Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

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

A NOVEL PRODUCT FILTERING AND PRODUCT RECOMMENDATION SYSTEM TO OPTIMIZE THE SEARCH SPACE AND SPARSITY IN REAL TIME ECOMMERCE ENVIRONMENT

[1]MR.T.SUBHA MASTAN RAO, [2]DR.G.RAMA KRISHNA

[1]Asstt Prof., Department of Computer Science and Engineering, KL University
[2]Professor., Department of Computer Science and Engineering, KL University

E-mail: [1]mastan@kluniversity.in, [2]ramakrishnag_cse@kluniversity.in

ABSTRACT

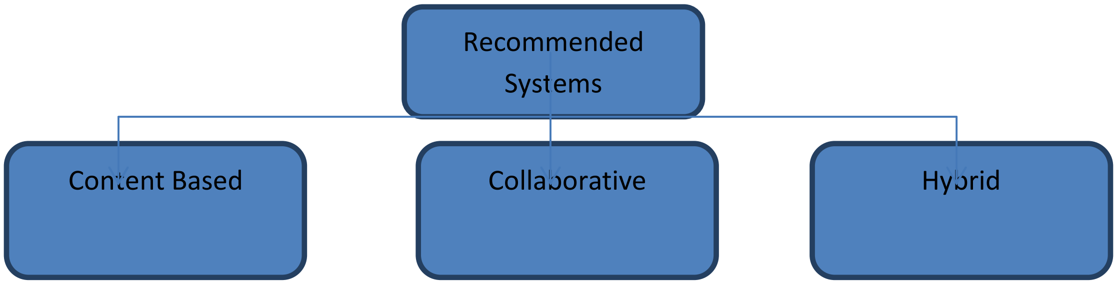
Online Product recommended system is the most effective prediction system in the e-commerce websites. Customized/Automated recommendation systems can assist the users to find relevant products within short time in large e-commerce databases. Several recommendation techniques have been proposed to filter the user interested products and to optimize e-commerce sales from different vendors. With the tremendous increase in the products and customers in e-commerce systems, the time taken to search the required product using traditional recommended techniques increases due to the large number of products features and its specifications. With the tremendous growth in the products and customers information however, the traditional systems encounter two key issues, enhanced response time and minimized recommending quality. New Qualitative recommended systems are essential to handle. The traditional techniques can't implement search space optimization when the item-space changes; on the other hand, the overall efficiency of the recommendation system will decrease as the items grow into a large amount. Sparsity of the product is the most significant basis which minimizes the inadequate quality. In this paper, a new product filtering and product recommended system is proposed to optimize the search space and sparsity. Product filtering can be used to filter the relevant products from the web using the conditional probability and similarity values. Product prediction based recommended system is proposed to predict the products ranking based on the products features as well as calculated predicted probabilities. Experimental results show that proposed filtering based recommended system outperforms well on web applications like e-bay in terms of search space and ranking are concern.
Keywords: *Online, Product Filtering, Qualitative, Sparsity, Product Filtering*

1. INTRODUCTION

In the e-commerce world everything goes on the web and depends on the internet, the purchasers and the vendors never meet one another but transaction deal can be finished through online via internet. Vendors design their webpages to maintain their product's information. Purchasers use the web-navigator to check out the website and they will admit the deal when those items meet their needs. Recommended filtering (RF) is one of the most potentially/promising recommending approaches. The major advantages of RF system are below: it possesses no special restriction on the highly recommended objects; it cannot only suggest text-based information, but additionally movies , music etc, whose data is hard to be assessed[4]; and it may also choose the users' new desires and priorities. But, recommending systems utilizing conventional recommended filtering are encountering two kinds of major issues: one is the processing workload; and the other is the forecasting precision. With the increase in large number of users and the items; the scale of the system becomes larger in size. Consequently, on one side, the processing workload boosts more and more rapidly; on the other side, the database of user ratings seems to be much more sparse. It is ever more difficult to determine the users' similarities precisely, which allows the decreases in the recommending quality. In some ways these two issues are in conflict[1-2].

