Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

RESULTANT VECTOR SIMILARITY MODEL (RVSIM) BASED COLLABORATIVE FILTERING FOR A COMPETENT RECOMMENDATION

VIKRAM MADITHAM *
Research Scholar, Department of CSE, JNTUA University, Ananthapuramu, Andhra Pradesh, India

SUDHAKAR REDDY N
Department of CSE, SV College of Engineering, Tirupati, Andhra Pradesh, India

MADHAVI KASA
Department of CSE, JNTUA University, Ananthapuramu, Andhra Pradesh, India
* Corresponding author's Email: vikram.maditham@gmail.com

Abstract

Recommender systems help the users filter the preferences when large volumes of data are present. However, recommender systems fail to render quality recommendations due to data sparsity problem where users are reluctant to accord feedback. Collaborative filtering technique is one of the quintessential techniques used in recommender systems to mitigate the data sparsity problem. It predicts users' ratings to convert sparse into dense matrix. Similarity models drive collaborative filtering to acquaint the prophecy of users' rating. The existing similarity models work on co-rated items which is not adept to ascertain nearest neighbors in the data. In this paper, we have proposed a new similarity model based on the consideration of user ratings as vectors. The new Resultant vector similarity model (RV Sim) is incorporated in the collaborative filtering algorithm to validate the outcomes using the popular data set Movie Lens. Firstly, resultant vector similarity model is derived based on the resultant vector mathematical model in which vector directions are considered to determine the similarity. Finally, the RV Sim based collaborative filtering algorithm is applied to generate the efficient recommendations. It is imminent that proposed similarity model proved to be superior when it is compared with the existing similarity models (Rjaccard,Cosine, Pearson, Mean square distance etc.) in line with the prominent evaluation metrics i.e. Precision, Recall, F1- Score in the domain of recommender systems. Upon increasing the nearest neighbors 5NN, 20NN, 50NN and 100NN, the RV Sim model has got F1-Score with 0.726, 0.74, 0.753 and 0.763 with considerable improvement

Keywords: Recommender systems, Resultant vector, Collaborative filtering, Similarity model

1. Introduction

Abundant amount of information is available in the current digital world and it is always cumbersome to the users to retrieve the data in accordance with their personal interest. Recommender system (RS) is one of the predominant tools to filter the data based on the user preferences when large catalog is available. According to the survey [28], thirty percent of the revenue is generated through Recommender systems. Now, Recommender systems extended to other areas like [3][5] digital marketing, medical,

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

insurance, news, movie etc. Predominantly, RS helps the customers when they are precarious about suggesting the items based on their past history. Collaborative filtering, content-based filtering and hybrid filtering [28] are indispensable techniques used in RS to predict the ratings. Collaborative filtering [5][7][11][13] has its significant role in the domain of RS to leverage performance and productive approach in real time applications[28] like Amazon, Flipkart etc.

However, collaborative filtering is further divided in to memory based and model approaches [16][20][29], in which memory based approach has been chosen for the current work. Similarity model [1] is the crux of memory based techniques, usually drive the entire recommendation process for better prediction results. Our work focuses on two objectives. Firslty, proposing a new similarity model (RVSim) and secondly determining the prophecy of unknown rating using RVSim based collaborative filtering.

In our new recommendation model, we have used vector-based similarity between two users. It is implemented according to the direction of the vector and the angle between two user-rating vectors. As we know that the resultant of two vectors is a vector that passes in between the given vectors. The magnitude of the resultant vector determines the similarity between two user-rating vectors. If both the vectors are near, then the magnitude of the resultant vector is maximum. In order to normalize the value of magnitude of resultant vector we are dividing it with maximum distance between vectors, which occurred when both are in opposite direction (i. e angle between them need to be 180 deg). Therefore, after normalizing the magnitude the similarity values range between -1 and 0. But in order to compute the unknown rating we need our values to range between 0 and 1. Therefore, we subtract the magnitude value from 1. If the value is closer to 0 it is more likely to be dissimilar and if the value is closer to 1 it is more similar. The detailed demonstration of RVSim is explained in methodology section.

After implementing the formula to find the similarity between two users, we have compared our new similarity metric with other traditional similarity metrics. It found that RVSim has exhibited better performance than other metrics which is proved in section (3). Later we fitted our similarity metric in collaborative filtering in order to find the unknown ratings of respective users. The popular dataset movielens -100K is used to validate the performance of the proposed method and metrics for prediction and top – n recommendations were taken and validated with metrics like F1-score, Precision and Recall, there it exhibits better results compared to other similarity metrics. The next sections comprise of literature survey, methodology and results.

