*International Journal of Computer Applications in Technology, Vol. x, No. x, 201X* 1

# Optimum Personalized Confirmed Quality Aware Recommendations

**Seema P Nehete**

Datta Meghe College of Engg, Navi Mumbai, Airoli, India

E-mail: [nehete.seema@gmail.com](mailto:nehete.seema@gmail.com)

# Satish R Devane

Datta Meghe College of Engg, Navi Mumbai, Airoli, India

E-mail: [srdevane@yahoo.com](mailto:srdevane@yahoo.com)

**Abstract:** Recommendation System(RS) save the time of users in their hectic life schedules for purchasing their interested products.RS faces challenges of data sparsity, cold start, efficiency of prediction of products and hence the proposed system is making use of Multi-kernel Fuzzy C Means (MKFCM) clustering to group together similar users having similar age, occupation, and gender into clusters. Clusters of similar users are optimized using the Fruit Fly (FF) optimization algorithm which gives high cluster accuracy and dynamically created subclusters of similar users and their favorite products, overcome sparsity issue which make the analysis easy. Collaborative Filtering(CF), one of the filtering method of RS is used to predict products for target users.This RS gains user’s faith by additionally performing analysis of textual reviews using optimized Artificial Neuron Network(ANN) to recommend the highest quality products, thus dual tested and quality confirmed products are recommended to the user. Experimentation is done on a standard movilense dataset used by many researchers to prove the efficiency of this RS and reviews of all users are extracted from online search engines for product quality analysis before recommendation. Experimentation proves highest recall and accuracy than existing recommendation systems.

## Introduction

Every buyer looks for the best product for himself before purchasing as a natural human tendency. Availability of ample products and shortage of time to research about the product, confuses user for purchasing. The buyers of every age group be it a teenager or old age are very busy with their hectic life schedules and hence prefer online shopping. All buyers rarely get to meet together friends,relatives due to these busy hectic schedules and lack of time, therefore they can not physically view products purchased by each others. Large size of data (product, movies, web series, books, songs) is available for online shopping, where buyers get confused for purchase. Buyers are not aware of actual defects in the products in which they are interested, as a result, might end up with low-quality products as they can’t see the physical product on online shopping. Among a large volume of users, products, and reviews available across the web, a Recommendation System (RS) helps target users, buyers for purchasing by predicting a list of products for them. The RS filters similar users for target user among millions of users using filtering methods.((Pazzani and Billsus, 2007)( Lu, Jie.,2010) (Bobadilla, 2013) have already presented different filtering methods for RS.

Collaborative Filtering (CF), one of the filtering method of the RS is mainly used to learn a user’s past purchasing history and recommend him list of products by finding similar users.

Once the target user is explored to different products from similar users, many new products can be added to his taste due to the advantage of CF over another famous filtering method,content-based filtering (Pazzani and Billsus, 2007). In content-based filtering, the user may reach the wrong product at the end not as per need, CF makes use of ratings given by users usually in the scale of 1-5 to their purchase products to calculate similar users with the belief that if two users have same liking in the past then they will agree in future also. Similarity measures used till now in CF are not efficient to give relevant predictions to the target user, if the dataset is sparse and the user might be end up with low- quality irrelevant recommendations.

The number of nearest similar users for the target user are directly proportional to number of relevant products that need to be recommended. Users who have just strated online purchasing with single product will never get proper recommendations. Traditional similarity measures, due to consideration of the average rating of customers does not give recommendations to these users. From a business point of view, RS should

help for adding new customers to increase revenue. Star ratings calculated as average of ratings given by users to their purchased products, which are provided by the e-commerce website do not show all negative textual reviews of products that might have experiences of dissatisfied customers for the same product. These negative textual reviews help to decide the quality of the product. Several approaches and systems for the recommendation were developed, but very few have focused on the integration of analysis of textual negative reviews in the recommendation process. From 2000 till 2019, we found only a few studies relevant to quality-aware product recommendations using sentiment analysis of reviews [Zhang,2013].

