Journal of Theoretical and Applied Information Technology
31st January 2023. Vol.101. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DIVERSITY ENHANCED ADAPTIVE CORRELATION CONNECTED CLUSTERING OF LONG TAIL ITEMS

SOANPET SREE LAKSHMI[1] , T. ADILAKSHMI[2] , BAKSHI ABHINITH[3]

[1]Assistant Professor, Department of Information Technology, Vasavi College of Engineering, India

[2]Professor, Department of Computer Science &Engineering, Vasavi College of Engineering, India.

[3]Student, Department of Computer Science &Engineering, Vasavi College of Engineering, India.

E-mail: [1]s.sreelakshmi@staff.vce.ac.in, [2] t_adilakshmi@staff.vce.ac.in , [3]bakshiabhinith@gmail.com

ABSTRACT

Recommender systems play a very important role in driving businesses. They recommend a set of items to the user which have a higher chance of getting consumed. The primary issue addressed in this work is to generate a recommendation list with items belonging to various categories so that the user can explore his different interests. The proposed method studies the diversity of the recommended list by enhancing the adaptive clustering method. In this approach, the dataset is partitioned into 3 sets namely the head part, the mid part, and the tail part . Then, different sets of methods are used to improve the diversity of the recommendation list. Popular items are extracted from the head part, items in the mid part are extracted using rating-based clustering method and the items in the tail part are extracted using correlation clustering-based method, thereby improving the diversity of the recommendation list.

Keywords: *Recommender Systems, Diversity, Clustering, Adaptive, Long Tail.*

1. INTRODUCTION

The goal of the RS is to generate a list of recommendations to be provided to the user . The list should be such that it increases the satisfaction of the user . Many approaches have been investigated in the literature . Some of the approaches include personalization or customization of the list according to the users' preferences, past history of the user, to name a few. Personalization is different from being context aware .Other approaches focus the inclusion of different items in the list so that the user does not get similar recommendations . The list may also contain other items which are not totally similar to the ones that the user likes and at the same time not totally dissimilar ones . In the proposed work, the recommendation list is generated applying different methods to different sets of the given data , thus enhancing the diversity in recommendation.

2. DIVERSITY OF RECOMMENDATIONS

Authors in [13,42,43,48] studied the accuracy and diversity and how they vary in the list. The authors discuss that focusing only on accuracy would lead to long term failure in case of usability of recommendation. The studies indicate that it is exceedingly difficult to improve diversity while maintaining accuracy in the recommendation.

The recommendation diversity can either be assessed at an individual level or an aggregate level. Most of the previous studies [44,45,46] focus on individual diversity , based on the average dissimilarity between items present in the Recommendation lists. Aggregate diversity is measured in [13,42,47] . Authors in [43] proposed a graph-based approach to obtain diversity. Their approach is limited to the data as a whole and there exists no specific approach to address the items in the long tail.

Authors in [63] proposed combining of genre coverage and the genres not repeating in the list. Such a method includes items from dissimilar genres, including human behavior in its calculation . authors in [64] include the items in the list, in such a way that those items that were successful in the past but forgotten.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. PROPOSED IDEA : DIVERSITY ENHANCEMENT THROUGH ENHANCED ADAPTIVE CLUSTERING

In the proposed approach, the recommendation list is generated by applying different methods to different set of items of varying popularity. Understanding that humans tend to like few genres and would be willing to accept and explore some newness in the recommendations, but not completely new and unexpected recommendations . Therefore, the proposed approach enlists the items in such a way that they are fairly expected with some newness. The split points can be identified by the possible existing elbows .

