii) Hybrid Filtering (HF) techniques is the grouping of the above mentioned techniques. The comprehensive description of these classification is explained as follows:

Content-based recommenders

To produce recommendations, content-based recommenders [2] look at the user's previous behaviour as well as the content items themselves. Pandora is an example of a recommendation system that is based on content. Their recommender algorithm takes into account the user's choices as well as song attributes such as length, instruments, harmonies, genre, and so on. It takes all of this information into account when recommending new music to a user. A user profile is generated using this information, which allows the system to deliver more personalised suggestions to the user.

Collaborative recommenders

In order to create recommendations, CF compares all users and user-item interactions (liked, disliked, and clicked) [3-6]. In general, the process of filtering for patterns or information utilize strategies that involve collaboration among several users, opinions, data sources, and other factors. When it comes to recommending information, collaborative recommenders outperform content-based recommenders.

* Corresponding author.
E-mail address: puru.mit2002@gmail.com

Tagging systems

Tags are terms that users may choose, and they are a simple yet effective way to organise, search, and explore the resources. It is also called Collaborative tagging [7-10]. The tagging systems are divided into two categories: static and dynamic tagging [11]. In static tagging the user may give interest or disinterest using thumbs up or thumbs down where in dynamic tagging the user gives his/her opinion or feedback on products or movies.

Generally, the recommendation techniques applied on social networking websites enhance the quality of recommendation. Here in this process, firstly the database collection is done of Social Networking Sites (SNS) [12] and this data is stored for proper analysis of data sets. Parallelly the user profile data is generated from these datasets either based on purchase history or using registration details. As shown in Figure 1, when the user enquires for any query the proper answer or solution is given to the user. The data collected is to be cleaned using some pre-processing techniques and using the recommendation engine the recommendations provided to the user.

movieId	Title	tag_list	tag	jaccard_sim
1704	Good Will Hunting	[bittersweet, boston, cerebral, character stud...	bittersweet,boston,cerebral,character study,ch...	1.000000
318	The Shawshank Redemption	[adaptation, adapted from:book, atmospheric, b...	adaptation,adapted from:book,atmospheric,based...	0.526718
1961	Rain Man	[1980s, amazing cinematography, autism, based ...	1980s,amazing cinematography,autism,based on a...	0.503759
3252	Scent of a Woman	[atmospheric, betrayal, bittersweet, blindness...	atmospheric,betrayal,bittersweet,blindness,cha...	0.503759
4995	A Beautiful Mind	[adapted from:book, based on a book, based on ...	adapted from:book,based on a book,based on a t...	0.492537
1784	As Good as It Gets	[affectionate, bittersweet, catastrophe, chara...	affectionate,bittersweet,catastrophe,character...	0.492537
4019	Finding Forrester	[adapted from:book, based on true story, baske...	adapted from:book,based on true story,basketba...	0.481481
8183	Educating Rita	[adaptation, adapted from:book, affectionate, ...	adaptation,adapted from:book,affectionate,alco...	0.470588
117176	The Theory of Everything	[adaptation, adapted from:book, based on a boo...	adaptation,adapted from:book,based on a book,b...	0.449275
89492	Moneyball	[1980s, adaptation, adapted from:book, art, ba...	1980s,adaptation,adapted from:book,art,bad end...	0.449275

Figure 1. Static Tagging vs Dynamic Tagging

Thesedatasets are trained using either supervised or unsupervised machine learning techniques to predict the future data or future behavior of the user. The recommendation engine again classified into offline recommendation or online recommendation. Using offline recommendation techniques, either ad alerts or email postings may get spread over the users. On the other hand, using online recommendation the Recommendation (RE) engine may analyze the ratings, tags and feedback of users and the same process repeated continuously to provide relevant recommendations to the users. Through this type of process, the recommendation quality and accuracy will be higher in RS. The architecture of recommendation systems process is shown in Figure 2.

CF is the main algorithm used in traditional RS. A CF makes recommendations to users based on their unique features, allowing them to tap into the knowledge of other users who are quite similar to them. Social data allows users to take into account new and diversified variables when calculating user similarity, which are not available in typical personalised RS [13]. In RS numerous notable studies have been conducted and the main findings of RS are as follows:

Synonymous Names

Synonymy is a problem that develops when a single product or thing is represented by two or more different names or lists of items (for example, action movie or action film). In this scenario, the recommendation algorithm is unable to distinguish if the terms display different items or the same item.

Scalability

The scalability of algorithms with real-world datasets is another challenge with recommendation systems. The traditional method has been overwhelmed by the wide range of products and clients in most cases, resulting in dataset problems and low reliability.

Latency Challenges

When new products are added to a recommendation engine's database more regularly, latency problems develop. Users are still recommended existing products because newly introduced products are not evaluated. To address the problem, organizations can utilize either a collaborative filtering approach or a category-based approach in combination with user-item interaction.