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

The system that predicts relevant products using robust filtering technique is known as Recommendation System(RS). The main objective of Recommendation System is to perform task to filter relevant products from the large databases. It much relies upon similarity prediction. In [1], Recommendation System(RS) can be categorized into three different types based upon application used; they are content based technique, collaboration technique, and hybrid technique.



Content Based Technique

Content-Based Technique filters products to the user from products favored by that user previously. It is carried out by first creating interaction between product and its features in term of Similarity Matrix then select among the most equivalent products to the target product by computing similarity from different mathematical functions. The most popular similarities like Cosine, Adjusted Cosine and Pearson coefficient. Good similarity techniques will result a high quality filtered items[4].

Collaborative technique

This is the most effective technique in offline Recommendation System(ORS). Collaborative technique tries to suggest products for target consumer from consumers with similar priorities. Most typical process of CF performs similarities computation on selection of user favorites that ecommerce websites typically gather rating towards products[3-6].

Hybrid technique

This technique starts from problems of above two approaches. In case, a customer purchases product X, it's not practical to suggest an entirely varying product Y to that buyer by Content Based technique. The prediction used in this system is same as Content Based approach but focus on peer experiences.

The conventional recommended filtering techniques include Item based, User based and Model based techniques. In this technique, we predict the user's behavior towards a particular item using the measured sum of variations from mean scores of consumers that formerly rated this item and the user mean-rate. Using clusters to eliminate the number of products and end-users was suggested to restore the large computation challenge; yet, this will reduce the good-quality of recommendations. Also, partitioning products to item space will restrict the recommendations to exact types of products. Moreover, if the cluster doesn't include the popular or unpopular products, they will never be suggested to users. Search based procedure depends on discovering related popular products to the one that the consumer has purchased or graded highly. This technique works effectively by creating a search-query that looks for items with the same capabilities that the consumer has graded extremely[2-7].

The well-advised solution is to use a sampling of transactions for recommendations which will certainly influence the recommendation good-quality. Discovering popular products that impart one or more capabilities is not a personalized system. Alternatively, Amazon uses Item to item Recommended Filtering. This technique, as mentioned in [5] is scalable and generates high good-quality recommendations. Also, they provide a characteristic where you will see what is highly recommended to you and revise your filter of feedback by product-line and subject. In addition they can get access to those items along with their prior purchases to rate them.

Issues of recommendation techniques

The issues for recommendation techniques increase to two core dimensions, noted as scalability and sparsity .

1. Sparsity: As the most of the recommendation techniques are dependent on similarities calculated over the co-occurred rated set of items with large intensity of sparsity

2. Scalability: Recommendation techniques appear to be highly effective in streaming items

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

that are interesting to consumers. Yet, they will require computations that are quite expensive and grow nonlinearly with the number of items and users in a database. Thus, in order to bring recommendation techniques efficiently on the web, and are successful in featuring recommendations with satisfactory delay, refined data structures and enhanced, scalable architectures are essential[8-14].

2. PROPOSED APPROACH

Product based recommended system predicts filtered based products list to the target users according to the view of other users. Assumption of the traditional recommended systems include, "if the ratings of the web based products rated by one or more users are similar, then the rating of the other products rated by same users will also be similar". Proposed workflow is defined in the figure 2. In this proposed approach, web based products along with features are taken as dataset to predict the relevant top-most filtered list to the users. In this work flow, real time web application like e-bay is used to extract products information along with different features like hit rate, associated information, category, title, image, link etc. The entire products lists along with features are stored in a file for preprocessing. Product filtering technique is proposed to find the relevant items from the large set of items along with high dimensional features. Finally, Products recommended system is proposed to predict the sparsity value and then predict the rank based items using conditional probability and product similarity values.

