Marketing E-Commerce by Social media using Product Recommendations and user Embedding

V V Ramalingam¹, A .Pandian², Kirthiga Masilamani³

Department of Computer Science and Engineering, S.R.M. Institute of Science and Technology, chennai, India
ramabi1976@gmail.com

Abstract. Marketing E-Commerce by Social media is the best way to improve marketing and business widely. The major issues faced with E-commerce and Social media interfacing is cold-start cross-site problem. The cold-start problem occurs at a situation when user is not having the history of purchase records. For the user who does not have a history of purchase records, we have introduced a method of finding the users’ interested product without knowing any of the demographic information of the user. The product is recommended on bases of visits i.e., the item which is most likely to be visited by the users occur in the hit list. This product is rated at the top position for the users to purchase. The e-commerce with social media sites uses the strategy of user embedding and product recommendations. The product recommendations are achieved by incorporating Latent Dirichlet Allocation (LDA), Re Ranking and Collaborative Filtering algorithms. The proposed framework can enhance the recommendation system by embedding products and users. This shows the potential of solving cold-start cross-site problem across the e-commerce and social media sites and enhances the marketing strategy.

1. Introduction
Recently, Recommendation systems works at serving to users realize relevant and attention-grabbing things from the knowledge era, are wide studied and applied in different fields starting from e-commerce to science prediction. Suppose one user needs to buy a product in which she/he doesn’t have much knowledge or exposure, the customer needs a suggestion poll to buy the product. In this paper, we introduced a concept of ‘Rating and Recommendation’ which helps the user to go for the product. The ratings and recommendations are closely associated with the customers’ visits to the product. This in turns builds trust for the customers on the product recommendation. The cold-start problem occurs in two different situations. The user is new to the system and hence the recommendations systems generally need the historical purchase records of the user to recommend on a product. If the item is new to the database which in turns is difficult for any user to review at first. This creates a problem of cold-start cross-site where recommendations are not possible without a history of purchase records. The well-known techniques and algorithms are discussed below.

BCS model helps in utilizing the user feedbacks from the information provided from secondary sources. Most of the existing works concentrates on information which is not very readily informative. Hence the data used here are demographics data and the item is formulated mode widely and easy to access [1]. The TrustSVD model, suggests that the data sets are not only taken from the explicit data but also inherits...
from the implicit influence from both the ratings and trust. The trust density is very less when compared with the ratings. Hence this results in considering both the trust information and user ratings [2]. The BiFu technique solves the cold-start problem by using cloud computing setting. The strategy introduced here is popular items mapping and frequency raters - raters who do it frequently. It imports smoothing and fusion technique. BiFu uses social media contents from both item and users. The single clustering interlinks one clustering items to cluster item or users. Whereas in Bi-clustering it interlinks both user and item dimension. It deals with sparse and high recommendation matrices [3]. The major issues faced are cold-start problems and data sparsity. In advance to existing methods, the proposed recommender algorithm, dubbed DecRec, decouples the following two aspects of the cold-start problem to effectively exploit the side information [4]:

- The rating sub-matrix is completed, this excludes cold-start users from the initially created rating matrix; and
- The transfer of knowledge from existing rating to cold-start items/users.

Granular Association approach is a new relational data mining approach which is used to extract the pattern from multiple tables. Experimental results show that:

- Discretization is an effective preprocessing technique in mining stronger rules
- The appropriate settings of interval numbers are critical to obtaining more rules
- The equal frequency approach outperforms the equal width and the k-means approaches
- The recommendation accuracy and the number of recommendations are improved significantly through the Discretization approaches [5].

The case based reasoning cycle is a process comprising of retrieve, reuse, revise, retain. This retrieves data from the case based on the most similar cases and uses their solutions to the problem to find a result [6]. Clustered Multilayer Networks assumes that information propagates among multiple networks. There are few difficulties faced by the clustering:

- Clustering makes more tedious for a single user to transfer information to masses
- Clustering will reduce the total fraction of individuals’ information [7].