2. Literature Survey

In the literature, extensive research has been taken place to elevate RS as one of prolific research domain. Relevant Jaccard (RJaccard) and Relevant Jaccard Mean square Distance (RJMSD) are the new similarity models introduced by S. Bag et al. [1] . But, in Jaccard they tried to change the most commonly used concept of co-rated items. Instead of co-rated items, they also want to include non-co-rated items between the users. In Enhanced multistage user-based collaborative filtering (EMUCF algorithm) proposed by

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

Jain et al. [2], they want to recommend desired items to new users also. So, they add a feature called rating certain items while the users are signing up. But, one of the major drawbacks of the EMUCF algorithm is many of us will rate the items as max rating or giving rating without testing is not fair. Model-based Collaborate Filtering Algorithm Based on Stacked Auto Encoder (MCFSAE) algorithm M. Yu et al [4], defined to convert the rating two dimensional information to superior dimensional classification dataset with the size of number of ratings. After converting the rating matrix, they filter the features of the items. This model basically recommending the items based on their feature rating. Herce-Zelaya.J et al.[6] introduced a new technique of cold start problem, it states that retrieving the activities of new user from social media. After getting the relative information about user, using different decision algorithms predict the particular items to new user. This algorithm only works for new users.

1. B. Walek and V. Fojtik [8] have come up with Predictory model, it firstly filters the favourite and un popular gernes of a particular user. After the final list of items are recommended using fuzzy expert system. One of drawbacks of the above model is it cannot filter the popular ones too. Instead, it recommends popular ones to all the users of the platform. In shopping application declaring certain items are popular is not possible and it depends on user's interest. Z. Tian and H. Liu[9] introduced and advent of collaborative filtering in order to reduce the problem of less recommendation score due to sparsity of the matrix cause by un rated items, they proposed an algorithm to add number of common scoring weight to the similarity calculation to improve the outcomes. But it doesn't solve the problem of sparsity because, the unrated items are not known and increase similarity by adding weights may increase the complexity. In novel social recommendation method C. H. Lai et al.[10], takes all the possible information about both the user and high rated products. It also checks the relationships around social media to classify the user preferences according to similarity between relationships. Users are recommended based on the factors like relationships, popularity of the product and user interactions on social media. L. Ardissono and N. Mauro [12] implemented compositional model of multi-faced recommendation model, it sates the getting information from social media will may lead to privacy issue to the particular information. Instead, they used the profile and contributions in social networks. One of the major drawbacks is it doesn't resolve the issue of privacy of user.
Recommender systems using Genetic algorithms [14] uses both item-item similarity and user-user similarity in order to recommend a list of items to user. It firstly filters the items based on the ratings got and second filters the items based on the similar ratings and finally, recommending the list of similar and high rated item list to the particular user. But it doesn't solve the problem of sparsity, if the user doesn't given rating to the high rated item, then the user got missed it. In trust based collaborating filtering methods proposed [15] in a three-step process to filter out the dissimilar users and retrieving similar users to recommend the perfect items to the target user. It firstly ranks the users according to the rating values and the social relationships, then filters the top performers similar to the target user through Ant Colony Optimization method. Finally recommend the set items those are rated higher by other filtered users.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

An advent of noise correction-based approach [17], it's main aim is to remove that much noise data from rating matrix. It divides the data into three parts called strong, average and weak and removing noise from it accordingly. Then applying one of the most effective solutions of spar sing "Bhattacharya coefficient" to find the similarity between the users. In Trustor clustering algorithm [18], they state that the solution of cold-start problem will be accessing the relationships over social media is not the right choice. Because, there might be the case that targeted user has no features in common to anther user. They tried to implement the feature that will help in identifying the correct users related to the targeted user. In Trust aware recommendation system [19], they have considered that the trust metric between two users might be great in more situations and might not work in other cases. In order to avoid this confusion weights are added properly to all the trust values between two users and optimized until we get more accurate recommendations. The main aim of trust aware network (TAN) [21] Recommendation system is to create trust clusters based on the social network service. Optimization of these trust clusters will improve the ability to make similarity between two users. They also explore the power user's impact out there in social media and recommending those products to the users. But if the targeted user is not impacted by power user, then the product recommended is not taken. By defining the correlation between items affect more than correlation between users. In genetic algorithm, they finalize that by creating a sub group of user item after filtering the correlations o items and users can create more similarity than normal user-user similarity. By using evolutionary algorithm [15] [22] to find the subset of high correlated items and making a subgroup using least squares. QPSO Recommendation system [23], it states that the collaborative filtering predicts the ratings much easily and effectively. But if we create cluster of users by grouping the similar users will enhance the efficiency of the algorithm. Clustering based on quantum behaved swarm optimization leads to better results.