Therefore a clustering-based CF approach for recommendation system is proposed in this research, which next aims at analyzing positive and negative reviews of products that need to be recommended to users. The RS uses Multi-kernel fuzzy c means (MKFCM) clustering to narrow down the volume of users to search for similar users.The clusters of users are optimized using fruit fly(FF) optimization algorithm. Technically, this RS is enacted around two stages. In the first stage, Users are clustered using user’s attributes . Then heterogeneous subclusters are created with dynamic clustering that collect together users and their favorite products to overcome data sparsity issue. At the second stage, a CF algorithm with new improved similarity measure which work with ratings given by users to their purchased products is imposed on the clusters which gives efficient solution to cold start problem. Resultant predicted products from CF will undergo feature extraction and review analysis using optimized artificial neural network(ANN) ,before a final prediction. The rest of the paper is organized as section 2 gives an idea about the literature survey, section 3 explains the proposed methodology followed by results and discussion in section 4.

## Related work

Literature survey is done in four parts.

### Clustering and Fruit Fly

The FF optimization algorithm is found in 2011. It is first used for clustering optimization by ( Xing, and g Meng, 2015) in 2015. The algorithm is easy for computation with less number of input parameters than all available swarm optimization algorithms. Authors have proposed shock factor for smell concentration value and tested initial range position on 10 datasets. The authors confirmed high precision, F-measure, and Dunn’s index. Author ( Zhou,Han, 2017) proposed a clustering algorithm based on FF concept. FF algorithm is used to calculate cut-off distance and cluster centers. Authors shown that FF optimization converges fast and helps in

correct clustering on seven UCI repository datasets.

Authors **(Zeng, 2017)** used multi-kernel fuzzy clustering for Multiview data efficiently. Authors used collaborative learning for individual views and multi kernels for the combination of all views in common kernel space. The authors tested the algorithm on synthesis datasets and proved that MKFCM obtained stable results than other existing clustering algorithms.

* 1. **D ynamic clustering**

Authors (Jianrui Chen, Uliji , Hua Wang, Zaizai Yan,2018) used heterogeneous clustering and CF. Users and their favorite products are clustered together and then CF is applied for giving recommendations to the active user. The author experimented on movielense and CiaoDVD dataset and obtained good performance of CF. Authors (Liji U, Yahui Chai, Jianrui Chenb,2018) used dynamic clustering for clustering users as per their states which are calculated from user features. Authors showed user states are changing as time evolves and become stable at some certain values. Authors have clustered users using stable states. This dynamic clustering on movilense dataset results in high precision and recall.

### Collaborative Filtering and Similarity Measures

(Wu, Xiaokun.,2017) has given a new similarity measure based on preferences uniformity of two users. The author has shown the proposed method is efficient than existing similarity measures. (Liu, 2014) also proposed a new user similarity model called new heuristic similarity measure(NHSM) to improve accuracy.Authors make use of normalization to differentiate all users by searching minute differences in their similarities. J.Bobaldia (Bobadilla, Jesu´s.,2012) calculated singularities for every product. He has considered agreement, and disagreements between users. Here the author has shown improvements in the recommendation quality.

### Textual Review Analysis

On the literature survey, it is found that RSs based on ratings can not test the quality of products. Textual reviews are needed to confirm the quality of products. Sentiment-aware recommendation considers reviews given by customers to assess the quality of products. A survey is done on sentiment aware recommendations.

Authors (Vladimir Despotovic, Dejan Tanikic,2017) have used an artificial neural network with a sigmoid activation function to analyze microblogs.Authors compared results with na¨ıve Bayes, Maximum entropy, and SVM techniques and obtained good performance. Authors (Vilares et al., 2017) have used syntax-based rules for semantic calculation and review analysis.They created dependency tree using compositional operations and semantic rules. They provided future scope to extract semantic rules between different objects.