Hashing is used to perform high-speed search with very less storage and expense. To generate hash code, Collective Matrix Factorization Hashing (CMFH) is used. Here the nearest neighbor is retrieved for a query in a data collection. The hash model can deal with a data or feature in a single type (text, image etc.). This works on connection between different modalities [8]. Most of existing e-commerce suggested systems aim to recommend the proper item to a user, supported whether the user is probably going to buy or sort of a product. On the other hand, the potential of recommendations depends on the time of the advice. Allow us to take a user World Health Organization simply purchased a laptop computer as an example. She might purchase a replacement battery in a pair of years (if the laptop computer's original battery typically fails to figure around that time) and get a brand-new laptop in another a pair of years. During this case, it's not a decent plan to suggest a brand-new laptop computer or an advanced laptop right when the user purchased the new laptop computer. It may hurt the user's satisfaction of the recommender system if she receives a doubtless right product recommendation at the incorrect time. We tend to argue that a system mustn't solely suggest the foremost relevant item, however conjointly suggest at the proper time [9].

Goering outlined a retail sales prediction and products recommendation system that was enforced for a sequence of retail stores. The relative importance of client demographic characteristics for accurately modeling the sales of every client kind square measure derived and enforced within the model. Knowledge consisted of daily sales data for 600 products at the shop level, broken out over a collection of non-overlapping client varieties. A recommender System was designed supported a quick on-line skinny Singular Worth Decomposition. It's shown that modeling knowledge at a finer level of detail by clump
across client varieties and demographics yields improved performance compared to one mixture model designed for the complete dataset [10]. Linden et al proposed recommendation algorithms area unit best glorious for his or her use on e-commerce internet sites, wherever they use input a couple of customer's interests to come up with an inventory of suggested things. Several applications use solely the things that customers purchase and expressly rate to represent their interests, however they'll additionally use alternative attributes, together with things viewed, demographic information, subject interests, and favorite artists. At Amazon.com, we tend to use recommendation algorithms to change the web store for every client. The shop radically changes supported client interests, showing programming titles to an engineer and baby toys to a replacement mother. There are unit 3 common approaches to resolution the advice problem: ancient cooperative filtering, cluster models, and search-based strategies. Here, we tend to compare these strategies with our algorithmic program that we tend to decision item-to-item cooperative filtering [11].

Zenithal discussed the underlying premise of this text is that dynamic demographics can result in a breakage of the mass markets for grocery product and supermarkets. A field study investigated the relationships between five demographic factors-sex, feminine operating standing, age, income, and matrimonial status-and a large vary of variables related to preparation for and execution of food market looking. Results indicate that the demographic teams dissent in important ways that from the standard food market shopper. Discussion centers on the ways in which dynamic demographics and family roles might influence retailers and makers of grocery product [12]. Zhao proposed a recommender system said to be ‘Breed’; this system has the following aspects:

- Breed was developed supported a micro blogging service platform. As such, it's not restricted by the knowledge obtainable in any specific e-commerce web site. Additionally, breed can trace users' purchase intents in close to time and build recommendations consequently.
- In breed, product recommendation is formed as a learning to rank drawback. Users' characteristics extracted from their social profiles in micro blogs and products' demographics learned from each on-line product reviews and micro blogs square measure fed into learning to rank algorithms for product recommendation [13].

Community question answering (CQA) is discussed by Guangyou Zhou. The questions retrieval has become the major challenge. This works on finding the recent questions solved by the users. The word ambiguity and word mismatch has become the major challenges here. The problems such as words mismatch between historical questions and queried questions have been faced. The words referred in queried and historical have two different meanings. The word-based translation model is suggested for the better performance [14]. Hashing is used to perform high-speed search with very less storage and expense. The performance of the data retrieval depends on the feature selection with different hashing methods. Also, a single feature wouldn't be helpful in hashing methods. So, they have proposed non-negative matrix factorization. NMF deals with data only with hidden features. And, it considers the data for which the relationships with the object are known [15]. The pair wise data is considered by the Gaussian functions. Here, they have applied the NMF with Multiview hashing for the first time in large-scale search. Multiple visual features from different angle are used to produce the optimized result. The results proved that the NMF outperforms the traditional ways of searching [15].

2. SYSTEM MODEL

In this paper, we have proposed a new model in which the products are be listed to the customers on basis of the users’ visits to the product. The products with the high visits are to be displayed on the top. These products are called Top K products. This attracts the customers’ eyes on choosing the product listed in top
position. And, the model adds a feature of recommending the product to the friends group. We have implemented it by mapping the users’ who have account with social networking sites and e-commerce sites.