Table 1: Sample Simpson Diversity Index Calculation For A User:

Class (genre)	Value-(n_i)	n_i (n_i - 1))
C1	16	240(16*(16-1))
C2	2	1
C3	1	0
C4	1	0
C5	7	42
	N = $\sum n_i$ = 27	$\Sigma(n_i(n_i - 1))$ = 283
(N* (N-1))	= 27*(27-1) = 702	
(Σ ($n_i(n_i$ - 1))) / (N(N - 1))	= 283/702 = 0.403	
Simpsons-diversity Index	= 1 - 0.403=0.596	

To diversify the items in the recommendation list, items in the recommendation list are given from all the three parts, the head, mid and tail parts. Enhanced adaptive clustering method is applied for long tail items. The recommended list contains popular items from head part, items from the mid part as well as the niche items from tail part. The items in the head part are recommended by their popularity measure- number of ratings .

The Mid–items are recommended, clustered based on the ratings-values. The items in long tail are less popular and therefore are difficult to handle in RS. Correlation connected clustering is employed for clustering movies in long tail based on their similarities. The recommendation list contains diverse items. Popular from head part, similar from Mid part and niche items can be recommended from the tail part.

This approach is implemented on a subset of the movielens dataset and the accuracy of rating predictions is evaluated.

4. PROPOSED ALGORITHM: ENHANCED ADAPTIVE CORRELATION BASED CONNECTED CLUSTERING - DIVERSITY (EADCCC-D)

The Proposed Algorithm EADCCC-D splits the items into 3 parts based on number of ratings. This work enhances the adaptivity of the algorithm to include diverse items from the different parts into the recommendation list. For the data set chosen, the items are movies. The movies in the Head Part are called Popular Movies (PM) , Mid part are Rating based Movies clusters (RC) and the tail part are Correlation Connected based Movies clusters (CCC). The Fig 1 below illustrates the separation of movies into three regions. The choice of the splits uses the elbow method . The fall point is termed α-min, and the lapse Point is α-max.

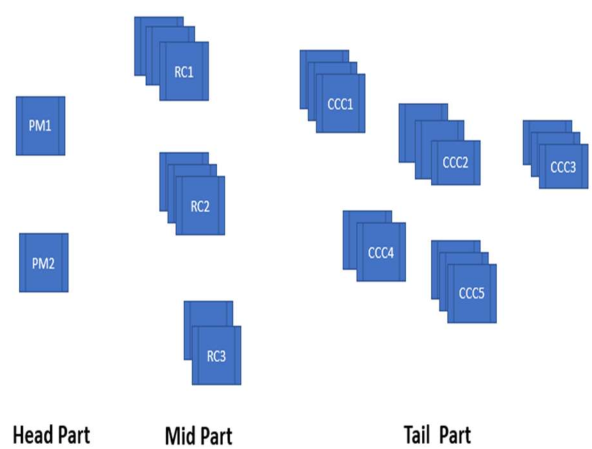

Figure 1: Showing Separation Of Items Into Three Parts

Algorithm: Enhanced_Diversity _Recommendations:

Step1: Derived variables that are based on users and movies are calculated-

Journal of Theoretical and Applied Information Technology
31st January 2023. Vol.101. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

User_favourite_genre,User_favourite_cluster,Movie_average_rating ,Movieid_count, Movieid_genre

Step2: Calculate the User_favourite_genre, User_favourite_cluster

Step3: Based on movieid_count, the data is split into three parts: head, mid and tail.
The criterion that is used to split the data is α-min and α-max. It is calculated by plotting the sorted no of ratings . The split points can be identified by the possible existing elbows.

Step4: Movieid_clusters from both mid and tail parts are generated.

Step5: A recommendation list is generated. the list consists of three sublists.

a) Sublist_head consists of movies present in the head part of data whose genre is the User_favourite_genre

b) sublist_mid consists of movies present in User_favourite_cluster (ratings - based)

c) sublist_tail consists of movies present in User_favourite_cluster-(tags based)

Step6: sublist_head, sublist_mid and sublist_tail is merged to give a final list.

The proposed method will generate the recommendation list with diverse items from the **A) Head part B) Mid part C) Tail part**

A) Head part Recommendation:

Popular movies are viewed by many , hence owing to those movies, head part recommendations are made . The movies in the head part are sorted in the descending order of ratings. The User_favourite_genre is found from the user_movie vector and the top movies in the head part that are of the same genre as that of favourite_genre is used to generate this sublist.