The literature survey concludes that existing clustering methods have given low cluster accuracy and

overlapping clusters. Clustering accuracy is dependent on initial centroids. If initial centroids are not accurate, million of users end up in the wrong clusters, therefore there is a need for optimized clustering. The clusters should be accurate to find the most similar users. Traditional similarity measures of CF face the problem of data sparsity and cold start so predicts irrelevant product list and lost user’s faith in the system. Only a few papers worked on sentiment aware recommendation. Before purchasing the product every user wants quality confirmation of interested product, so review analysis is needed and very important before recommending the products.

## Proposed Method

The proposed system of RS with CF needs user information and user’s rating for the past purchased product as input. For accurate clustering, RS uses an improved objective function of multi-kernel fuzzy clustering(MKFCM).MKFCM gives accurate clustering because of the correct and best initial centroids for clusters. Best centroids are calculated using the proposed fitness function of the FF optimization algorithm. Once accurate clusters are calculated, the sparse utility matrix problem of CF is solved by the evolutionary heterogeneous clustering of users and products.Heterogeneous clustering creates a filled utility matrix of users and their favorite product’s ratings. The new improved similarity measure is used to give recommendations to the single product purchased user and new users. Product quality analysis before the recommending the product is very important. This RS fetches textual reviews from online search engines and analyses using the proposed optimal artificial neural network(ANN). The RS gives three main methods - Optimized Multiple kernel Fuzzy C-means clustering(MKFCM) for clustering users, CF, and Online Review analysis using optimal ANN. In this RS, we proposed an improved objective function of MKFCM with the fitness function of FF optimization, new similarity measure for CF, and textual review analysis using activation function of optimal ANN. Proposed methods are explained in the following section. For explanation, the paper denotes the number of users by X and the number of products by P.

### 3.1 Optimised Multiple kernel Fuzzy C-means clustering(MKFCM for clustering users

Clustering is an important step in the RS. Different clustering methods are used by researchers, among them FCM is better with the expense of a large number of iterations. Clustering should be accurate and results are expected in a short time for millions of users.

Kernel methods can be used to get accurate clustering with less number of iterations. Kernel methods have the advantage of mapping data by easy separations into the

correct structure. The sigmoid kernel has the additional advantage of good performance. The RS uses MKFCM which is an extension to Fuzzy c means which is a combination of linear and sigmoid kernels. The objective function of traditional FCM is given in eqn.1

-----(1)

where,U represents user and C represents cluster center,users,y- Number of clusters and -Cluster membership of user i to cluster j. In traditional FCM if we elaborate on the termK(,)+K(,)-2K(,) (2)

If we put K(,Eqn (1) can be written as Eqn.3 and Where Mk=Avg(Linear, sigmoid).

Traditional FCM is not having the advantage of kernels and result in low cluster accuracy. To gain higher cluster accuracy, improved objective function with an average of sigmoid and linear kernels are taken as shown in Eqn 3. MKFCM proves more efficient than all other clustering algorithms. Centroid updation is done to reach accurate clusters with less iterations.

Million of users make purchases worldwide at all times. Clustering groups together similar users so that a large database of users can be handled efficiently. In this paper, we have analyzed different clustering algorithms and found improved clustering accuracy by MKFCM. It is found that the calculated average of linear and sigmoid kernel gives better accuracy. Also, MKFCM is protected from unrelated features added by different kernels and automatically adjust the kernel weights. For any clustering algorithm, centroids are needed to be chosen correctly. Many algorithms randomly generate these centroids initially. Cluster accuracy is highly dependent on these initial centroids. Therefore this RS uses the FF optimization algorithm to get initial and best centroids.

## FF optimization

Many algorithms of swarm intelligence cluster analysis are available where the proposed optimized FF is simple among them with less number of input parameters. To get the correct initial and best cluster centroids, the FF optimization algorithm is used which gives better cluster accuracy. The Fruitfly algorithm is implemented in the way flies search for their food. These fruit flies are intelligent than other birds as they have a good sense of smell and perception quality. The fruit flies fastly reach to the food and find similar food to the food they have reached efficiently. This behavior is used to find the best cluster centers and accurate clusters having similar users in this research. The FF optimization algorithm is proposed with an improved fitness function to get higher cluster accuracy. Using FF optimization we get different solutions out of which one best solution is selected which gives the best cluster accuracy. This solution gives correct initial and best centroids, using that optimum clusters are created. FF optimization is done by the following Eqn 5.