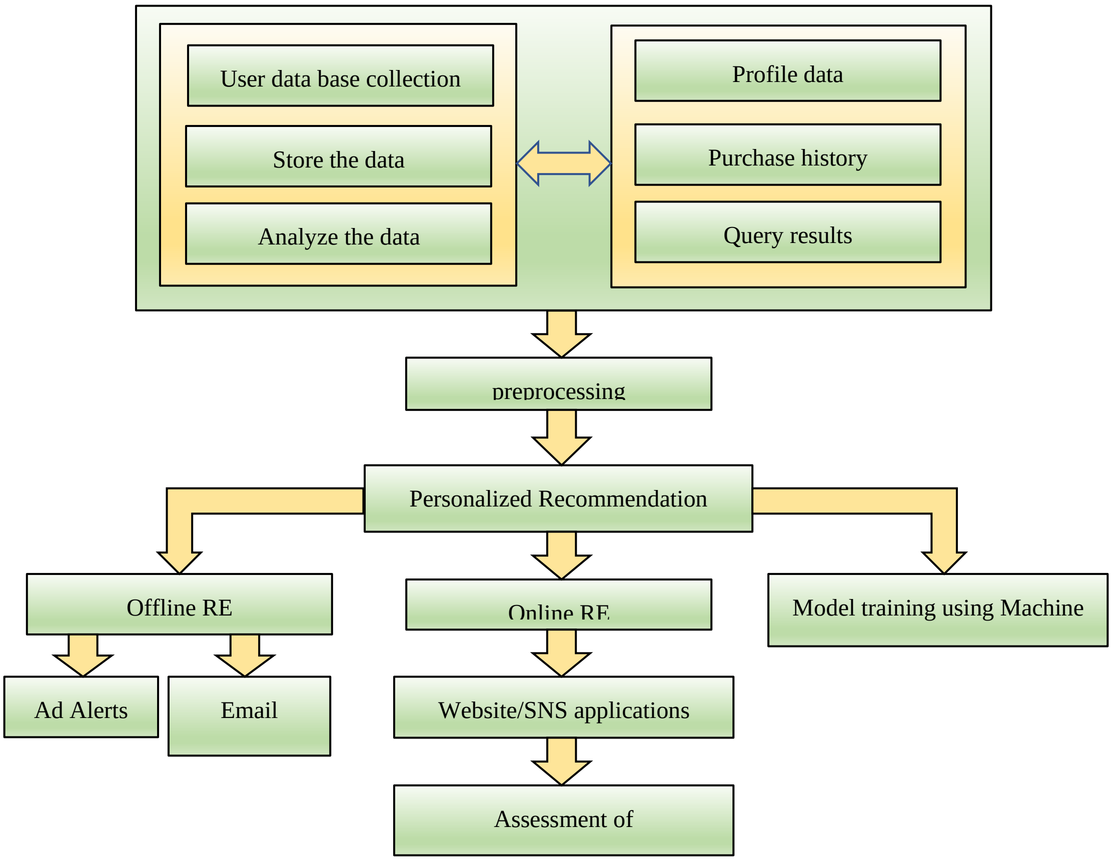

Figure 2. Architecture of Recommendation Systems

Privacy

Customers must usually provide personal information to the recommendation system in order to receive favourable services. However, it raises a number of data privacy and security concerns by making users provide personal information to recommendation algorithms.

Issue of Sparsity

The rating and review model becomes sparse when users do not submit ratings or reviews for the products they buy, resulting in data sparsity issues. It makes it more difficult to find a group of users with similar ratings or interests. To avoid these limitations, many researchers used various machine learning techniques to predict the user's behaviour and to enhance the accuracy of recommendations.

2. Related Work

The information about the preferences of its users for products is collected by the recommender system. It recommends items that its customers may want to buy based on their preferences. Previously, this knowledge could be retrieved by combining explicit and implicit data. These days, the need for Recommendation Systems is increasing in the social networking sites. Dhelim s et al (2022) [14] presented a survey on RS which focuses on personality aware RS. In this survey the exploration of design choices for personality aware RS was done by comparing personality modelling methods and RS techniques. As mentioned the RS is classified into content based, collaborative filtering and Hybrid filtering in which every conventional RS solve the problem of sparsity, scalability and cold start user/item issues.

The recommendations are provided to the user to get the best accuracy of social networking websites. In content based filtering (CBF) the user's past preference is considered to get recommendations. Donghui Wang et al (2018) [15] proposed a

CBF model with a chi square feature selection and softmax regression in which, CBF was applied for making recommendations of good journals and conferences to authors and these users may not decide where their research content was to be published. The web service used was http://www.keaml.cn/prs/ and in this, to fasten the search of journals and conferences, they have employed a web crawler with a continuous updating of training and testing models.

