

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

Prediction of Online Products by Hybrid Feedback Based Recommendation System

S.Aasha¹, Prof. Dr. A.R. Mohamed Shanavas,²

M.Phil Scholar, Department of Computer Science, Jamal Mohamed College, Trichy, India¹

Associate Professor, Department of Computer Science, Jamal Mohamed College, Trichy, India²

ABSTRACT: Online shopping has been known for a rapidly growing business, although online mobile shopping has not followed these same growth patterns in the past, but it is now being recognized for its potential. Recommender systems help the customers to find accurate product from a large database. For recommendation, admin train a database which has sentiment based keywords with positivity or negativity weight, so that when the user post comments for a product, these sentimental keywords are mined from user's comments and compare with the database. Based on the mining, if the user comment is matched with the positive keywords in database then admin will recommend the product for users. We use Hybrid filtering technique in recommendation system with feedback analysis to improve the recommendation system. These feedbacks include reviews, ratings and emoticons which are implemented by stochastic learning algorithm to enhance the sentimental analysis of a product more precisely for customer. For feedback analysis we build a dataset based on a survey of online mobile products comments from various shopping site. We also proposed to eliminate fake reviews of a product from the online site. This can be achieved by MAC address, where user can post comments only one time for the purchased product along with MAC address. This helps the customer to come out with genuine reviews of the product.

KEYWORDS: Recommendation System, Hybrid Filtering technique, Opinion Mining, Fake reviews, MAC address.

I. INTRODUCTION

The technological landscape is changing at a fast pace. Nowadays Mobile phones are becoming friends for humans. 90% of population have mobile phones in developed countries. Within a decade due to trend development of cell phone which is much smarter (called as Smartphone) Smartphone's have become an inseparable part of our daily lives. These Smartphone market has grown increasingly, not only in the conventional sales but has penetrated in the online shop. Nowadays more people prefer to buy the products online. Also numerous products are newly introduced in the market, Consumers generally check online reviews before purchase a product, it also helps the manufacturers to enhance the quality of the product. For purchasing right product they have to read all the reviews available for that product, since it is difficult for the users to read all the reviews. But if the reviews are not taken in consideration then user cannot buy a quality product.

Michael Jahrer, et.al.,... [1] Proposed recommender systems which help users to discover items within large web shops, to navigate through portals or to find friends with similar interests. The most interesting applications for recommender systems have thousands of users which generate huge amounts of data. For example, online shops collect purchase data and provide each user with a personalized shopping page on the login. The sources of information used for the recommender system can be widespread. Users generate events like the purchase of a product, rating a product, creating a bookmark or clicking on a specific item. Independently of the area of application or the type of information used, it is a major goal to increase the accuracy while retaining the capability of being able to use big datasets. Generating more accurate predictions is of general interest. For a subscription service like Netix, good recommendations are a key to customer loyalty. In the case of online stores better recommendations directly increase the revenue. In this article, we provide a systematic empirical analysis of different blending methods on the Netix dataset. The Netix dataset is one of the largest available benchmark datasets for collaborative filtering algorithms today. It contains about 108 ratings, collected in a time period of 7 years. We discuss and test several promising

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

algorithms for blending, including neural network, bagged gradient boosted decision trees, and kernel ridge regression. Our results show that linear blending is not optimal, and that it can be significantly outperformed by the presented methods. These methods are not limited to blending collaborative filtering predictors; they can be used for supervised regression problems in general.

Yu Zhang, et al., ... [2] implemented Collaborative filtering (CF) which is an effective recommendation approach based on the intuitive idea that the preference of a user can be predicted by exploiting the information about other users which share similar interests. In particular, CF techniques exploit past activities of the users, such as their transaction history or product satisfaction expressed in ratings, to predict the future activities of the users. In recent years, CF-based recommendation systems have become increasingly popular because it is generally much easier to collect the past activities of users than their profiles, partially due to privacy considerations. Collaborative filtering is an effective recommendation approach in which the preference of a user on an item is predicted based on the preferences of other users with similar interests. A big challenge in using collaborative filtering methods is the data sparsity problem which often arises because each user typically only rates very few items and hence the rating matrix is extremely sparse. In this paper, we address this problem by considering multiple collaborative filtering tasks in different domains simultaneously and exploiting the relationships between domains. We refer to it as a multi-domain collaborative filtering (MCF) problem. To solve the MCF problem, we propose a probabilistic framework which uses probabilistic matrix factorization to model the rating problem in each domain and allows the knowledge to be adaptively transferred across different domains by automatically learning the correlation between domains. We also introduce the link function for different domains to correct their biases.