B) Mid part Recommendation:

The movies are clustered based on Average_ratings . Also, they are assigned clusterid' s. A user id is given as an input. The User_favourite_genre is extracted using the users' rating patterns. User_favourite_cluster is identified from the ratings-cluster and the User_favourite_genre .For the given userId , the movies from the User_fav_cluster are extracted. This cluster is used to generate the sublist_r.

C) Tail part Recommendation:

Correlation connected clustering is employed in the long tail to cluster movies with respect to their tag-genome scores. A user id is given as input, The User_favourite_genre is extracted using the users' rating patterns .The ratings values (>3) are taken into consideration while finding the favourite genre of the user. The maximally occurring genre in this list is chosen as User_favourite_genre every userId, the users favourite cluster is found and its clusterid is extracted. The movies in this cluster are used to generate the sublist_t.

The Flow Chart in Fig 2 explains the overview of the steps used in EADCCC-D method .

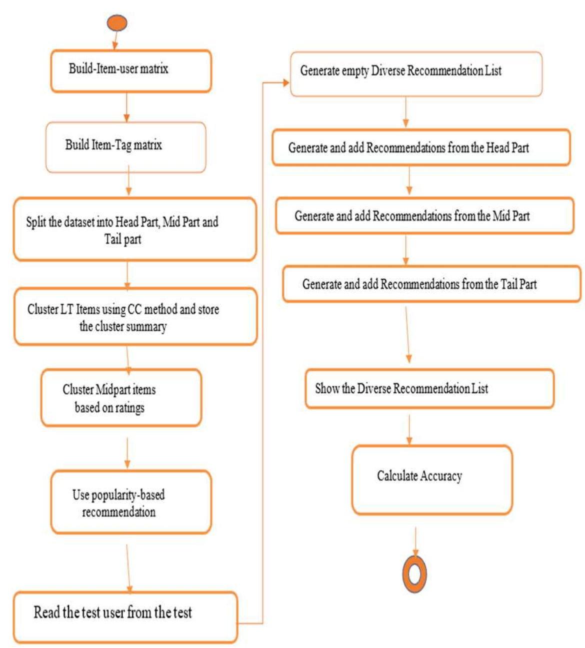

Figure 2 : Flow Chart Depicting The Steps Used In EADCCC-D

Journal of Theoretical and Applied Information Technology
31st January 2023. Vol.101. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. DATA SET DESCRIPTION

Every user in the dataset had watched and rated at least 20 movies. Demographic information of users is not given. A userId uniquely identifies a user. The data are contained in six files. The dataset can be downloaded from http://grouplens.org/datasets/.
The files listed below in table 2 describe the attributes available in the data set

Table 2: Dataset Description

FILE	ATTRIBUTES	Description
ratings.csv	userId , movieId , rating , timestamp	The ratings given by a User for the movie along with a timestamp
genome-tags.csv	tagId ,tag	Each Movie described in 1128 tags
genome-scores.csv	movieId ,tagId ,relevance	Relevance of each tag for every movie
movies.csv	movieId ,title , genres	Title of the movie along with the possible genres
tags.csv	userId , movieId , tag , timestamp	The tag justified by a User for the movie along with a timestamp

From this dataset , 100 users data were considered for experimentation . All movies rated by these 100 users are extracted. 10 is the number of ratings that was chosen as a minimum threshold for a movie to be considered. The count of Movies satisfying this criterion is 243.The criterion alpha for splitting data was calculated using possible elbows. The Elbow method detect points with maximum curvature .