In CF, similar users behaviour or ratings were chosen for recommendation. In comparison to other recommendation techniques, CF is mostly used and a successful technique. CF is used in online recommendations to get dynamic recommendations but there is a problem of accuracy due to its simple model. Nikolaos Polatidis et al (2016) [16] proposed a multi-level recommendation method which targets the accuracy issue of recommendations.

HF is also a part of RS in which the recommendations are provided to the user with merge of CBF and CF techniques. Ayyaz S et al (2018) proposed an approach named Hybrid Content based Fuzzy Conformal Recommender System (HCF-CRS). In this, the approach works as content based where it collects the demographic features of a user and fuzzy techniques have been applied on content to get similarity between users. In this, to target sparsity issues CBF and to achieve accuracy, fuzzy techniques were used and they were compared with baseline approaches as a result of the merge of both CBF as well as fuzzy techniques.

The main issues of RS are sparsity, cold start and scalability. The cold start issue is classified into either cold start user or cold start item problem [17]. Due to lack of information about a new user, cold start user issues may occur in RS. To target this cold start user issues, Jesus Bobadilla et al (2012) [18] proposed a new similarity measure to using optimization based on neural learning to achieve better results in terms of precision and recall. Online data is a massive amount of data where the sparsity problem occurs in RS. Here Senthilselvan Natarajan et al (2020) [19] proposed RS-LOD (Recommender System with Linked Open Data) to target the issues of sparsity and cold start issues. A LOD knowledge base "DBpedia" is used to find enough information about new entities for a cold start issue, and an improvement is made on the matrix factorization model to handle data sparsity.

The main issue of RS is the cold start issue which is classified into complete cold start (CCS) problems in which no ratings are available and incomplete cold start (ICS) problems that provide only a minimal amount of ratings. To target these cold start issues, Jian Wei et al (2017) [20] proposed a deep neural network SADE which extracts the content features of the items. This model utilizes the temporal dynamics of user preferences to target the cold start issues. The following Table 1 presents some of the author's Literature review based on parameters of limitation, technique used and their outcomes.

Table 1. Literature Survey of Recommender Systems

Author	Limitation	Techniques used	Outcomes
Tey, F.J., Wu, TY., Lin, CL. et al (2021) [21]	Used only yelp data set and recommendations based on friend's friends relationships. This may not give accurate recommendations.	Proposed a scheme to find friend's friends to achieve better recommendations.	Improved accuracy in recommendation due to the use of indirect relations
Jian Wei, Jianhua He, Kai Chen, Yi Zhou, Zuoyin Tang (2017) [20]	Requires extra storage and computation time	To predict the ratings of users, the author used Time SVD++	achieved the issue of cold start problems.
Suvash Sedhain, Aditya Krishna Menon, Scott Sanner, Lexing Xie, Darius Braziunas S (2017) [22]	Experiments were done on content based filtering rather than collaborative filtering whereas content based is an old technique in recommendation which is used only for offline recommendations.	linear regression used for recommendations	Targeted the problem of cold start users or items .
Hyeyoung Ko et al. (2022) [9]	It was only a survey and there was no proposed method for handling cold start issues	Review on recommendation systems	To make better understanding for the researchers
Ajaegbu, C. (2021)	It was only an item recommendation	three current traditional measurement metrics such as: Cosine-based similarity, Pearson correlation similarity and Adjusted cosine similarity, in the direction of cold-start situations.	strength of the three traditional measurement metrics with evidence shown when measured with Mean Absolute Error.

3. Proposed Work

The proposed framework alleviates the cold start problem of CF based on implicit and explicit ratings [16]. Here the set of users, items, tags and new users are defined as: U = $\{u_1, u_2, \ldots u_m\}$, I = $\{i_1, i_2, \ldots i_m\}$, T = $\{t_1, t_2, \ldots t_m\}$ and N = $\{n_1, n_2, \ldots n_m\}$. In this the user's demographic profiles are considered for predicting his/her interest with the similarity computations.
In these the proposed algorithm steps are as follows in Figure 3.

In the proposed framework the dataset is collected from *movielens.com* in which the user, item and tag data is collected. Various pre-processing techniques have been applied on this dataset to filter the data. After the cleaning, the data is constructed in the form of matrices such as user-item, item-tag and user-tag. In the same matrices, here the clusters are generated of similar users, items and tags such that the similarity computations are done based on demographic features.

Step 1: Dataset collection form movielens.com
Step 2: Pre-process the datasets
Step 3: Generate user-item, user-tag and item-tag matrices
Step 4: Generate user, item and tag clusters for providing similar computations.
Step 5: Construct the matrices of implicit ratings and explicit ratings
Step 6: Compute the similarity of these user, item and tags based on Jaccard and Diffusion based similarity using equations: JS = a/a+b+c and
Step 7: Apply machine learning algorithm for optimizing the solution
Step 8: Compute the optimized results with prediction and recommendation for the best results.