In hybrid recommendation system [24], it uses two or more recommendation systems in different ways to benefit from its advantages. But if we use two recommendations with same advantages lead to same recommendation as output. In order to achieve more accurate results, we need to combine the algorithms that has the solutions of both spar sing and cold start. In behavior or voted recommendation system [25], it states that the number of up votes decide the similarity between users. It allows the interest of users to be considered and voted accordingly. There is also a restriction that number of up votes won't cross more than 5. A framework for collaborative filtering recommendation system [26], in this algorithm it states the evaluating the novelty of the users and trust metric among the neighbors is most important that just normal user similarity through ratings. But if a user has good novelty score and bad correlation between ratings of users can lead to dissimilarity. In social collaborative filtering by trust [27], it states that in order to mitigate sparse data and cold start users they tried to integrate the two sparse data rendered by users and sparse social trust network data common to same user. But even combing the information of two sparse data it is difficult to achieve more accurate similarity between users. In factored similarity models [30], with social trust for superior – N item recommendation, it uses the factored similarity models along with the social media networks. It added matrix factorization technique to obtain the user interests based on

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

rated items and unrated items. A new similarity measure using Bhattacharya coefficient [32] tells that if we use normal collaborative filtering it will consider only co-rated items. Instead, we use all the rating items between pair of users and finding importance of each pair of ratings. But if we use every rating will lead to unwanted use of those ratings. Leveraging Multiview's of trust and similarity [33] states that the users are clustered based on the ratings of items and social network relationships. After completing the clusters, we predict the rating based on user-item rating prediction using support vector prediction. From the literature, it is resolved that our RVSim model is superior to all the existing similarity models.

3. Methodology

We know that the user item rating is a vector, then the distance between the user 'X' and user 'Y' rating vectors is d. Of late from literature, research is disseminated on co-rated items nevertheless not on vector data. When emphasis turns on vector data, the similarity model outshines in a better way.

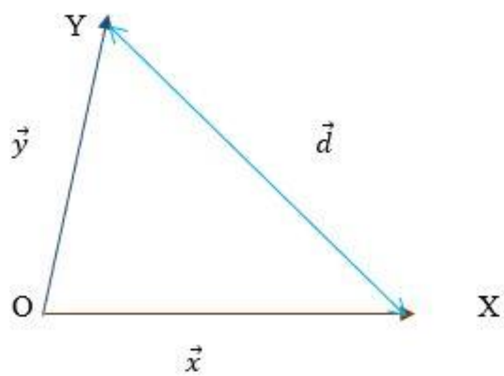

Figure. 1 X & Y Vectors with distance 'd'

As shown in the Figure 1., the vector that joins O(origin) to the X is $\vec{x}$, and the vector that joins O to the Y is $\vec{y}$, and the vector that joins the X and Y is $\vec{d}$. Here, if we want to find the distance between X and Y then either vector has to be moved the other vector's direction. In our model vector X moves towards O to reach Y.

The distance from X to O is $-\bar{x}$ and the distance from O to Y is $\bar{y}$. These two variants are specified in eq1 and eq2.

$$X \rightarrow O = -\vec{x} \quad \dashrightarrow \quad Eq_1$$

$$O \rightarrow Y = \vec{y} \quad \dashrightarrow \quad Eq_2$$

Note that the vector that joins X and O is $-\vec{x}$ because X moves in reverse direction. The vector $\vec{d}$ is calculated by adding above said two vectors. Hence, the result can be seen in eq3.

$$\vec{d} = -\vec{x} + \vec{y} \quad \dashrightarrow \quad Eq_3$$

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

As we already know the resultant of the two vectors X and Y is as shown below.

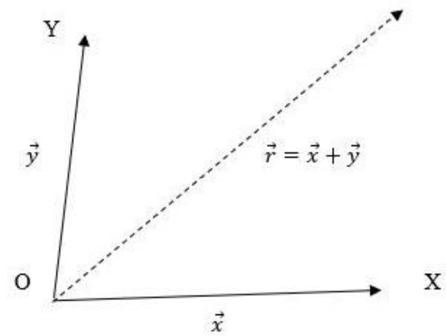

Figure. 2 Resultant of the Two Vectors X & Y

Now, we know that the resultant formula $\vec{R}$ is

$$\vec{R} = \sqrt{x^2 + y^2 + 2(x)(y)\cos\theta} \qquad \dashrightarrow \quad eq_4$$

But if we want to know the distance between X and y, then $\vec{D}$ will become

Substitute -x in place x in above equation 4:

$$\vec{D} = \sqrt{(-x)^2 + y^2 + 2(-x)(y)\cos\theta}$$

$$\therefore \vec{D} = \sqrt{x^2 + y^2 - 2(x)(y)\cos\theta} \qquad \dashrightarrow \quad eq_5$$

After finding the$\vec{D}$, if the distance $\vec{D}$ decreases then the two vectors are more similar. In order to normalize the above equation, we divide the above equation with the maximum distance between the users.

The distance between the users become maximum only when the angle θ between users is 180°. Let us consider maximum distance as $\vec{M}$. Then our formula becomes as follows.

$$\vec{M} = \sqrt{x^2 + y^2 - 2(x)(y)\cos 180^\circ}$$

$$(\because \cos 180^\circ = -1)$$

$$= \sqrt{x^2 + y^2 + 2(x)(y)}$$

$$\therefore \vec{M} = |x + y| \quad \dashrightarrow \quad eq_6$$

At last our similarity metric becomes sim_r as :

$$sim_r = \left[\frac{\vec{D}}{\vec{M}}\right] \qquad (\because \text{from } eq_5 \text{ and } eq_6)$$