6. RESULTS

Enhanced Adaptive correlation based Connected clustering- Long tail (EADCCC-LT)[13] is the approach in which the item set is divided into 2 parts and the correlation-based clustering is applied in the tail part. EADCCC-D is the extended approach in which the item set is divided into 3 parts , where popularity method is applied in head part, ratings-based approach is applied in the mid part and the correlation-based clustering is applied in the tail part. The recommendation list are generated for the sample users is shown in Fig 4 and 5. The Fig 4 shows the for recommendation list generated for sample userId 77.Fig 5 shows the for recommendation list generated for sample userId 92. Fig 6 shows the accuracy of the recommendation list for test users in EADCCC-LT. Fig 7 shows the accuracy of the recommendation list for test users in EADCCC-D. For example, If the recommendation list contains all the movies-watched by a sample user, the accuracy will be 100%. Similarly, If the recommendation list contains 5 movies that have been watched by a sample user out of 10 recommended movies, the accuracy would have been 50%. Thus, the accuracy is calculated for both the methods for the test users.

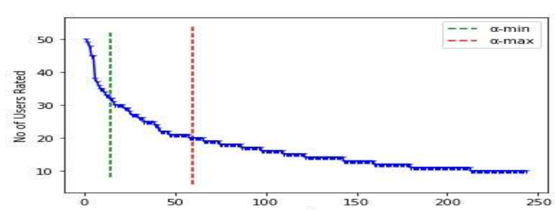

Figure 3: Rating Distribution Of 100 Users- 243 Movies - Split Points

Performance Comparisons

EADCCC-LT is the method in which the data is split into 2 parts as discussed in[26] , in comparison with the proposed approach where a 3-way split is proposed.

Journal of Theoretical and Applied Information Technology
31st January 2023. Vol.101. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Sample Accuracy Calculation For5 Users

UserId	User_ Wat	Watched_ recommended	Accur acy	Sample Average Accuracy
1	15	11	0.73	0.812
2	19	17	0.89	
3	33	21	0.63	
4	16	14	0.87	
5	18	17	0.94	

Table No.3 shows the sample accuracy calculation of the recommendation list for sample users.

The experiment is tested on 100 users dataset and the calculated accuracy for test users is shown in the accuracy column of Table No.4. It shows that EADCCC-LT is slightly higher than that of the using EADCCC-D method.

Figure 4: Output_Screen_Snip For Recommendation List Userid 77

Figure 5: Output _Screen_Snip For Recommendation List: Userid 92

Table 4: Performance EADCCC-D vs EADCCC-LT

Method	Accuracy	Simpson_ diversity_ index	Gini-ind
EADCCC-LT	0.834	0.558	0.551
EADCCC-D	0.796	0.604	0.581

Table No.4 compares the Accuracy and Diversity in recommendation list for test users using EADCCC-LT and using EADCCC-D .

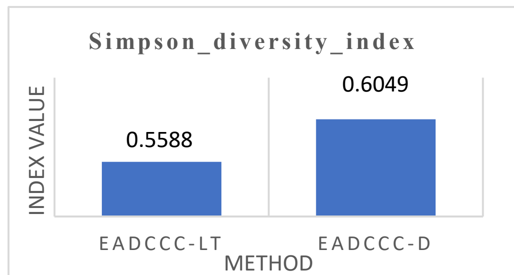

Figure 6: Accuracy Comparison Between EADCCC-LT and EADCCC-D

Fig 6 shows the accuracy of the recommendation list for sample test users using EADCCC-LT and also using EADCCC-D.

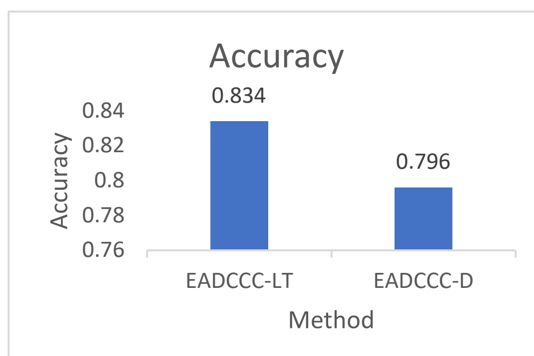

Figure 7: Simpsons Diversity Index Values Comparison Between EADCCC-LT and EADCCC-D

Fig 7 shows that the Simpsons diversity index of the method proposed EADCCC-D is higher than that of the EADCCC-LT method.