Figure 3. Proposed Algorithm for Recommendation Systems

In the user cluster the set is defined as U = $\{u_a, u_b, \ldots, u_n\}$, I = $\{i_a, i_b, \ldots, i_n\}$ and T = $\{t_a, t_b, \ldots, t_n\}$. In this, the same profiles of users are clustered together with demographic features. The demographic information of user profiles includeuser_gender, user_age, user_occupation_label, user_occupation_text, user_zip_code. Based on this profile the similar users are clustered together. In the same way similar items and tags were also collected. So, when a new user enters into the system/any social networking site or if any item launched in websites this would be an easy process to recommend items to new users. The cluster process of user, item and tag data is shown in Figure 4.

Here, the movies are clustered of the same genre or type and tags are also clustered based on the similar tags given on different movies or same movies. The general overview of the proposed framework is shown in Figure 5. In this the first step is a dataset collected from movielens.com. In step 2, pre-processing is done using various techniques. In step 3, the dataset is clustered as user, item and tag clusters. Step 4 consists of similarity computation using jaccard and ITR similarity. In step 5, a machine learning algorithm is applied for the purpose of optimization and prediction and then finally in step 6, recommendation is done.

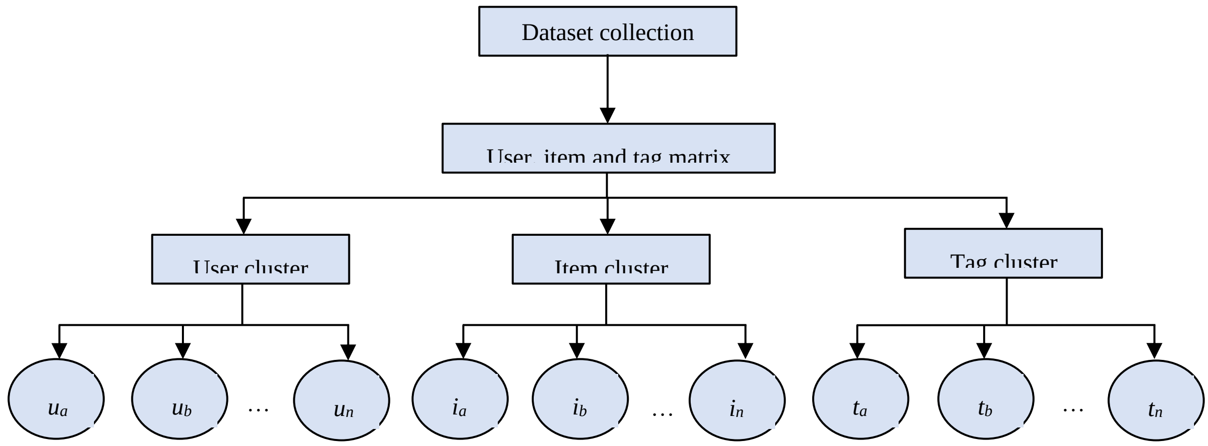

Figure 4. Cluster data of user, item and tag

There are two types of ratings such as implicit and explicit ratings in which the implicit rating is done based on user’s behavior whereas explicit ratings are user’s tagging or ratings on movies or products. Mainly the user gives feedback taken based on his/her behavior or ratings on items.

The collected feedback is also again classified into binary ratings or weightages of ratings. In this the binary is only counted if the user likes or dislikes any item; the rating counts as 1 or 0 respectively. The binary rating is only counting the rating in binary classification manner and even if a user browses on a particular item or hits on any movie or item then also it may count as 1.