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

$$= \left[\frac{\sqrt{x^2+y^2-2(x)(y)\cos\theta}}{(x+y)}\right]$$

$$= \left[\frac{\sqrt{x^2+y^2-2(x)(y)\frac{\sum x.y}{(x)(y)}}}{|x+y|}\right]$$

$$\left(\because\ \cos(x,y) = \frac{\sum x.y}{|x||y|}\right)$$

$$= \left[\frac{\sqrt{x^2+y^2-2\sum x.y}}{|x+y|}\right]$$

$$\therefore \text{sim}_r = \left[\frac{\sqrt{x^2+y^2-2\sum x.y}}{|x+y|}\right] \quad \dashrightarrow \quad eq_7$$

Where, $x = \sum_{i\in Ix,y)} R(x,i)^2$

$$y = \sum_{i\in I(x,y)} R(y,i)^2$$

$$\sum x.y = \sum_{i\in I(x,y)}[R(x,i).R(y,i)]$$

$$|x+y| = \sqrt{\sum_{i\in I(x,y)}\left(R(x,i)+R(y,i)\right)^2}$$

If we want to calculate the unknown rating the most similar user should have maximum rating. Then we subtract the above equation7 from 1. Therefore our similarity metric becomes:

$$\text{sim}_r = 1 - \left[\frac{\sqrt{\sqrt{\sum_{i\in Ix,y)} R(x,i)^2}^2 + \sqrt{\sum_{i\in I(x,y)} R(y,i)^2}^2 - 2\left(\sum_{i\in I(x,y)}[R(x,i).R(y,i)]\ \right)}}{\sqrt{\sum_{i\in I(x,y)}\left(R(x,i)+R(y,i)\right)^2}}\right]$$

$$\therefore \text{sim}_r = 1 - \left[\frac{\sqrt{\sum_{i\in Ix,y)} R(x,i)^2+\sum_{i\in I(x,y)} R(y,i)^2-2\left(\sum_{i\in I(x,y)}[R(x,i).R(y,i)]\ \right)}}{\sqrt{\sum_{i\in I(x,y)}\left(R(x,i)+R(y,i)\right)^2}}\right] \dashrightarrow eq_8$$

The Eq(8) represents the final RV Sim model which has been verified on small user and item matrix as specified in Table1 and compared with erstwhile techniques.

Table1: Sample Use and Item matrix

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

U/I	I1	I2	I3	I4
X1	3	2	4	3
X2	4	3	-	-
X3	5	4	3	2
X4	1	1	-	-

$$\begin{array}{cc} & \begin{matrix} \mathrm{X1} & \mathrm{X2} & \mathrm{X3} & \mathrm{X4} \end{matrix} \\ \begin{matrix} \mathrm{X1} \\ \mathrm{X2} \\ \mathrm{X3} \\ \mathrm{X4} \end{matrix} & \begin{bmatrix} - & 0.998460 & 0.905097 & 0.980580 \\ & - & 0.999512 & 0.989949 \\ & & - & 0.993883 \\ & & & - \end{bmatrix} \end{array}$$

(a) COS

$$\begin{array}{c} \begin{matrix} \mathrm{X1} & \mathrm{X2} & \mathrm{X3} & \mathrm{X4} \end{matrix} \\ \begin{bmatrix} - & 0.832050 & -0.316227 & 0 \\ & - & 0.894427 & 0 \\ & & - & 0 \\ & & & - \end{bmatrix} \end{array}$$

(b) PCC

$$\begin{array}{c} \begin{matrix} \mathrm{X1} & \mathrm{X2} & \mathrm{X3} & \mathrm{X4} \end{matrix} \\ \begin{bmatrix} - & 0.832050 & -0.316227 & 0 \\ & - & 0.875246 & 0 \\ & & - & 0 \\ & & & - \end{bmatrix} \end{array}$$

(c) CPC

$$\begin{array}{cc} \begin{matrix} \mathrm{X1} \\ \mathrm{X2} \\ \mathrm{X3} \\ \mathrm{X4} \end{matrix} & \begin{bmatrix} - & 0.444529 & 0.845471 & 0.303883 \\ & - & 0.468765 & 0.562347 \\ & & - & 0.380394 \\ & & & - \end{bmatrix} \end{array}$$

(d) JMSD

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

$$\begin{bmatrix} - & 0.889059 & 0.845471 & 0.607767 \\ & - & 0.937530 & 0.562347 \\ & & - & 0.760788 \\ & & & - \end{bmatrix}$$

(e) MSD

$$\begin{bmatrix} - & 0.5 & 1 & 0.5 \\ & - & 0.5 & 1 \\ & & - & 0.5 \\ & & & - \end{bmatrix}$$

(f)RJaccard

$$\begin{bmatrix} - & 0.835601 & 0.760268 & 0.653589 \\ & - & 0.9465 & 0.436907 \\ & & - & 0.778233 \\ & & & - \end{bmatrix}$$

(g) RVSIm)

Difficulty in Cosine Similarity:

As per the above observations we can state that x1 and x3 are more similar when compared to other users based on cosine similarity. But when we observe the cosine similarity between x3 and x4 is also higher, even though both are not similar to each other.