Fig 8 shows that the Gini index of the method proposed EADCCC-D is also higher than that of the EADCCC-LT method.

The accuracy of the EADCCC-D is less due to the increased diversity of the recommendation list. This helps the user in exploring different genres of movies that would generally have been suppressed by the recommendation list generated by the EADCCC-LT method.

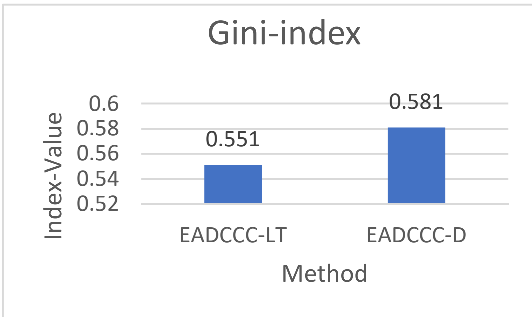

Figure 8: Gini Index Values Comparison Between EADCCC-LT and EADCCC-D

The Simpson diversity index score always varies between 0 and 1. A high score indicates a higher diversity, and a low score indicates a lower diversity. The comparison indicates that the diversity of the recommendation list generated by the EADCCC-D method is improved at the cost of slightly reduced accuracy.

7. SUMMARY

The diversity aspect of recommendations is an important consideration for the RS along with accuracy. The diversity of recommendations in the proposed work is improved by three-way split of the dataset that is based on the elbow points of the curve . Separate models for clustering were built for those parts and the recommendation lists are generated accordingly. These lists are then combined to make one single list for every user.

The diversity enhancement of the recommendation list is achieved by including items with a large number of ratings from the head part , items with similar ratings from the mid part , and items from long tail . The proposed methods were used, and the list would contain items, mostly from the same cluster in the long tail, and items with similar ratings to lessen the risk of including totally dissimilar items in the Recommendation List . Various approaches in literature focused on including dissimilar items whereas the proposed approach worked on producing lists with less dissimilar items, to reduce the risk factor.

In this work, we see that the three-way split has indeed increased the diversity within the recommendation list when compared to a two-way split that was used in previous works. This work can further be extended by using a multi-way split technique that divides the dataset into multiple parts and applying different methods of clustering for each part.

REFERENCES:

[1] G.Adomavicius And A.Tuzhilin, "Toward The Next Generation Of Recommender Systems: A Survey Of The State -Of-The Art And Possible Extensions, "Ieee Trans . Knowledge And Data Eng., Vol.17, No.6,Pp.734-749,June 2005.

[2] C. Anderson, The Long Tail, Hyperion ,Press,2006

[3] Hongzhi Yin Bin Cui Jing Li Junjie Yao Chen Chen , "Challenging The Long Tail Recommendation", Proceedings Of The Vldb Endowment, Vol. 5, No. 9.

[4] Y.J.Park And A.Tuzhilin, "The Long Tail Of The Recommender Systems And How To Leverage It," Proc.Acm Conf. Recommender Systems, Pp . 11-18, 2008.

[5] Y.J. Park,"The Adaptive Clustering Method For The Long Tail Problem Of Recommender Systems", Ieee Trans . Knowledge And Data Eng., Vol.25, No.8,August 2013.

[6] Http://Files.Grouplens.Org/Datasets/Movielens/Ml-20m-Readme.Html

[7] Paolo Cremonesi Franca Garzotto Roberto Pagano Massimo Quadrana, Recommending Without Short Head , Www'14 Companion, April 7–11, 2014, Seoul, Korea. Acm 978-1-4503-2745-9/14/04. Http://Dx.Doi.Org/10.1145/2567948.2577286

[8] Christian Bohm, Karin Kailing, Peer Kroger, Arthur Zimek," Computing Clusters Of Correlation Connected Objects" In Proc.Acm Int Conf On Management Of Data (Sigmod), Pp 455-466,Paris, France

[9] Soanpet. Sree Lakshmi ,Dr,T.Adi Lakshmi, "The Survey Of Recommender Systems", International Journal Of Engineering Trends And Technology (Ijett) – Special Issue – April 2017.