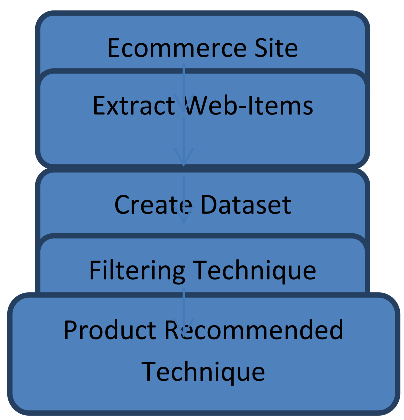


Fig 2: Proposed Workflow

Definitions:
User based item interaction matrix can be defined as "user interaction with the item along with rating".
Product based features interaction matrix is defined as "User interaction with the product along with the features list".
The general format of the product and its features matrix is tabulated below.

Features / Product	F_1	F_2	F_3	……	….	F_n
P_1	P_1F_1	P_1F_2	P_1F_3			P_1F_n
P_2	P_2F_1	P_2F_2	P_2F_3			P_2F_n
P_3	P_3F_1	P_3F_1	P_3F_3			P_3F_n
.						
.						
P_m	P_mF_1	P_mF_2	P_mF_3			P_mF_n

Each product rating can be defined by collaborating the features hit rate as:

$$P_iF_j = \begin{cases} Rate_{ij} & ;\text{exists} \\ 0; & \text{otherwise.} \end{cases}$$

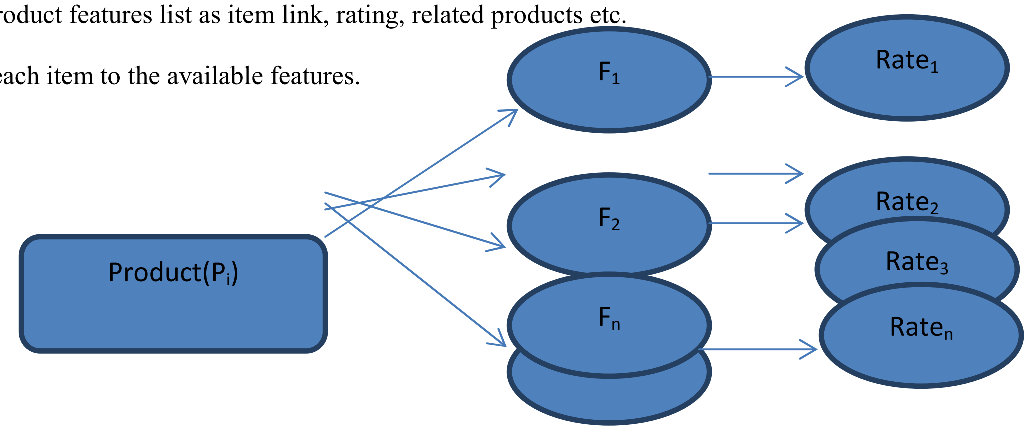
Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

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

Where P_i is the product in the web, F_j is the features list of the product and P_iF_j is the product feature rate.

$$Rate_{ij} = (Number\ of\ hits\ \text{or}\ Access\ rate\ in\ each\ category) / Total\ category\ hit\ rate\,.$$

Algorithm1:
Input: User Specified Product/Category, weight ω.
Output: Retrieves all relevant products/Category list.
Procedure:
Step 1:
Check internet connection or web access service.
Step 2:
Connect web service using the specified key pair as authentication.
Step 3:
List all the available products from the web.
For each item/product in the list.
Do
Get product features list as item link, rating, related products etc.
Done
Map each item to the available features.



Product To Feature Rate Mapping

Step 4:
Find the conditional probability to each pair of features in the product mapping.
For each pair (p_i,f_j) in the list.
Do
Conditional probability C.P
(p_i,f_j)= $(prob(p_i \cap f_j) / prob(p_i)) * \omega_j$
where ω_j is the weight of the each product feature.
Similarity of the product and its feature can be calculated using eq(2)
Sim(p_i,f_j)=

$$C.P(p_i, f_j) * (\sum_{i,j:=1} (\alpha_i - \bar{\beta}_i)(\alpha_j - \bar{\beta}_j) / \sqrt{\sum_{i:=1} (\alpha_i - \bar{\beta}_i)^2 \sum_{j:=1} (\alpha_j - \bar{\beta}_j)^2})$$