The RS uses distance within-cluster(DWC)and distance between clusters(DBC) for calculating cluster

fitness. From each FF, one solution is given. In the solution, if DWC is minimum and DBC is maximum, then clusters are accurate. In this way, each FF gives different clusters with different cluster centers. For different cluster centers ,given by each FF, aggregation of DBC is done.

Fitness =———(5) where DWC=———(6) DBC=—————–(7)

-scaling factor=1

According to minimum DWC and maximum DBC, every FF solutions are ranked. The solution with maximum DBC and minimum DWC is considered as the final solution. Final and best centroids will be taken from FF which has searched optimal clusters. These cluster centers will work as initial cluster centroids for Multiple fuzzy c means clustering.

## Evolutionary Heterogeneous Clustering

MKFCM clustering groups together similar users according to user details. The user purchases different products frequently. These products should be considered to calculate similar users. Also at different states, user’s tastes might change and dynamic clustering identifies these changes. The proposed system uses dynamic clustering. The dynamic clustering called as evolutionary heterogeneous clustering creates subclusters of users and products under each homogeneous cluster of users. The idea is that the user and only his most favorite and certified products are combined using an utility matrix as per the kuramoto model. Utility matrix has ratings given by users to their purchased products. This subclustering reduces million of user-products to a small subset of users and their most favorite products. This gives ignore sparse values and concentrates on the target user and his similar users. Thus for recommendation CF can concentrate the exact set of users and products.

### Finding most similar users using CF

In this research, we have analyzed different traditional similarity measures and found the following problems

* + 1. Cosine and adjusted cosine does not take into consideration only ratings given to common products between users, but it considers all ratings given by user to all products.
    2. Jaccard considers common products between users and gives only 2 values of similarity 1 and 0.5
    3. If two users are have commonly purchased products and both have given the same ratings to these products then also Pearson, cosine, and adjusted cosine give different similarity values of these users with the third user.
    4. For new users and sparse data, these similarity measures failed to give a solution.
    5. Tested F- measure accuracy for these similarity measures which is not satisfying.

To overcome all the above problems proposed RS finds out the most similar users for every target user.It calculate the similarity between two different users using the new measure given in eqn.8. In Eqn .8,numerator gives deviation of maximum rating among both the users for the common products with the span of the mean of ratings of both the users. Denominator considers only the maximum rating for the common product between user pair, because of that users with less product purchase history and new users get the most similar users.

where u-two different users,=uu,-uu‘s rating to- uu.products is predicted for the target

### Quality analysis of product using textual reviews extracted from online search engines

Textual review analysis is required to recommend quality confirmed products. The proposed method is used to give quality recommendations to the target user. This method is a dual confirmation to give quality recommendations. To recommend a list of products to target users CF is used with ratings of most similar users, but reviews given by all other users need to be analyzed to decide the quality and ranking of recommended products. This proposed method extracts all online reviews from all search engines for every product predicted in CF for the target user. Here, for active user’s interesting products, we have checked the quality of products by analyzing positive and negative reviews given by all users. The optimized ANN algorithm is used to perform textual review analysis.Then based on positive and negative review scores, predicted products in the list according to quality are recommended to the user.

## Algorithm

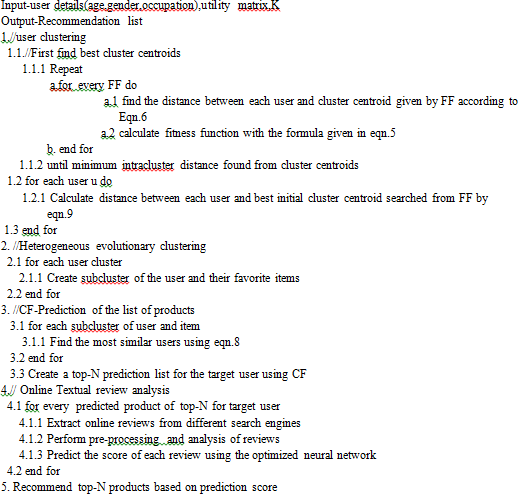
The proposed RS works in five main steps. Million of users and a large number of items are available online, user clustering helps in an optimized way of analyzing users. Heterogeneous clustering evolves as the liking of user changes or if users purchase new items. CF finds similar taste users together and predicts a selection list for users. Textual review analysis arranges the selected list according to best reviews. Finally, the top-N recommendation list is predicted for the user after dual testing.The algorithm for the proposed system is mentioned in Fig.1.