[10] M. Ramakrishna Murty,J.V.R Murthy, Prasad Reddy P.V.G.D, Suresh.C.Sapathy "Performance Of Teaching Learning Based Optimization Algorithm With Various Teaching Factor Values For Solving

Optimization Problems", International Conference Ficta-13 At Bhuveneswar, Springer Aisc (Indexed By Scopus, Isi Proceeding Dblp Etc),Vol:247,November 2013,Isbn 978-3-319-02931-3,Pp.207-216.

[11] M .Ramakrishna Murty, J.V.R Murthy, Prasad Reddy P.V.G.D, Suresh Satapathy "Homogeneity Separateness: A New Validity Measure For Clustering Problems" International Conference And Published The Proceedings In Aisc And Computing, Springer, (Indexed By Scopus, Isi Proceeding Dblp Etc)Vol 248, 2014,Isbn 978-3-319-03106, Pp 1-10.

[12] Http://Files.Grouplens.Org/Papers/Tag_Genome.Pdf

[13] J.B. Schafer, J.A. Konstan, And J. Riedl, "E-Commerce Recommendation Applications," Data Mining And Knowledge Discovery, Vol. 5, Nos. 1/2, Pp. 115-152, 2001

[14] B. Sarwar, G. Karypis, J. Konstan, And J. Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," Proc. 10th Int'l World Wide Web Conf., 2001.

[15] M. Pazzani And D. Billsus, "Learning And Revising User Profiles: The Identification Of Interesting Web Sites," Machine Learning, Vol. 27, Pp. 313-331, 1997.

[16] R.M. Bell And K. Yehuda, "Improved Neighborhood-Based Collaborative Filtering," Proc. Kdd Cup Workshop, Pp. 7-14, 2007.

[17] K.Q. Truong, F. Ishikawa, And S. Honiden, "Improving Accuracy Of Recommender Systems By Item Clustering," Ieice Trans. Information And Systems, Vol. E90-D, No. 9, Pp. 1363-1373, 2007..

[18] D.M. Fleder And K. Hosanagar, "Blockbuster Cultures Next Rise Or Fall: The Impact Of Recommender Systems On Sales Diversity," Management Science, Vol. 55 No. 5, Pp. 697-712, 2009.

[19] A. Hervas-Drane, "Word Of Mouth And Recommender Systems: A Theory Of The Long Tail," Working Paper, Harvard Business School, 2007.

[20] T. Tran And R. Cohen, "Hybrid Recommender Systems For Electronic Commerce," Knowledge-Based Electronic Markets, Pp. 78- 84, 2000

[21] L.H. Ungar And D.P. Foster, "Clustering Methods For Collaborative Filtering," Proc. Workshop Recommendation Systems, 1998

[22] Gediminas Adomavicius And Youngok Kwon. Maximizing Aggregate Recommendation Diversity: A Graph-Theoretic Approach. (Divers 2011), Held In Conjunction With Acm Recsys 2011. October 23, 2011, Chicago, Illinois, Usa

[23] Z. Szlavik, W.J. Kowalczyk, And M.C. Schut. 2011. Diversity Measurement Of Recommender Systems Under Different User Choice Models (Icwsm'11)

[24] Jacek Wasilewski And Neil Hurley ,"How Diverse Is Your Audience? Exploring Consumer Diversity In Recommender Systems (Recsys'2017)

[25] M. Servajean, E. Pacitti, S. Amer-Yahia, And P. Neveu. 2013. Profile Diversity In Search And Recommendation (Www '13 Companion).