Done
Step5:
Sort product list based on product feature similarity relationship.
Step6:
Find the conditional probability to each pair of features in the product mapping.
For each pair of products in the list.
Do
Conditional probability C.P
(pf_i,pf_j)=
$(prob(pf_i \cap pf_j) / prob(pf_i)) * \omega_j$
where ω_j is the weight of the each product feature.
Similarity of the pair of products can be calculated using eq(4)
SimList(p_i,p_j)=

$$C.P(p_i, p_j) * (\sum_{i,j:=1} (\alpha_i - \bar{\beta}_i)(\alpha_j - \bar{\beta}_j) / \sqrt{\sum_{i:=1} (\alpha_i - \bar{\beta}_i)^2 \sum_{j:=1} (\alpha_j - \bar{\beta}_j)^2})$$

Done

Algorithm2:
Input :
λ : Product weight threshold
P_i: product to be predicted in the list of highest similarity list.
U_i:User Access
Output: Predicted items.

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

Procedure:

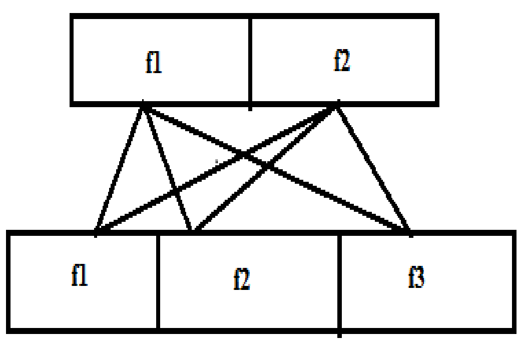
ProductKPredict(U_i, P_i, λ)
If Simlist!=null
Then
If P_i!=null
Then
λ =randomize(0,1);
List(k):=Selectkitems(Simlist)// select top k items from Simlist.
x:=0,y:=0;
for each product in List(k)
do
if C.P(P_i,F_{rate}) notexist //F_{rate} feature rating
then
x_i=x_i+$(\alpha_i - \bar{\beta}_i)$
y_i=y_i+$|\mathrm{Sim}(P_i, F_{rate})|$
end if
return x/y
else
x_i=x_i+ $\lambda * \sum_{i=1}^{n} f_{ir} / N$
y_i=y_i+ $\lambda * |\mathrm{Sim}(P_i, F_{rate})|$
return x/y
end if
// calculating items similarity between product features.

	F_1	F_2	.	.	F_n
P_1					
P_2					
.					
.					
P_n					

C.P($P_{i,}$,F_i)

	F_1	F_2	.	.	F_n
P_1					
P_2					
.					
.					
P_n					

Sim($P_{i,}$,F_i)



Product 1 And Product 2 Interaction

//Comparing each product similarity with each recommended product.
Top item in the list is predicted using the equation (6) as

$$TopSim(i, j) := \psi * C.P(i, j) + (1-\psi) * sim(i, j)$$

Where ψ is correlation coefficient between the two product features.

Predicted Probability:

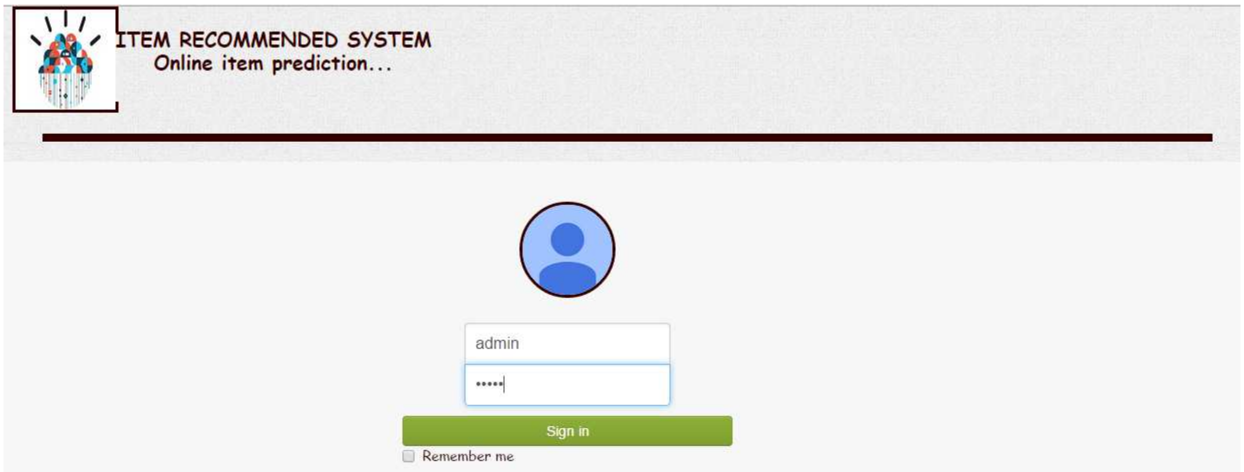
Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