Combination of Jaccard and MSD (JSMD):

It is observed that the similarities between x1 & x4 and between x3 & x4 are similar. However, their rating vector shows that the similarity between x1 and x4 is much higher than similarity between x3 and x4.

PCC & CPC Similarities:

According to rating vector x1 and x3 are similar in nature, but if we consider PCC and CPC similarities, the values are in negative which represents that they are dissimilar in behaviour.

RV Sim (New Similarity):

The correlation between all the users is identified accurately when compared with all other similarity models.

With the help Table2, RV Sim is evaluated manually how it responds for rating values in terms of similarity and non-similarity.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

Table 2 : Sample2 for User and Item rating matrix

Users / Items	I1	I2	I3	I4	I5	I6	I7	I8
X1	3	?	4	2	?	5	1	1
X2	3	-	-	2	-	-	-	1
X3	3	-	4	-	-	5	-	-
X4	-	-	-	2	-	-	1	1
X5	2	2	-	2	5	4	-	2
X6	1	3	-	-	4	-	3	4

In the above table, if we consider X1 as targeted user, where we need to find the rating of respective item I2. Here X1 and X2 are rated three items similarly, then both the users have most similarity. In our similarity metric, the distance between the X1 and X2 users is less than the value of similarity is nearer to 0. If the similarity value is nearer to 1 then they have high similarity, whereas similarity value is nearer to 0 then they has less similarity. Our similarity has values between 0 to 1.

Let us find the similarity between the users X1 and X2:

$$\mathrm{sim_r} = 1 - \left[\frac{\sqrt{3^2+2^2+1^2+3^2+2^2+1^2-2\big((3)(3)+(2)(2)+(1)(1)\big)}}{\sqrt{(3+3)^2+(2+2)^2+(1+1)^2}}\right]$$

$$= 1 - \left[\frac{\sqrt{9+4+1+9+4+1-28}}{\sqrt{36+16+4}}\right]$$

$$= 1 - \left[\frac{\sqrt{28-28}}{\sqrt{58}}\right]$$

$$= 1 - 0$$

$$= 1$$

As we can see the similarity of X1 and X2 is 1. Therefore X1 and X2 is most similar.

Whereas X1 and X5 are dissimilar because they rated items mostly dissimilar.

Then similarity between X1 and X6 is:

$$\mathrm{sim_r} = 1 - \left[\frac{\sqrt{3^2+1^2+1^2+1^2+3^2+4^2-2\big((3)(1)+(1)(3)+(1)(4)\big)}}{\sqrt{(3+1)^2+(1+3)^2+(1+4)^2}}\right]$$

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

$$= 1 - \left[\frac{\sqrt{9+1+1+1+9+16-2(3+3+4)}}{\sqrt{9+9+16}}\right]$$

$$= 1 - \left[\frac{\sqrt{37-20}}{\sqrt{34}}\right]$$

$$= 1 - \left[\frac{\sqrt{17}}{\sqrt{34}}\right]$$

$$= 1 - 0.7071067811865475$$

$$= 0.2928932188134525$$

As we can observe the similarity between X1 and X5 is nearer to 0 then we can say they are dissimilar.

After due calculations of the similarity value using sample values, it will be applied in standard user to user collaborative filtering algorithm. The steps are envisaged in Algorithm1 and step5 specifies our RVSim model and step14 specifies the actual collaborative prophecy algorithm where ratings are forecasted.

<u>Algorithm1</u>

Input: User - Item rating metric using RVSim

Output: Predicted rating values

1. RVSim (a, b) = [ ], Rt*(a, i) = [ ]
2. Repeat
3. Repeat
4. if a != b and $|I_a \cap I_b|$ then
5. RVSim(a, b) = $1 - \left[\frac{\sqrt{\sum_{i\in I(a,b)} Rt(a,i)^2 + \sum_{i\in I(a,b)} Rt(b,i)^2 - 2\left(\sum_{i\in I(a,b)}[Rt(a,i).Rt(b,i)]\right)}}{\sqrt{\sum_{i\in I(a,b)}\left(Rt(a,i)+Rt(b,i)\right)^2}}\right]$
6. else
7. RVSim (a, b) = 0
8. end if
9. Until user b = B
10. Until user a = A
11. Repeat
12. Repeat
13. if Rt (a, i) == 0 and RVSim(a)>0 then

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

14. Rt* (a, i) = $\overline{\mathrm{Rt(a)}} + \frac{\sum_{b \in S(b)} \mathrm{RVSim(a,b)}.\left(\mathrm{Rt(b,i)} - \overline{\mathrm{Rt(b)}}\right)}{\sum_{b \in S(b)} |\mathrm{RVSim(a,b)}|}$

15. else

16. Rt*(a, i) = 0

17.end if

18. Until the user b = B

19. Until the user a = A

After the productive completion of rating perdition for unknown ratings, the evaluation metrics are applied to affirm the model accuracy.