Raghunandan Keshavan, et al., ... [3] studied a low complexity algorithm, based on a combination of spectral techniques and manifold optimization, that we call here OPTSPACE. We prove performance guarantees that are order-optimal in a number of circumstances. Collaborative filtering was studied from a graphical models perspective which introduced an approach to prediction based on Restricted Boltzmann Machines (RBM). Exact learning of the model parameters is intractable for such models, but the authors studied the performances of a contrastive divergence, which computes an approximate gradient of the likelihood function, and uses it to optimize the likelihood locally. Based on empirical evidence, it was argued that RBM's have several advantages over spectral methods for collaborative filtering. An objective function analogous to the one used in the present paper was considered early on in Srebro and Jaakkola, which uses gradient descent in the factors to minimize a weighted sum of square residuals. Salakhutdinov and Mnih justified the use of such an objective function by deriving it as the (negative) log-posterior of an appropriate probabilistic model. This approach naturally leads to the use of quadratic regularization in the factors. Again, gradient descent in the factors was used to perform the optimization. Also, this paper introduced a logistic mapping between the low-rank matrix and the recorded ratings.

Morgan Harvey, et al., ... [4] implemented content filtering systems, based on techniques from information retrieval, are designed to assist in this process by narrowing down the number of items a user has to look through in order to fulfill a particular information need. These systems rely on textual descriptions of items and seek to match these descriptions with a user's profile in order to suggest useful items. One significant issue with this content-based filtering is that for some types of items it can be extremely difficult to choose suitable descriptive terms to search for. Another, more accurate, approach to discovering items of interest is provided by ratings-based collaborative filtering systems, which use past ratings to predict items the user may like. Such systems predict which items a given user will be interested in based on the information provided in their user profile. These profiles consist of votes or ratings for items in the system that the user has already viewed and evaluated. The profiles of other users are frequently also exploited to improve predictions for the target user. Profiles are generally constructed explicitly from user ratings, however they may also be compiled implicitly by considering a user's purchase or bookmark history. Explicit ratings systems are commonly found on movie and music recommendation sites such as MovieLens or imdb where users can give each item a rating from 0 to 5 stars. Zero indicates that the user strongly dislikes the item and five indicates that they really like the item, however any discrete set of values could be used. Implicit systems can also be used, for example in online retail stores such as Amazon where users purchase items or add them to a "wish list", indicating that they are interested in that kind of item.

HaoMa, et al., ... [5] provided the process of trust generation is a unilateral action that does not require user to confirm the relationship. This also indicates that user does not need to even know user in the real life. "Social friendships" refer to the cooperative and mutual relationships that surround us, such as classmates, colleagues, or

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

relatives, etc. Lots of social networking Web sites, like Facebook and Orkut, are designed for online users to interact and connect with their friends in the real life. From the definition, we can see that trust-aware recommender systems cannot represent the concept of “social recommendation”, since the idea of “social recommendation” anticipates to improve recommender systems by incorporating a social friend network. Secondly, trust-aware recommender systems are based on the assumption that users have similar tastes with other users they trust. This hypothesis may not always be true in social recommender systems since the tastes of one user’s friends may vary significantly. Some friends may share similar favors with this user while other friends may have totally different tastes. Hence, trust-aware recommendation algorithms cannot be directly applied to generate recommendations in social recommender systems. Thirdly, due to the rapid growth of Web 2.0 applications, online users spend more and more time on social network related applications since interacting with real friends is the most attractive activity on the Web. On the contrary, only few online systems, like Epinions, have implementations of trust mechanism. Thus, in order to provide more proactive and personalized recommendation results to online users, we should pay more attention to the research of social recommendation, in addition to the existing research of trust aware recommendation. The basic e-commerce reviews analysis is shown in fig 1.