[26] Soanpet .Sree Lakshmi, Dr. T Adi Lakshmi ,Bakshi Abhinith "An Adaptive Correlation Clustering-Based Recommender System For The Long-Tail Items",. Smart Innovation, Systems And Technologies. Springer ,. Volume 224 , 2021, 505-514 . Http://Www.Springer.Com/Series/8767

[27] Y. Shi, An Improved Collaborative Filtering Recommendation Method Based On Timestamp, 16th International Conference On Advanced Communication Technology. Ieee, 2014.

[28] L. Sun, E. I. Michael, S. Wang And Y. Li, "A Time-Sensitive Collaborative Filtering Model In Recommendation Systems," 2016 Ieee International Conference On Internet Of Things (Ithings) And Ieee Green Computing And Communications (Greencom) And Ieee Cyber, Physical And Social Computing (Cpscom) And Ieee Smart Data (Smartdata), Chengdu, 2016, Pp. 340-344

[29] Q. Yuan, G. Cong, Z. Ma Et Al., "Time-Aware Point-Of-Interest Recommendation," In Proceedings Of The International Acm Sigir Conference On Research & Development In Information Retrieval, Acm, Dublin, Ireland, July 2013.

[30] Wei S., Ye N., Zhang Q. (2012) Time-Aware Collaborative Filtering For Recommender Systems. In: Liu Cl., Zhang C., Wang L. (Eds) Pattern Recognition. Ccpr 2012. Communications In Computer And Information Science, Vol 321. Springer, Berlin, Heidelberg.

Https://Doi.Org/10.1007/978-3-642-33506-8_81

[31] Michael C. Thrun ∗ , Quirin Stier ,"Fundamental Clustering Algorithms Suite", Softwarex, Volume 13,2021,100642,Issn 2352-7110, Https://Doi.Org/10.1016/ J.Softx.2020.100642

[32] Dávid Nagy* , Laszló Aszalós, Tamás Mihálydeák," Finding The Representative In A Cluster Using Correlation Clustering" Pollack Periodica An International Journal For Engineering And Information Sciences Doi: 10.1556/606.2019.14.1.2 Vol. 14, No. 1, 2019, Pp. 15–24

[33] M. Pazzani, A Framework For Collaborative, Content-Based And Demographic Filtering, Artificial Intelligence Review, 13 (1999) ,393-408

[34] G. Adomavicius, A. Tuzhilin, Context-Aware Recommender Systems, In: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.) Recommender Systems Handbook, Springer Us2011, Pp. 217-253.

[35] J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative Filtering Recommender Systems, In: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.) The Adaptive Web, Springer Berlin Heidelberg2007, Pp. 291-324.

[36] M. Pazzani, D. Billsus, Content-Based Recommendation Systems, In: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.) The Adaptive Web, Springer Berlin Heidelberg 2007, Pp. 325-341.

[37] Bo Hu And Martin Ester. 2013. Spatial Topic Modelling In Online Social Media For Location Recommendation. In Proceedings Of The 7th Acm Conference On Recommender Systems, Recsys '13. Association For Computing Machinery, New York, Ny, Usa, 25–32. Doi:Https://Doi.Org/10.1145/2507157.2507174

[38] Fernando Mourão, Claudiane Fonseca, Camila Araújo, Wagner Meira Jr," The Oblivion Problem: Exploiting Forgotten Items To Improve Recommendation Diversity". Workshop On Novelty And Diversity In Recommender Systems (Divers 2011), Held In Conjunction With Acm Recsys 2011. October 23, 2011, Chicago, Illinois, Usa

[39] Landin, A.; Suárez-García, E.; Valcarce, D. When Diversity Met Accuracy: A Story Of Recommender Systems. *Proceedings* 2018, *2*, 1178. Https://Doi.Org/ 10.3390/ Proceedings2181178

[40] Younghoon Kim, Kyuseok Shim, Min-Soeng Kim, June Sup Lee, Dbcure-Mr: An Efficient Density-Based Clustering Algorithm For Large Data Using Mapreduce, Information Systems, Volume 42, 2014,Pages 15-35,Issn 0306-4379,Https://Doi.Org/10.1016/J.Is.2013.11.002.