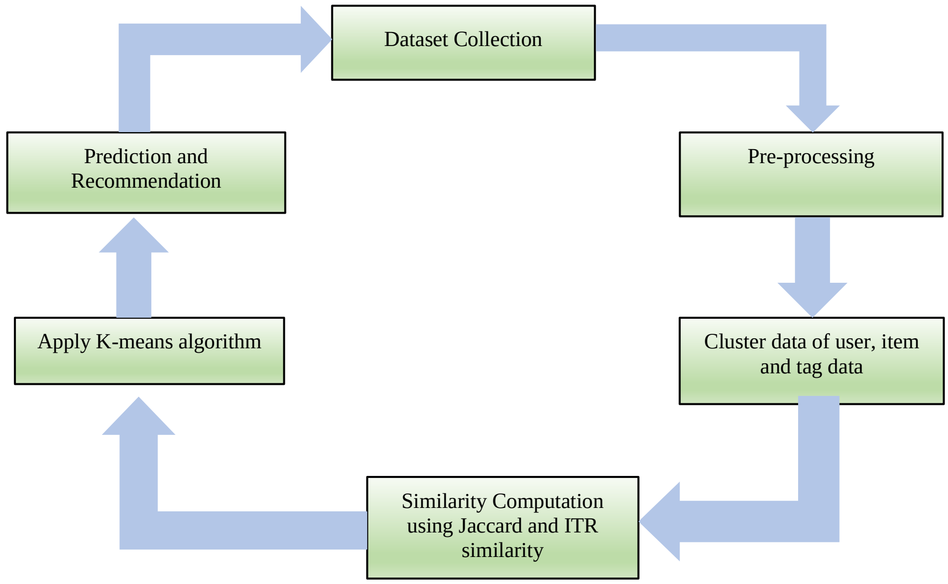

Figure 5. Proposed Framework of Collaborative tagging

The rating matrix in binary classification of user and item would be in the form of either 0 or 1. This matrix is classified into (i) user, item matrix – the user given rating on item which may count 1, otherwise it is 0. (ii) user, tag matrix – The user may assign any tag on a particular product then it counts as 1, otherwise 0. (iii) item, tag matrix – if tags are mentioned on items is 1 otherwise it is 0. Ratings are shown as explicitly in the matrices of user, item and tags. The binary rating matrix and weightages of their tags on items are shown in Figure 6.

$$\begin{array}{cc} & \text{item} \\ \text{user} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \quad \begin{array}{cc} & \text{tag} \\ \text{user} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \quad \begin{array}{cc} & \text{tag} \\ \text{item} & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$\begin{array}{cc} & \text{item} \\ \text{user} & \begin{bmatrix} 5 & 0 & 0 \\ 2 & 0 & 4 \\ 0 & 0 & 3 \end{bmatrix} \end{array} \quad \begin{array}{cc} & \text{tag} \\ \text{user} & \begin{bmatrix} 6 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix} \end{array} \quad \begin{array}{cc} & \text{tag} \\ \text{item} & \begin{bmatrix} 3 & 0 & 0 \\ 5 & 0 & 0 \\ 0 & 0 & 6 \end{bmatrix} \end{array}$$

Figure 6. Implicit and Explicit ratings of user, tag and item matrices

Here the similarity calculations are done using either Jaccard distance or Improved Triangle Similarity (ITR) [3]. When the binary values are considered, the similarity calculation is computed using Jaccard whereas for rating ITR is used. Later both similarity computations are merged together to get high accuracy in recommendation. The similarity computations are used as follows:

For binary data of user, item and tag rating matrices as Equation (1):

$$\frac{a_{11}}{a_{11}+b_{01}+c_{10}} \tag{1}$$

where a_{11} = the cell of matrix consists of 11, b_{01} = the cell consists of 01 and c_{10} = the cell consists of 10.

(a)

	i_1	i_2	i_3	i_4
u_1	1	0	0	1
u_2	1	0	1	0
u_3	0	0	1	1
u_4	1	1	0	1

(b)

	i_1	i_2	i_3	i_4
g_1	1	0	0	1
g_2	1	1	1	0
g_3	0	1	1	0
g_4	1	1	0	1

(c)

	i_1	i_2	i_3	i_4
t_1	1	0	1	1
t_2	1	0	1	0
t_3	0	1	1	1
t_4	1	1	0	1

Figure 7. Similarity Computations on Binary data of items (a) with user, (b) with genre, and (c) with tag data

For example, 7(a): Here in this table, we may compute similarity of {(i_1, i_2), (i_1, i_3), (i_1, i_4), (i_2, i_3), (i_2, i_4), (i_3, i_4)}.

For (i_1, i_2) = 1/1+0+2 = 0.33, (i_1, i_3) = 1/1+1+2 = 0.25, (i_1, i_4) = 2/2+1+1 = 0.5, (i_2, i_3) = 0, (i_2, i_4) = 1/1+2+0 = 0.33, (i_3, i_4) = 1/1+2+1 = 0.25.

Here in these computations i_1 is more similar to i_4 where the similarity value is 0.5. So, in this the blank cells are computed based on this similarity value. In the same way the similarity computations are done of items with genre data and items with tag data as shown in Figure 7 (b) and (c).

For explicit ratings of user, item, and tags the similarity measure used Improved Triangle Similarity (ITR) [23], which was recently proposed in which triangle similarity and user's rating preferences (URP) [23] were computed. The main feature of ITR focuces on both common ratings and non-common ratings. The ITR similarity between two users *x* and *y* is defined as follows in Equation (3):

$$Sim^{ITR}(s,t) = Sim^{TRI}(s,t) * Sim^{URP}(s,t) \tag{2}$$

$Sim^{TRI}(s,t)$ and $Sim^{URP}(s,t)$ are defined in Equations (3) and (4) respectively

$$Sim^{TRI}(s,t) = 1 - \frac{\sqrt{\sum_{i \in I_{st}} (r_{si} - r_{ti})^2}}{\sqrt{\sum_{i \in I_{st}} r_{si}^2} + \sqrt{\sum_{i \in I_{st}} r_{ti}^2}} \tag{3}$$

where I_{st} denotes the set of items and r_{si} and r_{ti} are ratings of users *s* and *t*.