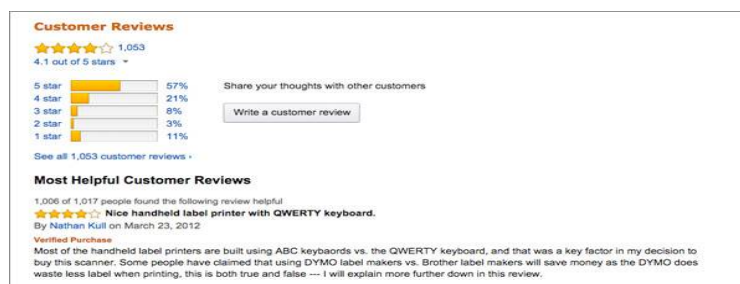


Figure1: Reviews in E-Commerce Website

Following the paper describes related work to our proposed system in Section 2, and followed by we present the proposed work in Section 3. In Section 4 we proposed the experimental results. In Section 5 provides conclusion and related future directions to this work.

II. RELATED WORK

Sentiment classification aims to overcome this problem by automatically predicting user review to be either positive or negative opinion.

Subjectivity Classification- It consists of evaluating whether a given sentence is subjective or not. It can be considered as the step that precedes sentiment classification.

Opinion summarization- It is focused on extracting the main features of an entity shared with one or several documents and the sentiments regarding them.

Opinion retrieval- The documents that express opinions or views are retrieved based on the given query. In this task, the documents are ranked as per the relevance score and the opinion score with respect to the query.

A recommendation system provides the customer to find products based on their requirements. In online shopping site there are many products to consider before purchasing and also the buyer does not have knowledge about the product. Hence Recommender system has the ability to predict whether a particular user would prefer an item or not, Decision making process and quality is improved by recommendation system. The recommender system is unique from a search engine in which, no keyword is given as input for every time to discover an items. The recommendation system can be done on the basis of buyer’s profile by which buyer’s interests can be acquired and it direct the buyers to find the items easily instead of going through large products. Hybrid filtering approaches can be implemented by combining content-based and collaborative-based to unifying the approaches into one model. This recommendation system combines multiple techniques to achieve better performance.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

III. PROPOSED FRAMEWORK

A recommendation system has implemented based on hybrid approach. We have tried to combine the existing techniques along with stochastic learning algorithm for recommendation to come up with a hybrid one. It improves the performance by overcoming the drawbacks of traditional recommendation systems. Among various recommendation approaches, collaborative filtering technique is very effective based on its high popularity. For recommendation of an item to the user, first we have to identify the relationships among various items, and then we use these relationships in order to compute the recommendations for the user. Using cosine similarity which is a similarity weight is going to play an important role in the collaborative item based filtering approach in order to select the trustable users from the given set of user. In the present methodology we are using adjusted similarity for computation of similar weights of an item. Sentiment analysis for a natural language is a large process and also it is a growing field. In the proposed framework we implement stochastic learning algorithm to analyse the feedback, where these feedbacks include reviews, ratings, and emoticons. The proposed approach can be classified into following steps:

1. Framework Construction
2. Dataset Collection
3. Sentiment analysis
4. Recommendation system
5. Fake reviews monitoring

1. Framework construction:

E-commerce framework is used to shop the products from online. This module is used to create web site for recommending best mobiles in online site. Admin is the responsibility for maintaining all products in the server. We created two accounts such as admin login and user login. In admin account, admin can login to the system and he can add a product and also he can post about product description. In User account they can login to shop the mobile products. By logging in user can view all the products along with the recommended product. In this module where admin can login to the system and post products along with its features and user can login to the system to view products in detail.

2. Dataset collection:

Datasets are collected on survey of various online shopping sites. For a product recommendation, admin will train a database which has sentiment based keywords. These keywords are trained such as positive sentiment it includes good, like, not bad, best, awesome, excellent etc. In negative sentiment it includes keywords such as bad, not like, waste and poor so on. These sentimental words are acquired from different user's comments from online site.

Opinion mining is used to predict the product is good or bad to shop in online by the users. For this admin will analyse the user feedbacks, these feedbacks include three types of comments such as ratings, reviews and emoticons (which is known as smiley's). The rating feedback has 5 stars, depending upon the satisfaction of a purchased product, user will give star rating according to it. In review feedback, it may contain user expressed textual reviews which describe about the product whether it is good or bad. Emoticon's helps us to add tone and clarity to our communication. Emoticons are used to reduce explanations. Due to social media has high growth, these picture representations of users feelings are playing an important role in communication. These Emoticons feedbacks express the user's feeling which help to conclude the product. These three feedbacks are collected and combined for sentimental analysis evaluation about a product, Also these feedbacks datasets are collected from online site based on the mobile product.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

3. Sentiment analysis:

Sentiment analysis aims to evaluate user review for a product to be either positive or negative classification. These predictions are done on the basis of text mining. To achieve the opinion about a product where admin train a database that has sentiment based keywords with positivity or negativity weight, so that when the user post comments for a product, these sentimental keywords are mined from user's comments and compare with database. Based on the mining it is easy to predict opinion about the product is good, bad, or worst for shopping. Admin recommend the products based on this sentimental analysis.