[41] C. Ziegler .S. Mcnee ,J.Konstan, G.Lausen, Improving Recommendation Lists Through Topic Diversification In Proc 14th Int. Conf. Www, Chiba, Japan,2005,Pp22-32

[42] Adomavicius, G., Y. Kwon. 2009. Toward More Diverse Recommendations: Item Re-Ranking Methods For Recommender Systems. Proc. Of The 19th Workshop On Information Technologies And Systems.

[43] Adomavicius, G., Y. Kwon. 2011. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques. Ieee Transactions On Knowledge And Data Engineering. Forthcoming.

[44] Bradley, K., B. Smyth. 2001. Improving Recommendation Diversity. Proc. Of The 12th Irish Conf. On Artif. Intelligence And Cognitive Sci.

[45] Hu, R., P. Pu. 2011. Enhancing Recommendation Diversity With Organization Interfaces. Proc. Of The 16th Int'l Conf. On Intelligent User Interfaces (Iui '11). 347-350.

[46] Mcsherry, D. 2002. Diversity-Conscious Retrieval. Proc. Of The 6th European Conf. On Advances In Case-Based Reasoning 219-233

[47] Herlocker, J.L., J.A. Konstan, L.G. Terveen, J. Riedl. 2004. Evaluating Collaborative Filtering Recommender Systems. Acm Transactions On Information Systems 22:1 5-53.

[48] Mcnee, S.M., J. Riedl, J.A. Konstan. 2006. Being Accurate Is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems. Conf. On Human Factors In Computing Systems 1097-1101

[49] Han And Kamber, Data Mining, Concepts And Techniques 3rd Edition.

[50] Https://Www.Amazon.Science/The-History-Of-Amazons-Recommendation-Algorithm

[51] Amazon.Com Recommendations Item-To-Item Collaborative Filtering, Greg Linden, Brent Smith, And Jeremy York • Amazon.Com. January - February 2003 Published By The Ieee Computer Society 1089-7801/03/$17.00©2003 Ieee Ieee Internet Computing

[52] R.D. Lawrence ,G.S. Almasi , V. Kotlyar ,M.S. Viveros , S.S. Duri , "Personalization Of Supermarket Product Recommendations ", Data

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Mining And Knowledge Discovery, 5, 11–32, 2001 °C 2001 Kluwer Academic Publishers. Manufactured In The Netherlands

[53] Kenta Oku, Shinsuke Nakajima, Jun Miyazaki, Shunsuke Uemura ," Context-Aware Svm For Context-Dependent Information Recommendation" Proceedings Of The 7th International Conference On Mobile Data Management (Mdm'06) 0-7695-2526-1/06 $20.00 © 2006 Ieee

[54] Weiyang Lin, Sergio A. Alvarez ,Carolina Ruiz ,"Efficient Adaptive-Support Association Rule Mining For Recommender Systems ".2002, Kluwer Academic Publishers . Printed In Netherlands.

[55] Diogo Vinícius De Sousa Silva , Frederico Araújo Durão ," A Hybrid Approach To Recommend Long Tail Items" In: Xviii Workshop De Teses De Dissertações (Wtd 2018), Salvador, Brasil. Anais Do Xxiv Simpósio Brasileiro De Sistemas Multimedia E Web: Workshops E Pôsteres. Porto Alegre: Sociedade Brasileira De Computação, 2018. © 2018 Sbc – Sociedade Brasileira De Computação. Isbn 978-85-7669-435-9

[56] Brynjolfsson, E., Y. Hu, And D. Simester. "Goodbye Pareto Principle, Hello Long Tail: The Effect Of Search Costs On The Concentration Of Product Sales." Management Science 57.8 (2011): 1373–1386