$$Sim^{URP}(s,t) = 1 - \frac{1}{1 + \exp(-|\bar{r}_s - \bar{r}_t| * |\sigma_s - \sigma_t|)} \tag{4}$$

where $\underline{r}_x$ and $\underline{r}_y$ denote the mean ratings of users *u* and *v* on item *i* in I_{xy}, respectively. σ_s and σ_t represent the standard variance of *s* and *t*, respectively as Equations (5).

$$\sigma_s = \sqrt{\frac{\sum_{i \in I_s} (r_{si} - \bar{r}_s)^2}{|I_s|}} \tag{5}$$

where I_s is the set of items rated by the user *s*. Also, the ITR is the measure to compute similarity of two items, genres and tags. Later these similarities if ITR is merged together to get better accuracy in recommendations is shown in Equation (6).

$$Sim^{ITR}(A,B) = Sim^{ITR}(s,t) + Sim^{ITR}(u,v) + Sim^{ITR}(a,b) + Sim^{ITR}(g,h) \tag{6}$$

Where '*s*' and '*t*' are similarity between two users, '*u*' and '*v*' are similarity between two items, '*a*' and '*b*' are similarity between two tags and '*g*' and '*h*' are similarity between two genres.

4. Results and Discussion

There are 100,000 ratings from 943 individuals on 1,682 movies in the MovieLens dataset. The MovieLens dataset in this version is the oldest. At least 20 movies have been rated by each user. The stars are assigned in whole-star increments. This dataset includes user demographic information as well as information on movies, tags, genres, and ratings. The training set for MovieLens datasets will comprise the first 150, 250, and 350 users. Such a random separation was designed for 10-fold cross validation executions, in which all of the tests are repeated a number of times for 150, 250, and 350 users. A total of 30% of all users were tested for the dataset.

The Mean Absolute Error (MAE) is a measure of the difference between RS predictions and user ratings. It is used to assess the performance of recommender systems. The following formula calculates the MAE as Equations (7).

$$MAE = \sum_{i=1}^{n} \frac{(Actual - Prediction\ Values)}{n} \tag{7}$$

On movieLens datasets, the various experiments are conducted. Users' data is modelled 70% for training and 30% for testing in these datasets. The similarity computations are carried out here utilising the average Jaccard and Improved Triangle Similarity (ITR) results applied to the base data of users, tags, genres, and objects with a recent timestamp. User-item, user-tag, and user-genre similarity are all calculated using this information. Several assessment criteria, such as MAE (see Table 2 and 3) precision and F1 measure, are used to compare the outcomes to baseline procedures as Equation (8)-(10).

$$Precision\ (P) = \frac{TR}{TR+FR} \tag{8}$$

$$Recall\ (R) = \frac{TR}{TR+FN} \tag{9}$$

$$F\ measure = \frac{(1+\propto^2).P.R}{\propto^2.P.R} \tag{10}$$

Table 2. MAE of IUGT-Euclidian Similarity (Item similarity computations based on user, genre and tag with Euclidian Distance), IUGT-Jaccard (Item similarity computations based on user, genre and tag with Jaccard Distance) and IUGT-Jaccard-ITR (Item similarity computations based on user, genre and tag with Jaccard Improved Triangle Similarity)

No. of users	IUGT-Euclidian Similarity	IUGT -Jaccard	IUGT -Jaccard-ITR
10	0.859	0.824	0.824
20	0.814	0.813	0.813
30	0.843	0.789	0.789
40	0.804	0.767	0.767
50	0.826	0.754	0.754
60	0.802	0.753	0.753
70	0.798	0.765	0.765
80	0.702	0.734	0.734
90	0.717	0.728	0.728
100	0.724	0.704	0.704

IUGT-Euclidian Similarity (Item similarity computations based on user, genre, and tag with Euclidian Distance), IUGT-Jaccard (Item similarity computations based on user, genre, and tag with Jaccard Distance), and IUGT-Jaccard-ITR (Item similarity computations based on user, genre, and tag with Jaccard Distance with Improved Triangle ratings) are several experiments computed in terms of MAE, F1 measure, and Precision. The above approaches are compared. In comparison, IUGT-Jaccard-ITR achieves the greatest outcomes in top N recommendations, precision, and F1 measure studies, as illustrated in Figures 8 and 9.

In Figures 8 and 9, the training sets are divided into 100, 200, and 300 users, with active users included in each set. In the split of 100 users, where active users are 20, the precision is high because the possible comparisons of number of users with active users are complicated. When the number of active users is between 50 and 80, the precision of IUGT-Jaccard-ITR is higher than the other techniques. Figures 8 and 9 exhibit the precision and F measure findings, which suggest that a system with fewer recommendations and higher accuracy will perform better, and the F measure with number of recommendations will be shown. In this the similarity computations were done among various users, items, tags and genres, where when a cold start user or item arrives in SNS it may give all computations based on UIGT (user, item, genre and tag). So, here the cold start issues might be resolved using different similarities between UIGT with jaccard similarity and improved triangle similarity. Finally the results of IUGT-Jaccard-ITR showed that it gives best results as compared to other approaches in terms of the above mentioned metrics such as precision and F1 measure. It also might avoid the issue of cold start problems.