4. Recommendation system:

A recommendation system is used for information filtering which provide user's requirement based on their preferences. To find a product, User can search for a product in search bar, and it display all the items related to the users search from that user can view the description of a product such as price and review details. The stochastic learning algorithm is implemented to classify the products such as positive or negative. Positive products are displayed in recommendation panel based on ratings and reviews for the registered users. If the product has negative review means, automatically the product will not be recommended.

PSEUDOCODE

Aim: Stochastic algorithm is implemented on the user feedback for accurate analysis to predict whether product is positive or negative.

Input: A review collection D, it's attached ratings R, its attached emoticons E, a user set U and item set V

Output: Recommend the positive mobile products for trusted users

Step 1: Read ratings, reviews and emoticons of each products

Step 2: If the rating is greater than 4 stars, it is considered as positive, otherwise it is negative

Step 3: Read the words from reviews datasets, Consider words as keywords

Step 4: Match the keywords with training datasets

Step 5: Based on training words labelled the review as "positive" if it belongs to positive keyword or else labelled the review as negative.

Step 6: Read the emoticons from the datasets

Step 7: If the emoticon symbol is happy then it is labelled as positive, otherwise labelled as negative

Step 8: By combining rating value, text review and emoticons labels, Feedbacks are stated as either positive product or negative for the product.

Step 9: Update each user reviews for each item sets

Step 10: Recommend positive label products

To calculate the polarity of the review, the polarity of each individual sentence needs to be calculated. Aggregation is finding out the polarity of each review to conclude if it falls in the positive class or negative class. However, to find out the overall response about the product, an evaluation of all the reviews is required.

5. Fake reviews monitoring:

In this module, fake reviews are analysed by admin. A MAC address which is known as media access control address of a computer is a unique identifier assigned to network interfaces for communications at the data link layer of a network segment. Admin can get user account details, Mobile address and Order id details. Based on user booking ID and MAC address one user can post only one time reviews which will be genuine reviews.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

IV. EXPERIMENTAL RESULTS

In experimental results we design our E-Commerce website using C#.NET framework as front end and SQL SERVER as back end.

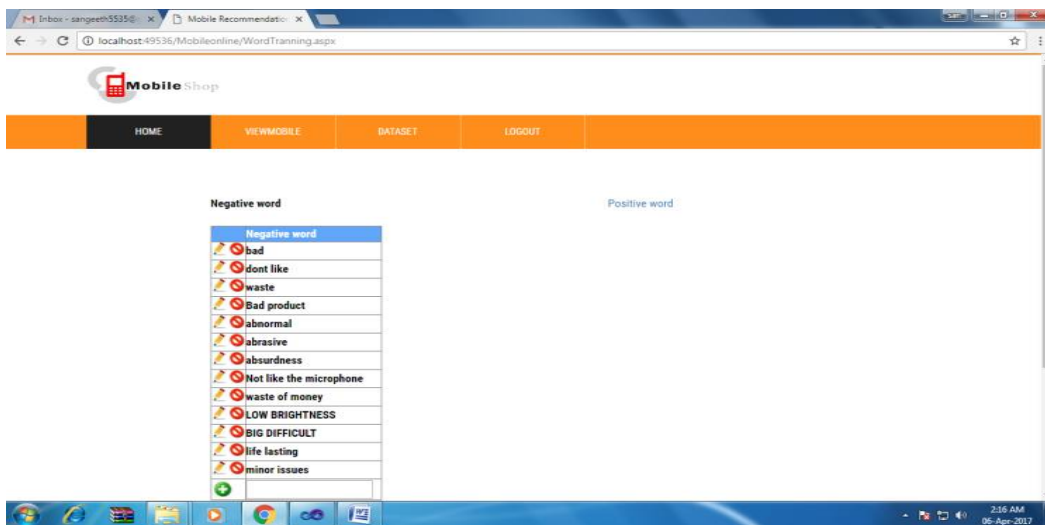


Fig 2. Words training

In Figure 2 it shows how the positive and negative words are trained in admin side such as positive sentiments includes good, like, better, best, not bad and negative sentiments includes bad, not like, waste and worst, disgusting so on.