3. Results

The efficacy of the proposed method is evaluated by working on popular dataset Group lens (100K) where minimum rating is 1 and maximum is 5. 943 users have rendered 1682 ratings. At the outset the RV Sim is compared with popular similarity model RJaccard [1] which is derived from Jaccard similarity model. F1- score is impressive in terms of efficacy for the proposed method when nearest neighbors are varying periodically in Fig1. In the subsequent steps, the proposed method is evaluated for other metrics (Recall, Precision and F1-score) for determining the top n-recommendations and is compared with erstwhile techniques. Table2 signifies recall, Table 3 signifies precision and Table signifies the F4-score results and comparative results of erstwhile techniques.

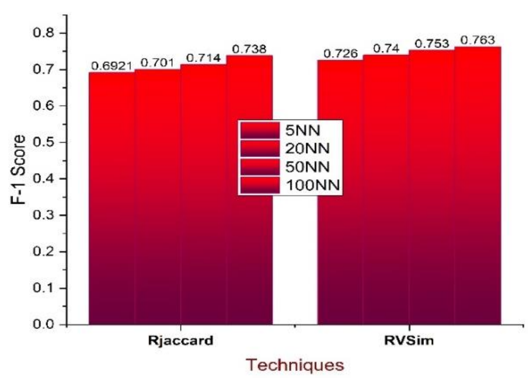

Fig 4. F1-Score evaluation among Rjaccard and RVSim

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

Table3. Recall results of different recommendation models

Dataset	Recommendation Model	N				
		Top-2	Top-4	Top-6	Top-8	Top-10
Movie lens	Rjaccard[1]	0.232	0.275	0.29	0.301	0.323
	EMUCF[2]	0.275	0.342	0.368	0.4	0.413
	MCFSAE[4]	0.43	0.527	0.57	0.612	0.631
	BILGA[14]	0.31	0.422	0.468	0.5	0.54
	TCFCAO[15]	0.295	0.331	0.4	0.468	0.55
	TAN[21]	0.579	0.58	0.626	0.659	0.678
	RVSim	0.543	0.592	0.664	0.752	0.755

Dataset	Recommendation Model	N				
		Top-12	Top-14	Top-16	Top-18	Top-20
Movie lens	Rjaccard[1]	0.35	0.3697	0.391	0.413	0.494
	EMUCF[2]	0.441	0.46	0.47	0.526	0.557
	MCFSAE[4]	0.665	0.67	0.7	0.71	0.751
	BILGA[14]	0.574	0.609	0.665	0.7	0.74
	TCFCAO[15]	0.59	0.62	0.67	0.697	0.721
	TAN[21]	0.711	0.745	0.776	0.808	0.834
	RVSim	0.771	0.762	0.794	0.821	0.854

Table4. Precision results of different recommendation models:

Dataset	Recommendation Model	N				
		Top-2	Top-4	Top-6	Top-8	Top-10
Movie lens	Rjaccard[1]	0.351	0.34	0.335	0.331	0.33
	EMUCF[2]	0.359	0.349	0.338	0.338	0.335
	MCFSAE[4]	0.612	0.595	0.591	0.587	0.581
	BILGA[14]	0.485	0.472	0.436	0.39	0.38
	TCFCAO[15]	0.413	0.385	0.358	0.35	0.355

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

	TAN[21]	0.735	0.711	0.7	0.691	0.682
	RVSim	0.864	0.895	0.852	0.752	0.755

Dataset	**Recommendation Model**	**N**				
		Top-12	Top-14	Top-16	Top-18	Top-20
Movie lens	Rjaccard[1]	0.323	0.328	0.32	0.315	0.31
	EMUCF[2]	0.331	0.33	0.329	0.324	0.32
	MCFSAE[4]	0.578	0.574	0.57	0.565	0.561
	BILGA[14]	0.37	0.362	0.363	0.36	0.34
	TCFCAO[15]	0.351	0.35	0.349	0.349	0.344
	TAN[21]	0.677	0.665	0.653	0.652	0.649
	RVSim	0.771	0.762	0.794	0.821	0.854

Table5. F1-Score results of different recommendation models:

Dataset	Recommendation Model	N				
		Top-2	Top-4	Top-6	Top-8	Top-10
Movie lens	Rjaccard[1]	0.269	0.305	0.310	0.325	0.326
	EMUCF[2]	0.312	0.345	0.352	0.359	0.369
	MCFSAE[4]	0.505	0.558	0.580	0.599	0.605
	BILGA[14]	0.378	0.445	0.451	0.438	0.446
	TCFCAO[15]	0.344	0.356	0.377	0.400	0.431
	TAN[21]	0.647	0.638	0.660	0.674	0.68
	RVSim	0.686	0.726	0.777	0.829	0.837

Dataset	**Recommendation Model**	**N**				
		Top-12	Top-14	Top-16	Top-18	Top-20
Movie lens	Rjaccard[1]	0.336	0.348	0.352	0.357	0.381
	EMUCF[2]	0.378	0.384	0.387	0.398	0.406
	MCFSAE[4]	0.618	0.618	0.628	0.629	0.642
	BILGA[14]	0.45	0.454	0.469	0.475	0.465