The system model is depicted with the below modules:

2.1 Product Embedding Module

The product embedding module is used to embed the product with the 'Product ID' which is used as a word token. The product id records the visits respective with the product and the user. The user visits are captured by this method. The product id is being recorded in the database along with the visits done on it. The product id key is increased by one, which makes the product to look like the most visited product of the customer. As the count mapped with the product ID increases, the one with the highest count is displayed in the top position.

2.2 User Embedding Module

The user embeddings in a similar way that we do for product embedding's. The user id is used as the token as such as product id. The visit list is considered to as a phrase consisting of a sequence of user ids and product ids. The user is embedded with the product that he has visited, this in turn create a product top list which is displayed at the top based on the visits and purchase.

2.3 Product Recommendation Module

The interesting problem here is recommending products for the users who don’t have history of purchase records. This situation is called ‘cold-start’ problem. We need to link users among the commercial and social media sites. The product with the highest count is listed in the top position for the user to purchase. The product id is being used for the calculation of count that is used to display as the ‘Top K’ product. On every visit, the product id count is increased by one. So, the product with the higher visit count is made to be displayed in the top list.

Figure 1. The system architecture.
3. Feature Extraction and Selection

The main issues faced by e-commerce Recommendation system is to recommend products for the users who do not have historical purchase records. In this paper, we implemented a model where recommendation is possible for the users who do not have historical purchase records. The model is implemented with a strategy of recommending products with highest count. The count is based on the number of visits logged on the product. The product ID and user ID is used in linking product and user for recommendations. We have used the following techniques to achieve the result:

3.1 Matrix Factorization Techniques for Recommender Systems

Recommender systems rely on different types of features, such as user’s demographic information, likes, shares, user’s interest etc. These different types of input data are placed as a matrix with one dimension representing users and other dimension representing the user’s interested products. In our model, the most visited products are gathered based on the visits. The product with high rating is displayed top in the product list. The number of times a user who visits a product is considered as the high rating, the product ID is linked with the count. The count with respect to the visit is increased by one. Thus, in the product database the count is recorded and the product with the higher count is displayed at the top for the users to purchase.

We proposed recommendation system through user embedding’s and product embedding’s. This is implemented by using the algorithms below:

3.2 Latent Dirichlet Allocation

Latent Dirichlet Allocation is a powerful learning algorithm for automatically and jointly clustering words into "topics" and documents into mixtures of topics. For instance, a topic in a collection of newswire might include words about "sports", such as "baseball", "home run", "player", and a document about steroid use in baseball might include "sports", "drugs", and "politics". Note that the labels "sports", "drugs", and "politics", are post-hoc labels assigned by a human, and that the algorithm itself only assigns associate words with probabilities.

3.3 Re Ranking Algorithm

Re Ranking algorithm is user to rank the products on based on the count. The product ID is used as an identifier for the products in the database. The user ID when visiting the specific product ID, the count linked with the product ID is increase by one. The re ranking algorithm is used to rank the products per the rank count. The product with count value greater is listed in the top most lists for the users.

3.4 Collaborative Filtering Algorithm

Collaborative filtering is a system model algorithm for making recommendations. This algorithm works on a theory that the similarities between different entities in the dataset are calculated by using the similarity measures. The similarities are selected based on the users’ preferences, actions, likes, comments, dislikes, ratings and reviews. A large data set is required for the collaborative filtering model. The rating is based on the visits the users’ make on the product. Thus, the product ID which has the highest count on visiting by users is displayed top in the product list.

4. Results and Discussions

To evaluate the effectiveness of the recommendation system, we have used the datasets from ‘http://env-4295793.cloud.cms500.com/Social/index.html’ domain. We have created user set and product set for our
The user set is created by launching the site and creating numerous user accounts. These accounts are used to purchase products. An identifier is mapped to user called User ID.

**Figure 2.** Product Details

The user in his account can add friends, communicate with friends etc., On the other hand the product DB is prepared by the admin which has the entire list of products to be listed. The admin is given the privilege to add the products in the database as well as to delete the product from the database. Each product in the database is mapped to a unique identifier as Product ID. The user embedding’s and product embedding’s done to interlink the database. The user when logs in to the system to purchase, he/she first view the list of products listed as per the requirements. Once the user makes his visit to the product, it is the first visit of the user to the specific product. The proposed system is built on a ground rule that; the recommendation system works without knowing the historical purchase records. So, the visits are recorded as the input for recommending products. The product catalog is displayed with the Product Name, Brand Name, Prize, Image, Ranks and Buy option.