Booking ID	Brand name	Rating	Review	Emoticon	MAC ADDRESS
BOOKID001	Nokia	5	Good product	😊	002511498562
BOOKID002	Samsung	4	Bad	😬	002511498761
BOOKID003	Iphone	3	Like	😄	002511508634
BOOKID004	Lenova	2	Not like	😡	002511708963
BOOKID005	Micormax	4	Better	😊	002511808271

Table 1: Product Purchased user dataset

In Table.1 it shows the details of purchased mobile products which include brand name, rating, reviews and emoticon with MAC address details, by this table we give feedback to analyse whether the product is good or bad to purchase.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

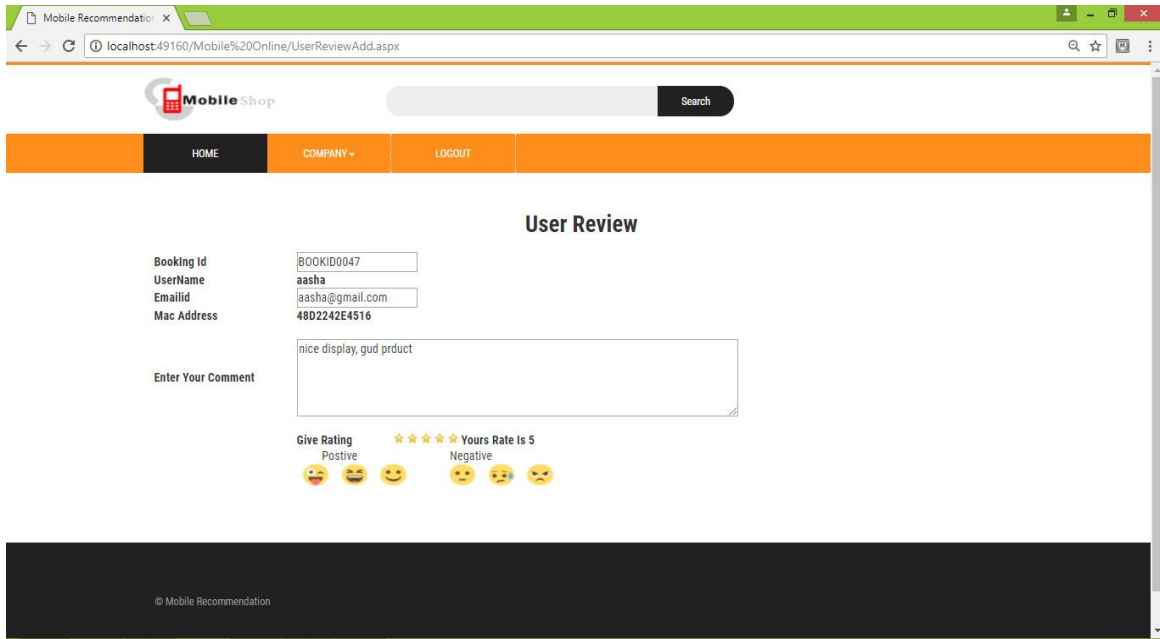


Fig 3: Purchased user reviews are shown

In Figure 3 it shows the user reviews for purchased products. This feedback includes booking id, user name, email id and MAC address, textual comments and emoticons.