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

	TCFCAO[15]	0.440	0.442	0.451	0.465	0.465
	TAN[21]	0.693	0.702	0.709	0.721	0.731
	RVSim	0.838	0.832	0.847	0.855	0.869

5. Conclusion and Future Enhancement

In this work, a new similarity model RVSim has been developed using the basic concept of resultant vector and tested for suitable recommendations. It was compared with erstwhile similarity models i.e COS, PEARSON, RJACCARD etc. and proven to be superior among all the techniques. The similarity model was incorporated in collaborative filtering algorithm in order to predict the ratings. The efficacy of the proposed RVSim based collaborating technique was compared in line with evaluation metrics Recall, Precision and F1-score. The test results affirmed that proposed method is performing quite impressively and can be extended to real time scenario. The work has been carried out using explicit feedback nevertheless, not considered implicit feedback. Considerable improvement could be observed upon combining the both feedback as a future work.

References

1) S. Bag, S. K. Kumar, and M. K. Tiwari, “An efficient recommendation generation using relevant Jaccard similarity,” Inf. Sci. (Ny)., vol. 483, pp. 53–64, 2019, doi: 10.1016/j.ins.2019.01.023.
2) A. Jain, S. Nagar, P. K. Singh, and J. Dhar, “EMUCF: Enhanced multistage user-based collaborative filtering through non-linear similarity for recommendation systems,” Expert Syst. Appl., vol. 161, p. 113724, 2020, doi: 10.1016/j.eswa.2020.113724.
3) Z. Su, Z. Lin, J. Ai, and H. Li, “Rating Prediction in Recommender Systems Based on User Behavior Probability and Complex Network Modeling,” IEEE Access, vol. 9, pp. 30739–30749, 2021, doi: 10.1109/ACCESS.2021.3060016.
4) M. Yu, T. Quan, Q. Peng, X. Yu, and L. Liu, “A model-based collaborate filtering algorithm based on stacked AutoEncoder,” Neural Comput. Appl., vol. 8, 2021, doi: 10.1007/s00521-021-05933-8.
5) A. Gazdar and L. Hidri, “A new similarity measure for collaborative filtering based recommender systems,” Knowledge-Based Syst., vol. 188, p. 105058, 2020, doi: 10.1016/j.knosys.2019.105058.
6) J. Herce-Zelaya, C. Porcel, J. Bernabé-Moreno, A. Tejeda-Lorente, and E. Herrera-Viedma, “New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests,” Inf. Sci. (Ny)., vol. 536, pp. 156–170, 2020, doi: 10.1016/j.ins.2020.05.071.
7) L. Al Hassanieh, C. A. Jaoudeh, J. B. Abdo, and J. Demerjian, “Similarity measures for collaborative filtering recommender systems,” 2018 IEEE Middle East North Africa Commun. Conf. MENACOMM 2018, pp. 1–5, 2018, doi: 10.1109/MENACOMM.2018.8371003.
8) B. Walek and V. Fojtik, “A hybrid recommender system for recommending relevant movies using an expert system,” Expert Syst. Appl., vol. 158, p. 113452, 2020, doi: 10.1016/j.eswa.2020.113452.
9) Z. Tian and H. Liu, “Collaborative filtering recommendation algorithm based on improved denoising auto encoder,” Proc. - 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. ITCA 2020, no. Itoec, pp. 23–28, 2020, doi: 10.1109/ITCA52113.2020.00012.
10) C. H. Lai, S. J. Lee, and H. L. Huang, “A social recommendation method based on the integration of social relationship and product popularity,” Int. J. Hum. Comput. Stud., vol. 121, no. April 2017, pp. 42–57, 2019, doi: 10.1016/j.ijhcs.2018.04.002.
11) S. Natarajan, S. Vairavasundaram, S. Natarajan, and A. H. Gandomi, “Resolving data sparsity and cold start problem in collaborative filtering recommender system using Linked Open Data,” Expert