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

$$\Pr ob(P_k, i) := \sum [TopSim(i,k) * \mathrm{C.P}(\mathrm{P}_r, k)]$$

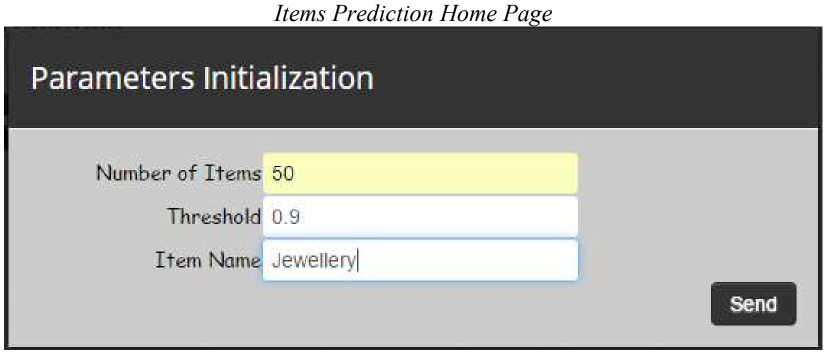
P_k is the k recommended products ,i=1…n features
Finally sort the list of items according to the probability values.

3. EXPERIMENTAL RESULTS:



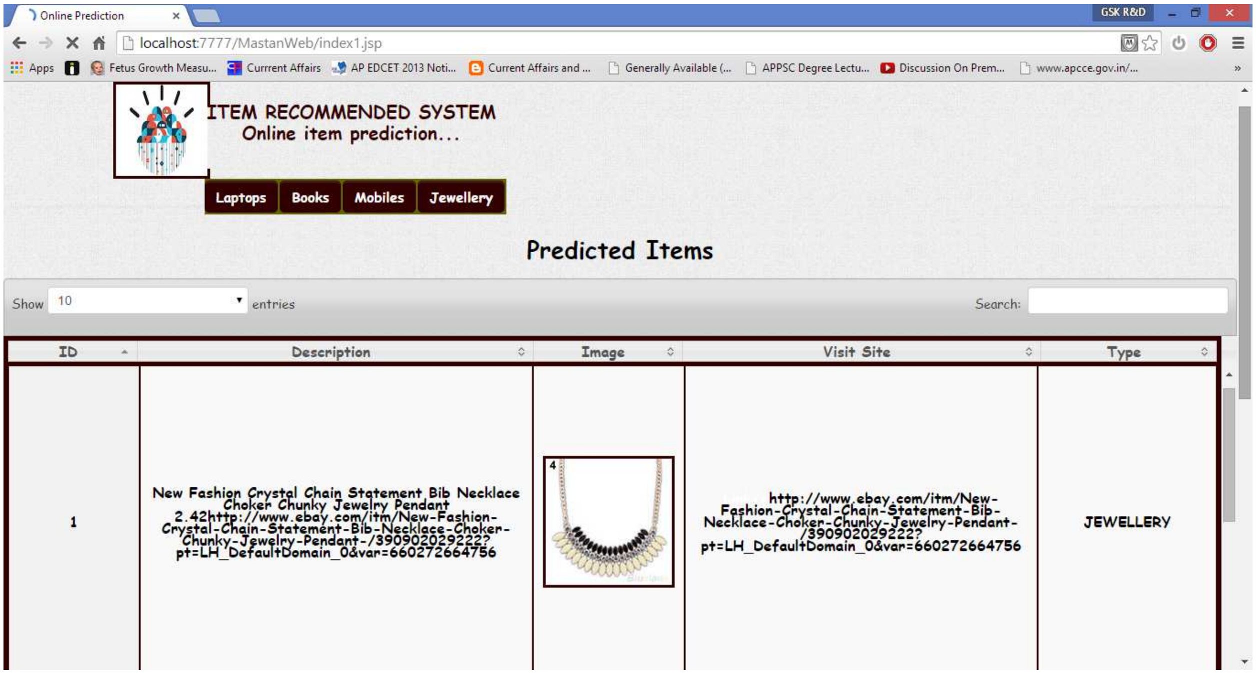
Item Prediction Login Page

Items Prediction Home Page

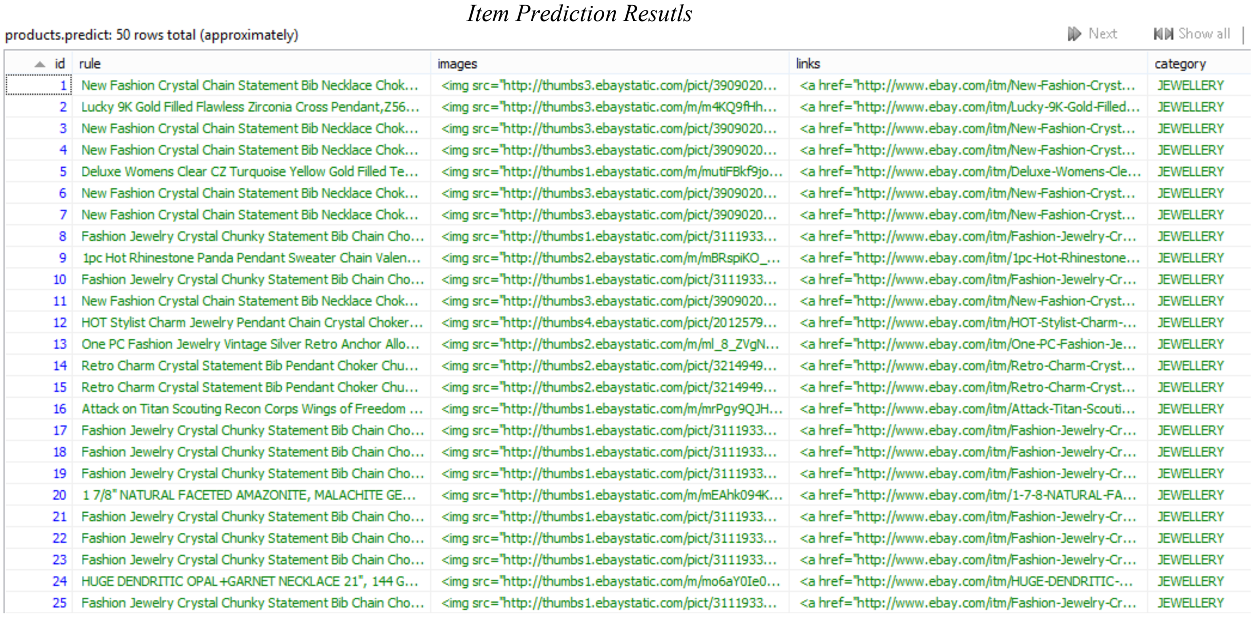


Parameters Setting For Item Prediction

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



Item Prediction Results


products.predict: 50 rows total (approximately)