The five steps of algorithm are elaborated further. Step 1:User Clustering

Users are clustered together using user details such as occupation ,gender and age number as input. Age and gender are taken as numeric, while the occupation is mapped among 0 to n. The distance of input data is calculated from the centroid of clusters using Eqn. 9.

Distance(user,centroid)=——–(9)

**Figure 1** Algorithm for optimum personalized confirmed quality aware recommendation



Initial and best centroid(ca,cg, cc) are for the occupation ,gender and age number attributes of the user. and searched using FF optimization algorithm explained in section a of 3.1.Then optimized MKFCM is applied to get accurate clusters.

Step 2:Hetrogeneous Evolutionary Clustering

In this step subclusters of each user, cluster are formed. User and their favorite products are grouped using heterogeneous evolutionary clustering.

Step 3: Prediction of the list of products

In every heterogeneous cluster of users and products, the similarity between two different users is found using the proposed formula given in Eqn.9

Using the highest similarity, most similar people are found for the target user. A list of products is predicted for target users from their most similar users.

Step 4: Online Textual Review Analysis

For every product of top-N ,textual reviews are searched and extracted from online search engines. Preprocessing of these reviews are done. The predicted score of each product is calculated using the optimized ANN algorithm.

Step5:Final Quality Recommendation

All the products in the predicted list of target users are ranked using a predicted score based on reviews. Finally, quality recommendations are given to the target user.

## Results and Discussion

The dataset of MovieLens (https://grouplens.org/dat asets/movielens/100k/) is used in our experiments.It contains 1 lac ratings given by 943 users to 1682 movies. Train and test dataset is given in standard dataset as base and test set. Here for every test user,80 % of his purchased products are taken into the training dataset, and the remaining 20% are taken into the test set for the same users. Like this movilense dataset is divided into five different combinations of the train and test data from u1 to u5 base and test sers. All five test sets contain different users. To evaluate the performance of the proposed method we considered all five base and test sets as five-fold cross-validation and took the average.

**Figure 2**



**Figure 3**



### Evaluation Metrics

Before suggesting list of products to users, the RS predicts ratings for recommended products of the target user. From recommended list of products, if a user actually purchases the product and the predicted rating of a that product matches the real rating given by user, then prediction and recommendation are accurate. In the standard dataset, the test set gives the user’s favorite products. So the RS is accurate if the predicted list of products is having maximum products from the test set of the user. On this concept first n recommended products are classified as relevant and not relevant. For such binary classification precision (Cremonesi et al., 2010) and recall (Fouss et al., 2007) metrics are used which will analyze actual purchased and predicted products by RS. Precision is validating how many relevant products are recommended in first N products and given by formula;

The recall is validating in first N products, how many relevant products are recommended out of the total number of relevant products. Recall can be stated as

Both measures should attain high value for good performance. One more evaluation criteria called F-measure is used which gives an accuracy of recommendation.

F-measure=2\*(P\*R)/(P+R)

To have better recommendation accuracy, F-measure should be high.

### Experimentation

The increase in the number of recommended products and the number of nearest neighbors, affect the precision and recall metrics. In this paper, we have analysed the efficiency of proposed algorithm with the existing algorithms.