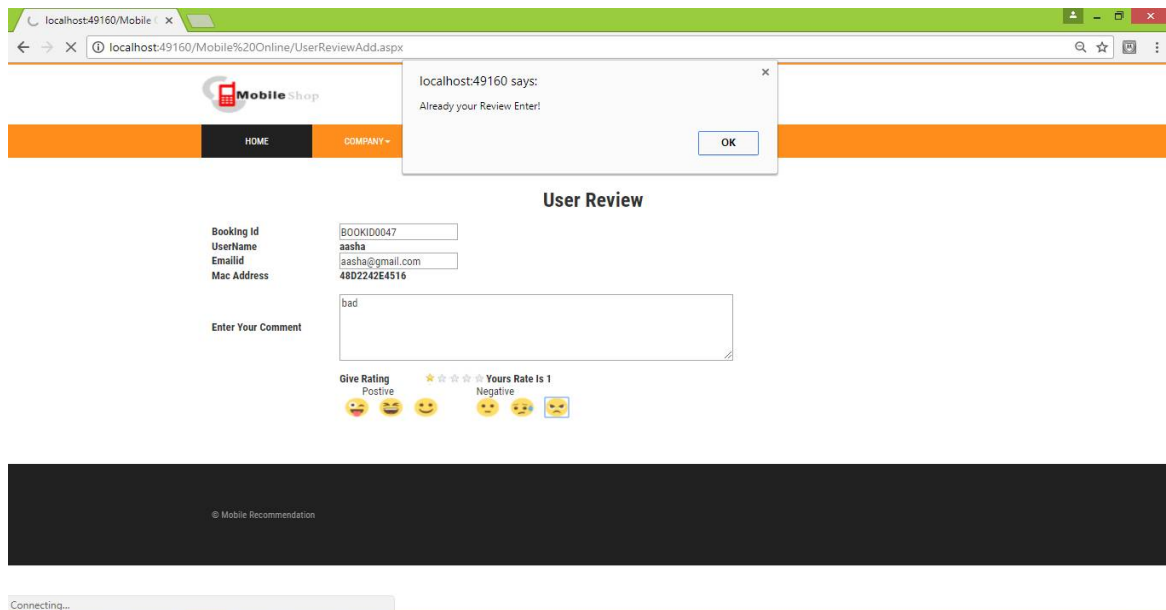


Fig 4: Fake reviews analysis

In Figure.4, Filtering of fake reviews is based on MAC address, booking id. The admin can take responsibility to filter fake reviews in online shopping site.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

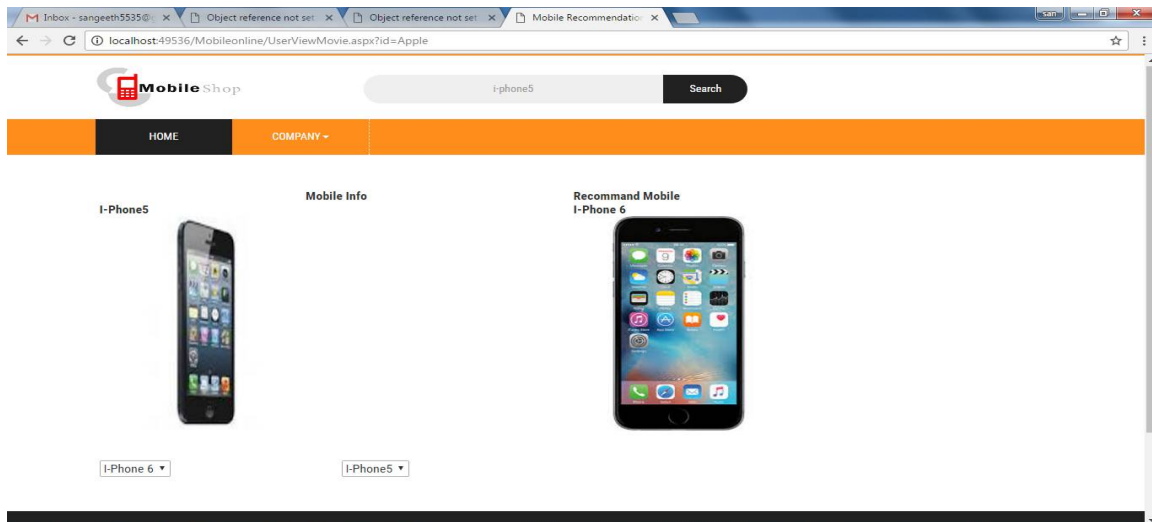


Fig 5: Recommended products by admin

In Fig.5, This datasets are trained by the admin, based on that analysis, each user feedback predict that the product is positive or negative, From this conclusion admin will recommend only positive products for registered users.

There are a number of evaluation metrics are used to evaluate the retrieval performance. For metric evaluation, we are using precision and recall. Where, precision measure can be calculate by the availability of relevant products from product database based on feedbacks and recall measure can be calculate by the availability of relevant products from the overall database.

$$\text{Precision} = \frac{\text{No of relevant products extracted}}{\text{Total no of products extracted}}$$

$$\text{Recall} = \frac{\text{No of relevant products extracted}}{\text{Total no of products in database}}$$

To analyse the recommendation system, we took mobile products in e-commerce based sentiments. Then calculate average precision and average recall for every sentiment. Result shown for proposed framework gives better result in comparison of existing framework. The performance chart is shown in Fig 6. The proposed stochastic learning algorithm provides improved results for product recommendation.