Table 3. F1-measure and Precision of IUGT-Euclidian Similarity (Item similarity computations based on user, genre and tag with Euclidian Distance), IUGT-Jaccard (Item similarity computations based on user, genre and tag with Jaccard Distance) and IUGT-Jaccard-ITR (Item similarity computations based on user, genre and tag with Jaccard Improved Triangle Similarity)

Top N	IUGT -Jaccard-ITR		IUGT -Jaccard		IUGT-Euclidian Similarity	
	Precision	F1 measure	Precision	F1 measure	Precision	F1 measure
5	0.843	0.814	0.784	0.807	0.803	0.813
10	0.839	0.823	0.786	0.812	0.806	0.817
15	0.836	0.828	0.792	0.823	0.822	0.832
20	0.831	0.843	0.812	0.832	0.833	0.844
25	0.841	0.860	0.823	0.842	0.846	0.857
30	0.840	0.852	0.821	0.841	0.839	0.849
35	0.834	0.846	0.827	0.832	0.832	0.843
40	0.832	0.844	0.822	0.835	0.829	0.841
45	0.831	0.837	0.814	0.812	0.821	0.835
50	0.823	0.838	0.812	0.802	0.819	0.832

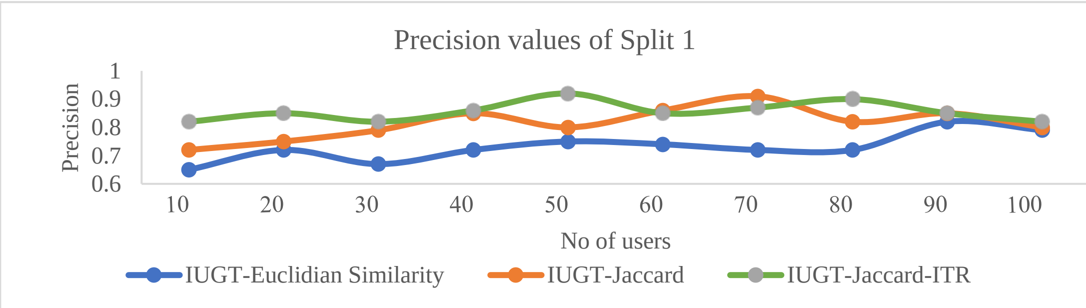

Figure 8. Comparison Precision results of Split 1 of IUGT-Euclidian Similarity, IUGT-Jaccard and IUGT-Jaccard-ITR

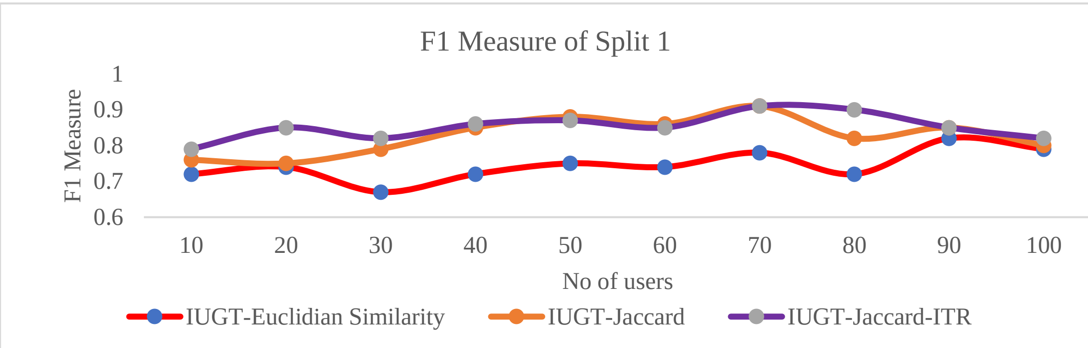

Figure 9. Comparison F1 measure results of Split 1 of IUGT-Euclidian Similarity, IUGT-Jaccard and IUGT-Jaccard-ITR

5. Conclusion

For providing more relevant recommendations to users, a variety of computational intelligence processes such as web based mining and fuzzy logic are being used across practically all sectors of online services. The drawbacks of traditional suggesting approaches such as knowledge-based, content-based, and collaborative-based recommendations are cold start issues where a new user enters into the system or application then analysing these new users is difficult as well as providing recommendations for them. To target these cold start issues, the proposed IUGT-Jaccard-ITR is used for recommendation. In this paper, a framework was developed for addressing cold start issues in collaborative filtering with tagging which results in high-quality suggestions with greater prediction accuracy. To improve the quality and accuracy of recommendations, similarity measures such as the average results of Jaccard and ITR in collaborative tagging with recent timestamp are utilised. Experiments indicate that our proposed technique may greatly enhance prediction quality and accuracy while also addressing cold start challenges. In the future we may focus on personalised recommendation systems with implementation of deep neural networks.