If the user is interested to buy the product or to view the additional information, he can click on ‘Buy’. The detailed view of the product is displayed which includes Product Name, Image and prize. If you see the Figure 2, there is a column ‘Ranks’ to display the rank of the product. The rank is calculated based on the number of visits made by the user on the product. For every visit happening, the product in the database is incremented by one. This adds on the count of the product which is mapped with the product ID. The product ID with the highest count is listed in the top positions; these products are called as ‘Top K’ Products. This achieves in recommending a product to a user without knowing the history of purchase records.

As discussed in the section 3, the recommendation is based on the user embedding’s and product embedding’s, the user vector and product vector is mapped in to a matrix. Without having any historical purchase records of the user, the product is recommended on basis of ‘Top K products’. The product ID is mapped with the count, which is increased by one on each visit of the user to the product. Adding to the recommendation system on the product database, the user is also allowed to recommend his purchased products to the friends of their own. The product has an option of ‘Recommend To’ where he/she can recommend the product to his circle of friends. This is termed to be peer recommendations. This proves on a potential approach to solve the cold-start situation and to enhance the recommendation system.

5. Conclusion
In this paper, we have proposed a Recommendation system on solving the cold-start situations. The system is proposed on recommending products from e-commerce websites to micro blogging users who do not have historical purchase records. Our main idea is to develop a model with overcoming cold-start product recommendation. The users from e-commerce and social sites are linked with the network. The product is recommended based on the visits made by the users. By this approach the users need not to have any historical purchase records. The interested and most visited products of the users are displayed at the top of list for the users to purchase. This in turn favors users’ side on marketing. This helps them to keep track of their most visited products and enriches e-commerce. Thus, we conclude that we have proposed a recommendation system overcoming the cold-start problem.

References
1. Anupriya Gogna and Angshul MajumdarA Comprehensive Recommender System Model: Improving Accuracy for both warm and Cold Start users
2. Guibing Guo, Jie Zhang and Neil Yorke-Smith 2016 A Novel Recommendation Model Regularized with user Trust and Item Ratings IEEE Transactions On Knowledge And Data Engineering28
3. Daqiang Zhang, Ching-Hsien Hsu, Min Chen, Quan Chen, Naixue and Jaime Lloret Cold-Start Recommendation using Bi-Clustering and Fusion for Large-Scale Social Recommender Systems
4. Iman Barjasteh, Rana Forsati, Dennis Ross, Abdol-Hossein Esfahanian and Hayder Radha 2016 Cold-Start Recommendation with Provable Guarantees: A Decoupled Approach IEEE Transactions On Knowledge And Data Engineering28
5. Xu He, Fan Min and William Zhu 2014 Comparison of Discretization Approaches for Granular Association Rule MiningComparison Des Approches De Discrétisation Pour L’association Granulaire Dans l’extraction Des Données IEEE Canadian Journal of Electrical And Computer Engineering27
6. Anna Gatzioura and Miquel Sànchez-MarrëA Case-Based Recommendation Approach For Market Basket DataUniversitat Politècnica Decatalunya
7. Yong Zhuang and Osman Ya_Gan 2016 Information Propagation in Clustered Multilayer Networks IEEE Transactions on Network Science and Engineering3
8. Hashingguiguang Ding, Yuchen Guo, Jile Zhou and Yue Gao 2016 Large-Scale Cross-Modality Search Via Collective Matrix Factorization IEEE Transactions On Image Processing 25
9. Wang J and Zhang Y Opportunity Model for E-Commerce Recommendation: Right Product; Right Time
10. Giering M, Retail Sales Prediction and Item Recommendations using Customer Demographics at Store Level
11. Linden G, Smith B and York J Amazon.Com Recommendations: Item-To-Item Collaborative Filtering
12. Zeithaml V A The New Demographics and Market Fragmentation
13. Zhao W X, Guo Y, He Y, Jiang H, Wu Yand Li X We Know What You Want To Buy: A Demographic-Based System for Product Recommendation on Micro Blogs
14. Guangyou Zhou, Zhiwen Xie, Tingting He, Jun Zhao and Xiaohua Tony Hu 2016 Learning The Multilingual Translation Representations for Question Retrieval in Community Question Answering Via Non-Negative Matrix Factorization IEEE Transactions on Audio, Speech and Language Processing24
15. Li Liu, Mengyang Yu and Ling Shao 2015 Multi view Alignment Hashing for Efficient Image SearchIEEE Transactions on Image Processing24