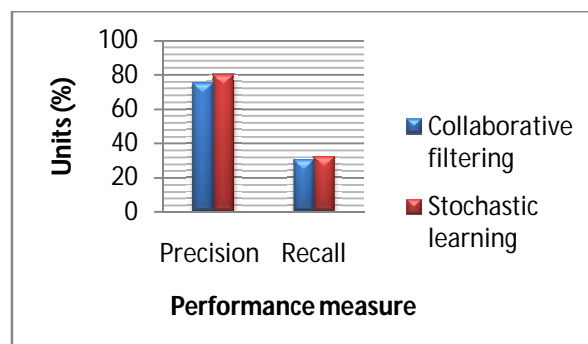


Fig 6: Performance chart

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

V. CONCLUSION

In this proposed work, we have presented a novel implementation of a mobile product recommendation based on hybrid filtering with feedback analysis to improve the recommendation system. This feedback includes Ratings, reviews and emoticons are analysed and categorized the product such as positive or negative. MAC based filtering approach can be used to avoid fake reviews to be posted. Our method was evaluated by real user data collected through an online website.

In future we can extend the work by consider recommendations not only based on user's interest also based on the user's age and lifestyle, we can also implement various mobile related products for recommendations. To enhance the sentimental classification, we can take whole phrases instead of considering keyword for sentimental classification by which it clearly convey the accurate sentiments. Another challenging task in opinion mining is each user expresses their feedback in different language such as English, Arabic, and French etc. To analyse aspect keyword from the user feedback is quite difficult for many languages. This can be implemented in future enhancement.

REFERENCES

- [1] Jahrer, M., Toscher, A., Legenstein, R., "Combining predictions for accurate recommender systems", In KDD'2010, pp. 693-702, 2010.
- [2] Zhang, Y., Cao, B., Yeung, D.Y., "Multi-domain collaborative filtering", In: Proc. of UAI 2010, pp. 725-732, 2010.
- [3] Keshavan R. H., Montanari A., Oh S., "Matrix completion from noisy entries," Journal of Machine Learning Research, Volume 11 pp.2057-2078, July 2010.
- [4] Harvey, M., Carman, M., Ruthven, I., and Crestani, F., "Bayesian latent variable models for collaborative item rating prediction", in CIKM'11, pp. 699-708, 2011.
- [5] Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I., "Recommender systems with social regularization", In: WSDM, pp. 287-296, 2011.
- [6] Bedi, P., Kaur, H., and Marwaha, S., "Trust based recommender system for semantic web", In Proc. of IJCAI'07, pages 2677-2682, 2007.
- [7] Cui, P., Wang, F., Liu, S., Ou, M., Yang, S., and Sun, L., "Who should share what?: item-level social influence prediction for users and posts ranking," in Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11), pp. 185-194, July 2011.
- [8] Scellato, S., Noulas, A., and Mascolo, C., "Exploiting place features in link prediction on location-based social networks," of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '11), pp. 1046-1054, ACM, August 2011.
- [9] Wang, Z., Sun, L., Zhu, W., Yang, S., Li, H., and Wu, D., "Joint social and content recommendation for user-generated videos in online social network," IEEE Transactions on Multimedia., vol. 15, no. 3 , pp. 698-709, 2013.
- [10] Chen, Y., Cheng, A., and Hsu, W.H., "Travel Recommendation by Mining People Attributes and Travel Group Types from Community-Contributed Photos", IEEE Transactions on Multimedia, vol. 15, no. 6, pp. 1283-1295, 2013.
- [11] Osmar R. Zaiane, Building a Recommender Agent for e-Learning Systems, in Proc. of the 7th International Conference on Computers in Education (ICCE 2002) pp 55-59, Auckland, New Zealand, December 3 - 6, 2002.
- [12] Dimitris, P., Bengt, J. Nilsson., Hollander, J., "Comparative Evaluation of Top-N Recommenders in e-Commerce: An Industrial Perspective" IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Pages: 1024 - 1031, 2015.