id	rule	images	links	category
1	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
2	Lucky 9K Gold Filled Flawless Zirconia Cross Pendant,Z56...	<img src="http://thumbs3.ebaystatic.com/m/m4KQ9fHh...	<a href="http://www.ebay.com/itm/Lucky-9K-Gold-Filled...	JEWELLERY
3	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
4	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
5	Deluxe Womens Clear CZ Turquoise Yellow Gold Filled Te...	<img src="http://thumbs1.ebaystatic.com/m/mutiFBkf9jo...	<a href="http://www.ebay.com/itm/Deluxe-Womens-Cle...	JEWELLERY
6	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
7	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
8	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
9	1pc Hot Rhinestone Panda Pendant Sweater Chain Valen...	<img src="http://thumbs2.ebaystatic.com/m/mBRspiKO_...	<a href="http://www.ebay.com/itm/1pc-Hot-Rhinestone...	JEWELLERY
10	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
11	New Fashion Crystal Chain Statement Bib Necklace Chok...	<img src="http://thumbs3.ebaystatic.com/pict/3909020...	<a href="http://www.ebay.com/itm/New-Fashion-Cryst...	JEWELLERY
12	HOT Stylist Charm Jewelry Pendant Chain Crystal Choker...	<img src="http://thumbs4.ebaystatic.com/pict/2012579...	<a href="http://www.ebay.com/itm/HOT-Stylist-Charm-...	JEWELLERY
13	One PC Fashion Jewelry Vintage Silver Retro Anchor Allo...	<img src="http://thumbs2.ebaystatic.com/m/ml_8_ZVgN...	<a href="http://www.ebay.com/itm/One-PC-Fashion-Je...	JEWELLERY
14	Retro Charm Crystal Statement Bib Pendant Choker Chu...	<img src="http://thumbs2.ebaystatic.com/pict/3214949...	<a href="http://www.ebay.com/itm/Retro-Charm-Cryst...	JEWELLERY
15	Retro Charm Crystal Statement Bib Pendant Choker Chu...	<img src="http://thumbs2.ebaystatic.com/pict/3214949...	<a href="http://www.ebay.com/itm/Retro-Charm-Cryst...	JEWELLERY
16	Attack on Titan Scouting Recon Corps Wings of Freedom ...	<img src="http://thumbs1.ebaystatic.com/m/mrPgy9QJH...	<a href="http://www.ebay.com/itm/Attack-Titan-Scouti...	JEWELLERY
17	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
18	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
19	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
20	1 7/8" NATURAL FACETED AMAZONITE, MALACHITE GE...	<img src="http://thumbs1.ebaystatic.com/m/mEAhk094K...	<a href="http://www.ebay.com/itm/1-7-8-NATURAL-FA...	JEWELLERY
21	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
22	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
23	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY
24	HUGE DENDRITIC OPAL +GARNET NECKLACE 21", 144 G...	<img src="http://thumbs1.ebaystatic.com/m/mo6aY0Ie0...	<a href="http://www.ebay.com/itm/HUGE-DENDRITIC-...	JEWELLERY
25	Fashion Jewelry Crystal Chunky Statement Bib Chain Cho...	<img src="http://thumbs1.ebaystatic.com/pict/3111933...	<a href="http://www.ebay.com/itm/Fashion-Jewelry-Cr...	JEWELLERY

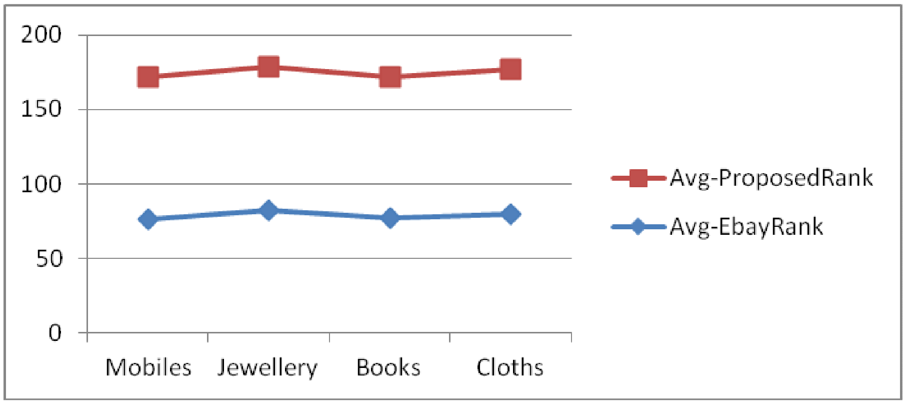
Performance Results:

Table 1: Average Ranking Of Items In Proposed Approach And Ebay Website

Items	Avg-EbayRank	Avg-ProposedRank
Mobiles	76.34	95.76
Jewellery	82.45	96.32
Books	77.23	94.54
Cloths	79.34	97.45

Journal of Theoretical and Applied Information Technology
20th November 2015. Vol.81. No.2

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



4. CONCLUSION

In this work, EBay web services are used to analyze ranking and popular items from the large set of database. .Sparsity of the product is the most significant basis which minimizes the inadequate quality. In this paper, a new product filtering and product recommended system is proposed to optimize the search space and sparsity. Product filtering can be used to filter the relevant products from the web using the conditional probability and similarity values. Product prediction based recommended system is proposed to predict the products ranking based on the products features as well as calculated predicted probabilities. Proposed approach gives better accuracy in terms of item prediction and filtering infrequent items from the large set of categorical products. Probability estimators used in this method can optimize the product importance based on the features. Experimental results on the web services gives better accuracy in terms of item prediction and accuracy factors are concerned.

REFERENCES

[1] Songjie Gong, Chongben Huang, "Employing Fuzzy Clustering to Alleviate the Sparsity Issue in Collaborative Filtering Recommendation Algorithms", 2008, pp.449-454.
[2] Xiaoyuan Su, Taghi M. Khoshgoftaar, Xingquan Zhu, and Russell Greiner." Imputation-boosted collaborative filtering using machine learning classifiers". In SAC '08: Proceedings of the 2008 ACM symposium on Applied computing, pages 949–950, New York, NY, USA, 2008. ACM.
[3] K. Lakiotaki, S. Tsafarakis, and N. Matsatsinis, "UTA-Rec: A Recommender System Based on Multiple Criteria Analysis," in Proceedings of the 2008 ACM Conference on Recommender Systems, New York, NY, USA, 2008, pp. 219–226.
[4] G. Adomavicius and A. Tuzhilin. "Personalization technologies: a process-oriented perspective."Communications of the ACM, 48(10):83--90, 2005.
[5] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins." Eigentaste: A constant time collaborative filtering algorithm" Information Retrieval, 4(2):133--151, 2001.
[6] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. "Evaluating collaborative filtering recommender systems". ACM Transactions on Information Systems (TOIS), 22(1):5--53, 2004.
[7] Xiaoyuan Su, Russell Greiner, Taghi M. Khoshgoftaar, and Xingquan Zhu. "Hybrid collaborative filtering algorithms using a mixture of experts". In Web Intelligence, pages 645–649, 2007.
[8] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl. Grouplens "applying collaborative filtering to usenet news". Communications of the ACM, 40(3):77--87, 1997.
[9] C.-H. Lai, D.-R. Liu, and C.-S. Lin." Novel personal and group-based trust models in collaborative filtering for document recommendation." Information Sciences, 239:31--49, 2013.
[10] Deepak Agarwal , Bee-Chung Chen, "Regression-based latent factor models", Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, June 28-July 01, 2009, Paris, France.

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

[11] M. Ge, C. Delgado-Battenfeld, and D. Jannach, "Beyond Accuracy: Evaluating Recommender Systems by Coverage and Serendipity," in Proceedings of the Fourth ACM Conference on Recommender Systems, New York, USA, 2010, pp. 257–260.

[12] R.-R. Liu, C.-X. Jia, T. Zhou, D. Sun, and B.-H. Wang." Personal recommendation via modified collaborative filtering" Physica A: Statistical Mechanics and its Applications, 388(4):462--468, 2009.

[13] V. Manolis and M. Konstantinos G., "On the Combination of Collaborative and Item-based Filtering," Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece.

[14] A. Umyarov and A. Tuzhilin, "Improving Rating Estimation in Recommender Systems Using Aggregation- and Variance-based Hierarchical Models," in Proceedings of the Third ACM Conference on Recommender Systems, New York, NY, USA, 2009, pp. 37–44.