### Performance Analysis of MKFCM

The proposed MKFCM algorithm of user clustering for top-N recommended products is compared with two existing algorithms Evolutionary heterogeneous clustering-CF[Chen, Jianrui,2018) and PR-ACMF(Liji U, Yahui Chai, Jianrui Chenb ,2018) . All these

algorithms attained good performance. but are very much calculative and complicated. Also we have compared proposed algorithm with traditional k-means clustering having 3,8 and 15 clusters. The precision

,recall and F-measure for mentioned algorithms are given in table 1.From the table1 it can be seen that proposed MKFCM algorithm outperforms than all existing algorithms. For Top-4 and Top-6 products K-means3 and PR-ACMF have same recall, but our proposed algorithm perform better than both. For all clustering algorithms including our proposed algorithm precision goes on decreasing as Top-N products increased but recall and F-measure increase as top-N increases.

### Performance analysis of proposed similarity measure

RS uses new similarity measure to find the similarity between two different users. The standard dataset is having sparse data. To overcome the sparsity of the dataset we proposed evolutionary heterogeneous clustering which creates subclusters of every user cluster. As subclusters are less sparse than the overall dataset ,new proposed similarity measures outperform traditional similarity measures as shown in table

2. For all traditional similarity measures including our proposed similarity measure, precision goes on decreasing as the number of recommended products are increased. Recall and F-measure is increased as top –N products increases. For top-10 products new similarity measure performs very well. As the number of top-N products is increased, our similarity measure performed very efficiently than other measures. Using the proposed similarity measure, users with single or less product purchase history are properly getting most similar users and hence recommendations as compared to existing measures.

### Performance analysis of optimized ANN

For quality recommendations, we have analyzed online reviews of all users. In this paper, we have compared the proposed method with traditional ANN for performance analysis. Review analysis works on real data extracted from online search engines. New reviews are extracted

**Table 1** Comparision of MKFCM with existing algorithms

Evaluation Metrices

Precision Recall F-measure

Method Top-N Top-N Top-N

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2 | 4 | 6 | 8 | 10 | 2 | 4 | 6 | 8 | 10 | 2 | 4 | 6 | 8 | 10 |
| kmeans3 | 92.71 | 91.74 | 90.94 | 90.4 | 90.03 | 8.48 | 16.2 | 23.03 | 29.07 | 34.47 | 15.47 | 27.36 | 36.48 | 43.68 | 49.51 |
| kmeans8 | 92.18 | 91.23 | 90.49 | 90 | 89.66 | 8.43 | 16.1 | 22.91 | 28.94 | 34.32 | 15.37 | 27.2 | 36.31 | 43.48 | 49.3 |
| kmeans15 | 91.58 | 90.92 | 90.37 | 89.9 | 89.57 | 8.38 | 16.06 | 22.88 | 28.92 | 34.29 | 15.28 | 27.12 | 36.26 | 43.44 | 49.25 |
| EHC-CF | 92.28 | 91.43 | 90.65 | 90.1 | 89.65 | 8.43 | 16.14 | 22.95 | 29 | 34.32 | 15.39 | 27.26 | 39.37 | 43.56 | 49.3 |
| PRACMF | 93.21 | 91.78 | 91.11 | 90.4 | 90.1 | 8.52 | 16.2 | 23.03 | 29.08 | 34.5 | 15.55 | 27.36 | 36.55 | 43.69 | 49.61 |
| PROP | 93.68 | 92.02 | 91.44 | 90.9 | 90.22 | 9.03 | 16.52 | 23.62 | 29.33 | 34.62 | 16.48 | 28.02 | 37.54 | 44.35 | 50.04 |