References

1. Adomavicius, G. and Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, vol. 17, no. 6, pp.734-749, 2005.
2. Ayyaz, S., Qamar, U. and Nawaz, R. HCF-CRS: A Hybrid Content based Fuzzy Conformal Recommender System for providing recommendations with confidence. *PloS one*, vol. 13, no. 10, p.e0204849, 2018.
3. Ajaegbu, C. An optimized item-based collaborative filtering algorithm. *Journal of ambient intelligence and humanized computing*, pp. 1-8, 2021.
4. Banda, L. and Bharadwaj, K.K. Evaluation of collaborative filtering based on tagging with diffusion similarity using gradual decay approach. In *Advanced Computing, Networking and Informatics-Volume 1: Advanced Computing and Informatics Proceedings of the Second International Conference on Advanced Computing, Networking and Informatics (ICACNI-2014)*, Springer International Publishing, pp. 421-428, 2014.
5. Bellogín, A., Cantador, I., Díez, F., Castells, P. and Chavarriaga, E. An empirical comparison of social, collaborative filtering, and hybrid recommenders. *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 4, no. 1, pp. 1-29, 2013.
6. Chang, J.H. Mining weighted sequential patterns in a sequence database with a time-interval weight. *Knowledge-Based Systems*, vol. 24, no. 1, pp.1-9, 2011.
7. Banda, L., Singh, K., Abdel-Basset, M., Thong, P.H., Huynh, H.X. and Taniar, D. Recommender systems using collaborative tagging. *International Journal of Data Warehousing and Mining (IJDWM)*, vol. 16, no. 3, pp. 183-200, 2020
8. Halpin, H., Robu, V. and Shepherd, H. The complex dynamics of collaborative tagging. In *Proceedings of the 16th international conference on World Wide Web*, pp. 211-220, 2007.
9. Ko, H., Lee, S., Park, Y. and Choi, A. A survey of recommendation systems: recommendation models, techniques, and application fields. *Electronics*, vol. 11, no. 1, pp. 141, 2002.
10. Banda, L. and Bharadwaj, K.K. An approach to enhance the quality of recommendation using collaborative tagging. *International Journal of Computational Intelligence Systems*, vol. 7, no. 4, pp. 650-659, 2014.
11. Sen, S., Vig, J. and Riedl, J. Tagommenders: connecting users to items through tags. In *Proceedings of the 18th international conference on World wide web*, pp. 671-680, 2009.
12. Heymann, P., Ramage, D. and Garcia-Molina, H. Social tag prediction. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 531-538, 2008.
13. Lika, B., Kolomvatsos, K. and Hadjiefthymiades, S. Facing the cold start problem in recommender systems. *Expert systems with applications*, vol. 41, no. 4, pp. 2065-2073, 2014.
14. Dhelim, S., Aung, N., Bouras, M.A., Ning, H. and Cambria, E. A survey on personality-aware recommendation systems. *Artificial Intelligence Review*, pp. 1-46, 2022.
15. Wang, D., Liang, Y., Xu, D., Feng, X. and Guan, R. A content-based recommender system for computer science publications. *Knowledge-Based Systems*, vol. 157, pp. 1-9, 2018.
16. Polatidis, N. and Georgiadis, C.K. A multi-level collaborative filtering method that improves recommendations. *Expert Systems with Applications*, vol. 48, pp. 100-110, 2016.
17. Bakhshandegan Moghaddam, Farshad & Elahi, Mehdi. Cold-start solutions for recommendation systems. DOI: 10.1049/PBPC035G, 2019.
18. Bobadilla, J., Ortega, F., Hernando, A. and Bernal, J. A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems*, vo. 26, pp. 225-238, 2012.
19. Sen, S., Vig, J. and Riedl, J. Tagommenders: connecting users to items through tags. In *Proceedings of the 18th international conference on World wide web*, pp. 671-680, 2009.
20. Wei, J., He, J., Chen, K., Zhou, Y. and Tang, Z. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, vol. 69, pp. 29-39, 2017.
21. Tey, F.J., Wu, T.Y., Lin, C.L. and Chen, J.L. Accuracy improvements for cold-start recommendation problem using indirect relations in social networks. *Journal of Big Data*, vol. 8, no. 1, pp. 1-18, 2021.
22. Sedhain, S., Menon, A., Sanner, S., Xie, L. and Braziunas, D. Low-rank linear cold-start recommendation from social data. In *Proceedings of the AAAI Conference on Artificial Intelligence* , vol. 31, no. 1, 2017.
23. Yadav, U., Duhan, N. and Bhatia, K.K. Dealing with pure new user cold-start problem in recommendation system based on linked open data and social network features. *Mobile Information Systems*, pp. 1-20, 2020.