Syst. Appl., vol. 149, 2020, doi: 10.1016/j.eswa.2020.113248.
12) L. Ardissono and N. Mauro, "A compositional model of multi-faceted trust for personalized item recommendation," Expert Syst. Appl., vol. 140, p. 112880, 2020, doi: 10.1016/j.eswa.2019.112880.
13) R. Chen, Q. Hua, Y. S. Chang, B. Wang, L. Zhang, and X. Kong, "A survey of collaborative filtering-based recommender systems: from traditional methods to hybrid methods based on social networks," IEEE Access, vol. 6, pp. 64301–64320, 2018, doi: 10.1109/ACCESS.2018.2877208.
14) B. Alhijawi and Y. Kilani, "A collaborative filtering recommender system using genetic algorithm," Inf. Process. Manag., vol. 57, no. 6, p. 102310, 2020, doi: 10.1016/j.ipm.2020.102310.
15) H. Parvin, P. Moradi, and S. Esmaeili, "TCFACO: Trust-aware collaborative filtering method based on ant colony optimization," Expert Syst. Appl., vol. 118, pp. 152–168, 2019, doi: 10.1016/j.eswa.2018.09.045.
16) Z. Khan, N. Iltaf, H. Afzal, and H. Abbas, "Enriching Non-negative Matrix Factorization with Contextual Embeddings for Recommender Systems," Neurocomputing, vol. 380, pp. 246–258, 2020, doi: 10.1016/j.neucom.2019.09.080.
17) S. Bag, S. Kumar, A. Awasthi, and M. K. Tiwari, "A noise correction-based approach to support a recommender system in a highly sparse rating environment," Decis. Support Syst., vol. 118, no. August 2018, pp. 46–57, 2019, doi: 10.1016/j.dss.2019.01.001.
18) J. Lee, G. Noh, H. Oh, and C. kwon Kim, "Trustor clustering with an improved recommender system based on social relationships," Inf. Syst., vol. 77, pp. 118–128, 2018, doi: 10.1016/j.is.2018.05.008.
19) S. Yadav, V. Kumar, S. Sinha, and S. Nagpal, "Trust aware recommender system using swarm intelligence," J. Comput. Sci., vol. 28, pp. 180–192, 2018, doi: 10.1016/j.jocs.2018.09.007.
20) X. Wang, Y. Liu, J. Lu, F. Xiong, and G. Zhang, "TruGRC: Trust-Aware Group Recommendation with Virtual Coordinators," Futur. Gener. Comput. Syst., vol. 94, pp. 224–236, 2019, doi: 10.1016/j.future.2018.11.030.
21) G. Noh, H. Oh, and J. Lee, "Power users are not always powerful: The effect of social trust clusters in recommender systems," Inf. Sci. (Ny)., vol. 462, pp. 1–15, 2018, doi: 10.1016/j.ins.2018.05.058.
22) A. Laishram, V. Padmanabhan, and R. P. Lal, "Analysis of similarity measures in user-item subgroup based collaborative filtering via genetic algorithm," Int. J. Inf. Technol., vol. 10, no. 4, pp. 523–527, 2018, doi: 10.1007/s41870-018-0195-z.
23) R. Logesh, V. Subramaniyaswamy, V. Vijayakumar, X. Z. Gao, and V. Indragandhi, "A hybrid quantum-induced swarm intelligence clustering for the urban trip recommendation in smart city," Futur. Gener. Comput. Syst., vol. 83, pp. 653–673, 2018, doi: 10.1016/j.future.2017.08.060.
24) E. Çano and M. Morisio, "Hybrid recommender systems: A systematic literature review," Intell. Data Anal., vol. 21, no. 6, pp. 1487–1524, 2017, doi: 10.3233/IDA-163209.
25) J. Bobadilla, F. Serradilla, and J. Bernal, "A new collaborative filtering metric that improves the behavior of recommender systems," Knowledge-Based Syst., vol. 23, no. 6, pp. 520–528, 2010, doi: 10.1016/j.knosys.2010.03.009.
26) J. Bobadilla, A. Hernando, F. Ortega, and J. Bernal, "A framework for collaborative filtering recommender systems," Expert Syst. Appl., vol. 38, no. 12, pp. 14609–14623, 2011, doi: 10.1016/j.eswa.2011.05.021.
27) B. Yang, Y. Lei, J. Liu, and W. Li, "Social Collaborative Filtering by Trust," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 8, pp. 1633–1647, 2017, doi: 10.1109/TPAMI.2016.2605085.
28) J. Bobadilla, F. Ortega, A. Hernando, and J. Bernal, "A collaborative filtering approach to mitigate the new user cold start problem," Knowledge-Based Syst., vol. 26, pp. 225–238, 2012, doi: 10.1016/j.knosys.2011.07.021.
29) J. Bobadilla, F. Ortega, and A. Hernando, "A collaborative filtering similarity measure based on singularities," Inf. Process. Manag., vol. 48, no. 2, pp. 204–217, 2012, doi: 10.1016/j.ipm.2011.03.007.
30) G. Guo, J. Zhang, F. Zhu, and X. Wang, "Factored similarity models with social trust for top-N item recommendation," Knowledge-Based Syst., vol. 122, pp. 17–25, 2017, doi: 10.1016/j.knosys.2017.01.027.
31) X. Ma, H. Lu, Z. Gan, and J. Zeng, "An explicit trust and distrust clustering based collaborative filtering recommendation approach," Electron. Commer. Res. Appl., vol. 25, pp. 29–39, 2017, doi:

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:55 Issue:05:2022
DOI 10.17605/OSF.IO/QXFGZ

10.1016/j.elerap.2017.06.005.
32) B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, “A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data,” Knowledge-Based Syst., vol. 82, no. March, pp. 163–177, 2015, doi: 10.1016/j.knosys.2015.03.001.
33) G. Guo, J. Zhang, and N. Yorke-Smith, “Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems,” Knowledge-Based Syst., vol. 74, pp. 14–27, 2015, doi: 10.1016/j.knosys.2014.10.016.