**Table 2** comparison of proposed similarity measure with existing measures

Evaluation Metrices

Precision Recall F-measure

Method Top-N Top-N Top-N

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 12 | 14 | 16 | 18 | 10 | 12 | 14 | 16 | 18 | 10 | 12 | 14 | 16 | 18 |
| cos | 89.91 | 89.37 | 89.19 | 88.8 | 88.6 | 34.4 | 39.09 | 43.39 | 47.19 | 50.7 | 49.44 | 54.04 | 58.02 | 61.27 | 64.14 |
| pcc | 89.06 | 88.7 | 88.44 | 88.3 | 87.99 | 34.1 | 38.8 | 43.04 | 46.89 | 50.35 | 48.97 | 53.65 | 57.54 | 60.88 | 63.69 |
| acos | 89.28 | 88.89 | 88.63 | 88.3 | 88.15 | 38.9 | 38.88 | 43.12 | 46.94 | 50.44 | 49.09 | 53.74 | 57.65 | 60.95 | 63.81 |
| icos | 89.95 | 89.39 | 89.2 | 88.9 | 88.65 | 34.4 | 39.16 | 43.42 | 47.23 | 50.72 | 49.5 | 54.15 | 58.41 | 61.68 | 64.52 |
| prop | 90.23 | 90.06 | 89.65 | 89.1 | 89.31 | 39.1 | 39.92 | 44.12 | 47.96 | 51.27 | 54.57 | 55.32 | 59.14 | 62.35 | 65.14 |

**Table 3** Comparision of traditional ANN and optimized ANN o textual reviews

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 10 | 12 | 14 | 16 | 18 |
| ANN | 70.3 | 71.21 | 71.51 | 71.7 | 72.51 |
| Optimized ANN | 71.36 | 72.3 | 72.45 | 73 | 73.52 |

Evaluation Metric F-measure Method Top-N

for all target users. This real-time review analysis is compared by using traditional and proposed optimized ANN. The optimized ANN method works efficiently than traditional ANN as shown in table 3. The proposed optimized method works efficiently when reviews are considered for top-10 and more products.

### Cold start

RSs suffer from a cold start problem. RSs fail to give quality recommendations to a new user as he has no history. The proposed system solves this cold start issue. A new user is allocated to a respective user cluster according to his age, gender, and occupation. Then this user goes to evolutionary heterogeneous clustering and CF. For this, we considered the default highest rating for all products which are purchased by other users in the same cluster having the same age, gender, and occupation combination. Then target new users get quality and appropriate recommendations. This solution acquires user trust in the RS.

6. **Conclusion**

Clustering is important to improve the scalability of the CF algorithm. It reduces million of users into different small clusters. These clusters can be handled efficiently than entire user data in this RS using MKFCM. Other than available methods, this RS used a dynamic heterogeneous clustering to deal with sparse

values in the data. The subclusters are evolved as a user makes a new purchase or if the user’s taste changes with time. CF works around the new similarity measure. The proposed new similarity measure used ratings of common products purchased between different users which should be considered actually to calculate similarity. Our RS is more efficient with high recall and recommendation accuracy as compared to k-means,EHC-CF,PRACMF. Review analysis is important for quality confirmation before final recommendation. This RS extracts real- time textual reviews of products from online search engines to analyze quality. Negative reviews confirm the low-quality features of products,such products may not be recommended to the target user. The analysis shows that the optimized ANN gives an accurate review classification compared to traditional ANN. An optimum personalized RS is provided in this work which gives dual confirmation to target users about recommended products first by CF and second by review analysis.

## References

Bobadilla, J. (2013).

Cremonesi, P., Koren, Y., and Turrin, R. (2010). ‘Performance of recommender algorithms on top-n

recommendation tasks’. *Proceedings of the fourth ACM conference on Recommender systems*.

Fouss, F., Pirotte, A., michel Renders, J., and Saerens, M. (2007). ‘Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation’. *IEEE Transactions on Knowledge and Data Engineering*, Vol 19, No 3, pp. 355–369.

Liu, H. (2014). ‘A new user similarity model to improve the accuracy of CF’. *Knowledge-Based Systems*, Vol 56, pp. 156–166.

Pazzani, M. J. and Billsus, D. (2007). *Content-based RSs*. Springer, Berlin, Heidelberg.

Vilares, D., G´omez-Rodr´ıguez, C., and Alonso, M. A. (2017). ‘Universal, unsupervised (rule-based), uncovered sentiment analysis’.

Zeng, S. (2017). ‘A unified collaborative multikernel fuzzy clustering for multiview data’. *IEEE Transactions on Fuzzy Systems*, Vol 26, pp. 1671–1687.